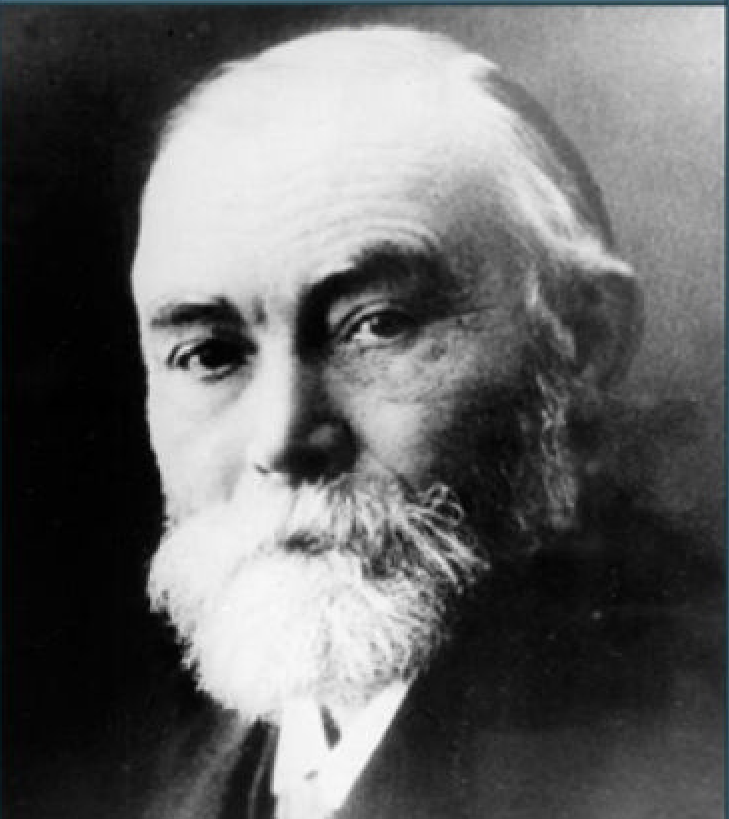
The Cambridge Companion to FREGE



EDITED BY
MICHAEL POTTER AND
TOM RICKETTS

THE CAMBRIDGE COMPANION TO
FREGE

Each volume of this series of companions to major philosophers contains specially commissioned essays by an international team of scholars together with a substantial bibliography, and will serve as a reference work for students and non-specialists. One aim of the series is to dispel the intimidation such readers often feel when faced with the work of a difficult and challenging thinker.

Gottlob Frege (1848–1925) was unquestionably one of the most important philosophers of all time. He trained as a mathematician, and his work in philosophy started as an attempt to provide an explanation of the truths of arithmetic, but in the course of this attempt he not only founded modern logic but also had to address fundamental questions in the philosophy of language and philosophical logic. He is generally seen (along with Russell and Wittgenstein) as one of the fathers of the analytic method, which dominated philosophy in English-speaking countries for most of the twentieth century. His work is studied today not just for its historical importance, but also because many of his ideas are relevant to current debates in the philosophies of logic, language, mathematics and the mind. *The Cambridge Companion to Frege* provides a route into this lively area of research.

New readers will find this the most convenient detailed guide to Frege currently available. Advanced students and specialists will find a conspectus of recent developments in the interpretation of Frege.

MICHAEL POTTER is Reader in the Philosophy of Mathematics at the University of Cambridge and a Fellow of Fitzwilliam College. He is the author of *Reason's Nearest Kin* (2000), *Set Theory and its Philosophy* (2004) and *Wittgenstein's Notes on Logic* (2009).

TOM RICKETTS is Professor of Philosophy at the University of Pittsburgh. He is the author of numerous articles on the development of analytic philosophy, especially Frege, Wittgenstein and Carnap.

OTHER VOLUMES IN THE SERIES OF CAMBRIDGE COMPANIONS

ABELARD *Edited by* JEFFREY E. BROWER and KEVIN GUILFOY
ADORNO *Edited by* THOMAS HUHN
ANSELM *Edited by* BRIAN DAVIES and BRIAN LEFTOW
AQUINAS *Edited by* NORMAN KRETZMANN and ELEONORE STUMP
ARABIC PHILOSOPHY *Edited by* PETER ADAMSON and RICHARD C. TAYLOR
HANNAH ARENDT *Edited by* DANA VILLA
ARISTOTLE *Edited by* JONATHAN BARNES
ATHEISM *Edited by* MICHAEL MARTIN
AUGUSTINE *Edited by* ELEONORE STUMP and NORMAN KRETZMANN
BACON *Edited by* MARKKU PELTONEN
BERKELEY *Edited by* KENNETH P. WINKLER
BRENTANO *Edited by* DALE JACQUETTE
CARNAP *Edited by* MICHAEL FRIEDMAN and RICHARD CREATH
CRITICAL THEORY *Edited by* FRED RUSH
DARWIN 2ND EDN *Edited by* JONATHAN HODGE and GREGORY RADICK
DARWIN'S *ORIGIN OF SPECIES* *Edited by* MICHAEL RUSE and ROBERT J. RICHARDS
SIMONE DE BEAUVOIR *Edited by* CLAUDIA CARD
DESCARTES *Edited by* JOHN COTTINGHAM
DUNS SCOTUS *Edited by* THOMAS WILLIAMS
EARLY GREEK PHILOSOPHY *Edited by* A. A. LONG
EARLY MODERN PHILOSOPHY *Edited by* DONALD RUTHERFORD
FEMINISM IN PHILOSOPHY *Edited by* MIRANDA FRICKER and JENNIFER HORNSBY
FOUCAULT 2ND EDN *Edited by* GARY GUTTING
FREUD *Edited by* JEROME NEU
GADAMER *Edited by* ROBERT J. DOSTAL
GALEN *Edited by* R. J. HANKINSON
GALILEO *Edited by* PETER MACHAMER
GERMAN IDEALISM *Edited by* KARL AMERIKS

Continued at the back of the book

The Cambridge Companion to FREGE

Edited by

Michael Potter
University of Cambridge

and

Tom Ricketts
University of Pittsburgh

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo, Delhi, Dubai, Tokyo, Mexico City

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521624794

First published 2010

Printed in the United Kingdom at the University Press, Cambridge

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

Library of Congress Cataloguing in Publication data
The Cambridge companion to Frege / [edited by] Michael Potter, Tom Ricketts.
p. cm. – (Cambridge companions to philosophy)
Includes bibliographical references and index.
ISBN 978-0-521-62428-2 – ISBN 978-0-521-62479-4 (pbk.)
1. Frege, Gottlob, 1848–1925. I. Ricketts, Tom. II. Potter, Michael D. III. Title. IV. Series.
B3245.F24C35 2010
193–dc22 2010011242

ISBN 978-0-521-62428-2 Hardback
ISBN 978-0-521-62479-4 Paperback

CONTENTS

CONTRIBUTORS

CORA DIAMOND is University Professor and Kenan Professor of Philosophy Emerita at the University of Virginia. She is the author of *The Realistic Spirit: Wittgenstein, Philosophy, and the Mind* (1991) and the editor of *Wittgenstein's Lectures on the Foundations of Mathematics, Cambridge, 1939* (1976).

WARREN GOLDFARB is W. B. Pearson Professor of Modern Mathematics and Mathematical Logic at Harvard University and has been on the faculty there since 1975. His research interests centre on mathematical logic, the development of analytic philosophy (particularly Frege, Russell, Carnap, Quine and Wittgenstein), the interrelationships between logic and philosophy and the issues in metaphysics and philosophical logic that are at the heart of the analytic tradition. He edited Jacques Herbrand's *Logical Writings* (1971), co-authored (with Burton Dreben) *The Decision Problem: Solvable Classes of Quantificational Formulas* (1979) and co-edited Kurt Gödel's *Collected Works*, vol. III (1995), vols. IV–V (2003). His textbook *Deductive Logic* was published in 2003.

MICHAEL HALLETT has been a member of the Philosophy Department at McGill University since 1984. In that year he published *Cantorian Set Theory and Limitation of Size*, an historical analysis of Cantor's work and of the emergence of modern set theory, concentrating on the philosophical and conceptual difficulties with the power-set operation. His subsequent work centres on Frege, Hilbert and Gödel, with the geometrical and foundational work of Hilbert as its current focus. He is one of four general editors of a project to publish (in six volumes) the notes of major lecture courses Hilbert gave on the foundations of mathematics and physics between 1891 and 1933. The first volume (on geometry) appeared in 2004.

RICHARD HECK is Professor of Philosophy at Brown University, where he has taught since 2005. He has worked on several aspects of Frege's philosophy, as well as on general issues in philosophy of language, logic, mathematics and mind. He lives in Massachusetts with his wife, daughter and four cats.

PETER HYLTON was educated at Kings College, Cambridge and at Harvard University. He is currently Professor of Philosophy and Distinguished Professor at the University of Illinois, Chicago. His publications include *Russell, Idealism, and the Emergence of Analytic Philosophy* (1990), *Quine* (2007) and numerous essays on the history of analytic philosophy: some of these essays are collected in *Propositions, Functions, and Analysis* (2005).

MICHAEL KREMER is Professor of Philosophy at the University of Chicago. He received his PhD from the University of Pittsburgh in 1986 and taught at the University of Notre Dame prior to joining the University of Chicago faculty. He has published widely on early analytic philosophy, especially Frege, Russell and the early Wittgenstein, as well as philosophical logic and the philosophy of language. He is currently working on a book on Frege's sense–reference distinction.

PETER MILNE is Professor of Philosophy at the University of Stirling. He has published articles on the history and philosophy of logic, on formal epistemology and on the foundations of probability.

ALEX OLIVER is Reader in Philosophy at Cambridge University and a Fellow of Gonville and Caius College. He has written extensively on metaphysics, philosophical logic and the philosophy of mathematics.

MICHAEL POTTER is Reader in the Philosophy of Mathematics at Cambridge University and a Fellow of Fitzwilliam College. His research interests span the philosophy of mathematics and the history of early analytic philosophy. His books include *Reason's Nearest Kin* (2000), *Set Theory and Its Philosophy* (2004) and *Wittgenstein's Notes on Logic* (2009). He is also co-editor of *Mathematical Knowledge* (2007) and of a forthcoming collection on *The Tractatus and Its History*.

THOMAS RICKETTS has been a professor at the University of Pittsburgh since 2005, having previously held faculty appointments at Northwestern University, the University of Pennsylvania and Harvard University. His research interests focus on the development of analytic philosophy, especially Frege, Russell, Wittgenstein, Carnap and Quine.

PETER SULLIVAN is Professor of Philosophy at the University of Stirling. He has published articles on most of the leading figures in the early history of analytic philosophy and has a particular interest in Ramsey. He is co-editor of a forthcoming collection on *The Tractatus and Its History*.

WILLIAM W. TASCHEK is Professor of Philosophy at the Ohio State University. His principal research is in the philosophy of language and the history of early analytic philosophy, with special emphasis on Frege, Russell and the early Wittgenstein.

JOAN WEINER is the author of *Frege in Perspective* (1990) and *Frege Explained* (2004), along with numerous scholarly articles. She has been awarded fellowships by the Guggenheim Foundation, the American Philosophical Society and the Mellon Foundation, and a research grant by the National Science Foundation. She is currently Professor of Philosophy at Indiana University.

MARK WILSON works largely in the intersection between traditional philosophical concern and scientific practice; his recent book *Wandering Significance* (2008) is devoted to such topics. He teaches philosophy at the University of Pittsburgh.

PREFACE

This volume has been many years in the making. It was begun by one of us (Thomas Ricketts), who commissioned most of its chapters. The other editor (Michael Potter) joined the project at a later stage, commissioning several more chapters and seeing the volume through to press. Although it differs in some respects from what either of us might have designed on his own, we are delighted to see it published now. We are grateful to the contributors for their patience and understanding during the volume's long gestation: we hope they will judge the result worth the wait.

M.D.P.
T.R.

NOTE ON TRANSLATIONS

There are several words in German that Frege used with technical meanings. The various English translations of his work (and, as a result, the secondary literature in English) are not agreed about how to translate some of them. The following table shows the translations that have generally been used in this volume. Quotations have sometimes, where the sense allows, been silently altered to conform to these conventions.

Frege	This volume	Other translations
Sinn	Sense	Meaning
Bedeutung (post-1891)	Reference	Meaning, nominatum, denotation
Wertverlauf	Value-range	Course of values
Vorstellung	Idea or representation	
Begriffsschrift	Conceptual notation	Concept-script

The correct translation of *Bedeutung* is a matter of particular controversy for some commentators, and here complete uniformity has not been feasible.

CHRONOLOGY

8 November 1848	Frege born in Wismar
1866	Death of Frege's father from typhus
1869–71	Frege attends University of Jena
1871–4	Frege attends University of Göttingen
1874	Frege begins teaching at University of Jena
1879	Publication of *Begriffsschrift* (*Conceptual Notation*)
	Frege begins to receive a stipend from University of Jena
1884	Publication of *Grundlagen der Arithmetik* (*Foundations of Arithmetic*)
14 Mar 1887	Frege marries Margarete Lieseberg
1893	Publication of first volume of *Grundgesetze der Arithmetik* (*Basic Laws of Arithmetic*)
1896	Frege promoted to a professorship
1898	Death of Frege's mother
16 June 1902	Russell writes to Frege about the contradiction
1903	Publication of second volume of *Grundgesetze der Arithmetik*
1904	Frau Frege dies
1918	Frege retires
26 July 1925	Frege dies in Bad Kleinen

For a list of Frege's publications, please consult the bibliography at the end of this volume.

OTHER VOLUMES IN THE SERIES OF CAMBRIDGE COMPANIONS

GREEK AND ROMAN PHILOSOPHY *Edited by* DAVID SEDLEY
HABERMAS *Edited by* STEPHEN K. WHITE
HAYEK *Edited by* EDWARD FESER
HEGEL *Edited by* FREDERICK C. BEISER
HEGEL AND NINETEENTH-CENTURY PHILOSOPHY *Edited by* FREDERICK C. BEISER
HEIDEGGER 2ND EDN *Edited by* CHARLES GUIGNON
HOBBES *Edited by* TOM SORELL
HOBBES'S *LEVIATHAN* *Edited by* PATRICIA SPRINGBORG
HUME 2ND EDN *Edited by* DAVID FATE NORTON and JACQUELINE TAYLOR
HUSSERL *Edited by* BARRY SMITH and DAVID WOODRUFF SMITH
WILLIAM JAMES *Edited by* RUTH ANNA PUTNAM
KANT *Edited by* PAUL GUYER
KANT AND MODERN PHILOSOPHY *Edited by* PAUL GUYER
KEYNES *Edited by* ROGER E. BACKHOUSE and BRADLEY W. BATEMAN
KIERKEGAARD *Edited by* ALASTAIR HANNAY and GORDON DANIEL MARINO
LEIBNIZ *Edited by* NICHOLAS JOLLEY
LEVINAS *Edited by* SIMON CRITCHLEY and ROBERT BERNASCONI
LOCKE *Edited by* VERE CHAPPELL
LOCKE'S *ESSAY CONCERNING HUMAN UNDERSTANDING* *Edited by* LEX NEWMAN
LOGICAL EMPIRICISM *Edited by* ALAN RICHARDSON and THOMAS UEBEL
MAIMONIDES *Edited by* KENNETH SEESKIN
MALEBRANCHE *Edited by* STEVEN NADLER
MARX *Edited by* TERRELL CARVER
MEDIEVAL JEWISH PHILOSOPHY *Edited by* DANIEL H. FRANK and OLIVER LEAMAN
MEDIEVAL PHILOSOPHY *Edited by* A. S. MCGRADE
MERLEAU-PONTY *Edited by* TAYLOR CARMAN and MARK B. N. HANSEN

MILL *Edited by* JOHN SKORUPSKI
MONTAIGNE *Edited by* ULLRICH LANGER
NEWTON *Edited by* I. BERNARD COHEN and GEORGE E. SMITH
NIETZSCHE *Edited by* BERND MAGNUS and KATHLEEN HIGGINS
OCKHAM *Edited by* PAUL VINCENT SPADE
PASCAL *Edited by* NICHOLAS HAMMOND
PEIRCE *Edited by* CHERYL MISAK
THE PHILOSOPHY OF BIOLOGY *Edited by* DAVID L. HULL and MICHAEL RUSE
PLATO *Edited by* RICHARD KRAUT
PLATO'S *REPUBLIC Edited by* G. R. F. FERRARI
PLOTINUS *Edited by* LLOYD P. GERSON
QUINE *Edited by* ROGER F. GIBSON JR
RAWLS *Edited by* SAMUEL FREEMAN
RENAISSANCE PHILOSOPHY *Edited by* JAMES HANKINS
THOMAS REID *Edited by* TERENCE CUNEO and RENÉ VAN WOUDENBERG
ROUSSEAU *Edited by* PATRICK RILEY
BERTRAND RUSSELL *Edited by* NICHOLAS GRIFFIN
SARTRE *Edited by* CHRISTINA HOWELLS
SCHOPENHAUER *Edited by* CHRISTOPHER JANAWAY
THE SCOTTISH ENLIGHTENMENT *Edited by* ALEXANDER BROADIE
ADAM SMITH *Edited by* KNUD HAAKONSSEN
SPINOZA *Edited by* DON GARRETT
THE STOICS *Edited by* BRAD INWOOD
TOCQUEVILLE *Edited by* CHERYL B. WELCH
WITTGENSTEIN *Edited by* HANS SLUGA and DAVID STERN

MICHAEL POTTER

1 Introduction

EARLY LIFE

Frege was born in 1848 in Wismar, a small port on the Baltic coast in Mecklenberg.[1] His father, who ran a private school for girls there, died when he was eighteen, and his mother took over the running of the school in order to be able to provide for the university education of Frege and his younger brother. Frege was encouraged in this by a young teacher at his father's school called Leo Sachse. Sachse had attended university in Jena, and Frege went there too in 1869, lodging in the same room that Sachse had rented there before him. Frege's studies in Jena consisted mainly of courses in mathematics and chemistry. The only philosophy was a course on Kant's critical philosophy given by Kuno Fischer.

From Jena Frege went on to Göttingen, where he took further courses in mathematics and physics and wrote a dissertation, 'On a Geometrical Representation of Imaginary Forms in the Plane'. His only philosophy course at Göttingen was one on the philosophy of religion given by Hermann Lotze. After five semesters, Frege returned to Jena to submit a further dissertation for his venia docendi (i.e. licence to teach in the university). The title of this second dissertation was 'Methods of Calculation based on an Extension of the Concept of Quantity'. Neither dissertation exhibits more than a passing interest in logic or the philosophy of mathematics.

One of Frege's mathematics lecturers at Jena, Ernst Abbe, acted as a sort of mentor, supporting him, for instance, in his efforts to

[1] For information about Frege's life I have relied throughout this Introduction on Lothar Kreiser, *Frege: Leben, Werk, Zeit* (Hamburg: Meiner, 2001).

gain promotion. But it is hard to find anyone in Frege's education who might count as a philosophical teacher of central importance. The nearest to a direct influence is perhaps Lotze, not because of his lectures on the philosophy of religion but because he published a book on logic in 1874. Dummett has convincingly argued[2] that an undated list of seventeen numbered observations about logic which has survived in Frege's hand was written in response to reading Lotze's book; internal evidence strongly suggests that these notes are probably among the earliest of Frege's unpublished writings on logic to have survived (although perhaps not quite pre-dating *Begriffsschrift*, as Dummett suggested).[3]

In the notes, Frege makes a distinction, which was to be central to his thinking about logic throughout his career, between thoughts and ideas: a thought is something such that 'it makes sense to ask whether it is true or untrue', whereas 'associations of ideas are neither true nor untrue'. Truth is objective. As Frege puts it, '2 times 2 is 4' is true, and will continue to be so even if, as a result of Darwinian evolution, human beings were to come to assert that 2 times 2 is 5. Every truth is eternal and independent of being thought by anyone and of the psychological make-up of anyone thinking it.[4]

Frege does not yet quite say, as he would later, that the subject-matter of logic is truth, but he does say that logic 'only becomes possible with the conviction that there is a difference between truth and untruth'. Following close on this, given that truth is objective, is that logic is not a branch of psychology. 'No psychological investigation can justify the laws of logic.' But truth, which is on Frege's presentation fundamental to logic, cannot be defined. 'What true is,' he says, 'is indefinable.' Frege does not at this stage give an argument to explain *why* truth is indefinable, but he later held that any attempt to define it would inevitably be circular, because one would have to understand the definition as being *true*.

If what I have said about the dating of these notes is correct, then Frege formed some of his fundamental views about logic remarkably early. It is worth stressing, moreover, that the views just mentioned

[2] M. Dummett, 'Frege's Kernsätze zur Logik', in his *Frege and Other Philosophers* (Oxford: Clarendon Press, 1991).

[3] See Frans Hovens, 'Lotze and Frege: The dating of the "Kernsätze"', *History and Philosophy of Logic*, 18 (1997), pp. 17–31.

[4] *PW*, p. 174.

constitute a response to Lotze's book, not a summary of it. It is true, for instance, that Lotze distinguished between logic and psychology, but his reason for doing so was that logic deals with the value of our thoughts whereas psychology deals with their genesis. This is obviously rather distant from Frege's anti-psychologism, which was based on the objectivity of truth, not on its value.[5]

BEGRIFFSSCHRIFT

Frege's short book *Begriffsschrift*, which he published in 1879, marks the beginning of modern logic. The word 'Begriffsschrift' is not Frege's own, but seems to have been coined by Humboldt in 1824:[6] it is usually translated 'conceptual notation' or 'concept-script'. Here we shall call the book by its italicized German title and use the word unitalicized for the formal language it describes. The idea of a formal language is not itself new with Frege. But Frege's Begriffsschrift has a number of features that were quite new in 1879.

The 'seventeen key sentences' already show Frege treating logic as a subject whose central concern is truth, and regarding thoughts as of relevance to logic because they are what truth applies to. In the first chapter of *Begriffsschrift* ('Definition of the symbols'), Frege uses the term 'judgeable content' for what he previously called a thought. Moreover, he straightaway highlights an issue which was to remain of concern to him throughout his philosophical writings, namely that of identifying the structure of a judgeable content. Since what follows logically from

The Greeks defeated the Persians at Plataea

and what follows from

The Persians were defeated by the Greeks at Plataea

are identical, logic need not distinguish between these two propositions: they have the same judgeable content.

[5] For the view that Frege should be seen as a neo-Kantian who was heavily influenced by Lotze, see G. Gabriel, 'Frege als Neukantianer', *Kant-Studien*, 77 (1986), pp. 84–101. See also Hans Sluga, *Gottlob Frege* (London: Routledge and Kegan Paul, 1980).

[6] See M. Beaney and Erich H. Reck (eds.), *Gottlob Frege: Critical Assessments of Leading Philosophers* (London: Routledge, 2005), vol II, p. 13.

One of Frege's innovations was to introduce a sign to mark the act of judging that something is the case. The sign he used was a vertical line which he called the judgement stroke. He also made use of a horizontal line which he called the content stroke, whose purpose was to turn what follows the stroke into a judgeable content. However, it is not entirely clear what this amounts to. A charitable reader[7] might see this as an implicit recognition that anything which expresses a judgeable content is of necessity complex, and hence in need of binding into a unity before it is capable of being judged. This is at any rate something which Frege was in his later writings keen to assert. A less charitable reader might think that if I have expressed a content then that is all there is to it: if the content I have expressed is judgeable, nothing more is needed to indicate that; if it is not, then preceding it with a stroke cannot make it so.

Because in practice the vertical judgement stroke never occurs without being immediately followed by the horizontal content stroke, the combination of the two strokes inevitably came to be treated as a symbol in its own right. This is the origin of the turnstile symbol $\vdash$ that is ubiquitous in modern logic. However, it is worth stressing that this symbol, although it originated with Frege, is often now used in ways that he would not have recognized. In particular, Frege did not recognize a notion of conditional assertion, so would not have allowed the turnstile to be embedded, as in expressions such as

$$A_1, A_2, \ldots, A_n \vdash B.$$

The second major innovation which Frege's conceptual notation encapsulates – and the one for which it is nowadays renowned – is a method for expressing multiple generality. However, Frege not only provides such a notation; he also displays a firm grasp of the principles that underlie it. He is clear, for instance, that in a quantified expression such as $\forall x \exists y Rxy$ the letters 'x' and 'y' do not function like names. Frege conspicuously avoids the unfortunate usage inherited from mathematics which refers to them as variables: as he makes clear, they are not variable names but placeholders.

[7] E.g. Peter Sullivan, 'Frege's logic', in Dov M. Gabbay and John Woods (eds.), *Handbook of the History of Logic*, vol. III (Amsterdam: North-Holland, 2004), pp. 659–750.

If, in an expression (whose content need not be assertible), a simple or complex symbol occurs in one or more places and we imagine it as replaceable by another … then we call the part of the expression that shows itself invariant a function and the replaceable part its argument.[8]

Notice, incidentally, that on this account predicates are a particular kind of function, namely those derived from expressions whose content is assertible (i.e. from sentences).

Frege's choice of symbols shows awareness, too, of the desirability of notational economy. He has a sign for the universal quantifier (nowadays always notated $\forall$), but he does not also have a sign for the existential quantifier $\exists$, since $\exists x$ can easily be regarded as an abbreviation for $\sim\forall x\sim$. The same economy is evident too in his choice of propositional connectives. He has signs for negation (nowadays $\sim$) and for material implication (nowadays $\rightarrow$) but not for the other connectives, which can be defined in terms of them. He also notes explicitly that he could just as well have used negation and conjunction, although he stops just short of asserting that they are adequate to express all the others. Although he did not actually make use of the device of truth-tables in presenting his account, he might as well have done, as his presentation of the meanings of the logical connectives is explicitly truth-functional in character.

The other thing for which the *Begriffsschrift* is especially notable is the axiom system for predicate calculus contained in the second chapter ('Representation and derivation of some judgements of pure thought'). He had already in the first chapter formulated *modus ponens*

From $\vdash B \rightarrow A$ and $\vdash B$ derive $\vdash A$

as well as the quantifier rule

From $\vdash \forall x(A \rightarrow \Phi(x))$ derive $\vdash A \rightarrow \forall x \Phi(x)$.

Now he added the logical axioms, which he arranges in four groups:

$\vdash a \rightarrow (b \rightarrow a)$

$\vdash (c \rightarrow (b \rightarrow a)) \rightarrow ((c \rightarrow b) \rightarrow (c \rightarrow a))$

$\vdash (d \rightarrow (b \rightarrow a)) \rightarrow (b \rightarrow (d \rightarrow a))$

[8] *Bs*, §9.

$$\vdash (b \rightarrow a) \supset ({\sim}a \rightarrow {\sim}b)$$

$$\vdash {\sim}{\sim}a \rightarrow a$$

$$\vdash a \rightarrow {\sim}{\sim}a$$

$$\vdash c = d \rightarrow f(c) = f(d)$$

$$\vdash c = c$$

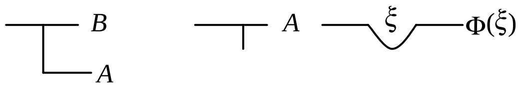
$$\vdash \forall x f(x) \rightarrow \mathrm{f}(c)$$

Frege's presentation of his axiom system is curiously understated, however. The axioms appear in the text as numbered formulae, not distinguished by any mark from the other formulae which he states as being derivable from them. He identifies his preferred axiom system only indirectly, by listing the numbers of the formulae which form what he calls the 'core' of his system.

The third chapter of *Begriffsschrift,* 'Some topics from a general theory of sequence', is a treatment of the theory of ancestral relations, expressed in the Begriffsschrift. Frege intended this chapter as an illustration of the power and elegance of his notation. There is no denying, however, that there is something unsatisfactory about the presentation. It offers its treatment of mathematical induction as an example of the ability of the Begriffsschrift to capture mathematical concepts and arguments, but then the chapter ends abruptly and in an oddly inconclusive manner. And the principle of mathematical induction itself is offered in a curiously understated way: it is not labelled as such, and the only indication in the text that this is what it is is Frege's observation that the Sorites paradox may be derived using it; mathematical induction is mentioned by name only in a laconic footnote.

RECEPTION

For all its many remarkable features, *Begriffsschrift* is undoubtedly a flawed work. One weakness, already noted, is its lack of clarity about the axiomatization of logic that it contains. Another is the rather lame presentation of the third chapter. But the feature that was of overriding importance in determining how the book would be received is one that we have not yet mentioned. In the exposition in the last section I used the symbols ~, →, ∀ for the logical constants that are now common among logicians. But Frege did not use these symbols. What we would write as $A \rightarrow B$, $\sim A$ and $\forall x\ \Phi(x)$ he wrote as



respectively. Now Frege's two-dimensional notation no doubt has its advantages. Once the eye has become used to it, it exhibits the logical structure of a complicated expression more vividly than does the bracketing of the conventional, one-dimensional alternative. But the plain fact is that it was too radical a departure from what was familiar to have any hope of adoption, and no one other than Frege ever used it. Moreover, he himself was curiously stubborn about it. A more concessive personality than his might have responded to criticism by separating out the part that is most unfamiliar (the two-dimensionality) and asking his readers to focus on his other innovations, which are independent of it.

Perhaps financial pressure contributed to Frege's decision to publish the *Begriffsschrift* when he did, despite its evident incompleteness. Not only was the University of Jena in a poor financial state, but his own position within that institution was by no means secure. He was surviving as a *Privatdozent*, financially dependent on the fees paid by his students. Since the courses Frege gave were not popular, his income was small and highly variable from semester to semester. During the academic year 1878–9, for instance, it amounted to 249 marks.

The publication of *Begriffsschrift* seems to have had the desired effect of helping Frege's career. At any rate, in 1881 the university granted him an annual stipend of 300 marks. At this time his mother moved from Wismar to Jena, and they shared a house together for some years, which may also have aided his financial position.

The preface to *Begriffsschrift* promises that a work which applies the Begriffsschrift to arithmetic is imminent. And in the summer of 1882 Frege wrote to Stumpf, a contemporary of his then working at Prague:[9] 'I have now nearly completed a book in which

[9] The letter is presented in *PMC* as being to Anton Marty, but the editors acknowledge that the addressee may well have been Stumpf, since the letter from him quoted below is evidently a reply to it.

I treat the concept of number and demonstrate that the first principles of computation which up to now have generally been regarded as unprovable axioms can be proved from definitions by means of logical laws alone.'

Frege's confidence that he could indeed derive the truths of arithmetic 'from definitions by means of logical laws alone' arose from the application of his Begriffsschrift, which, he said, 'will not let through anything that was not expressly presupposed, even if it seems so obvious that in ordinary thought we do not even notice that we are relying on it for support'.

Frege's letter also shows the first signs of what was to be a continuing theme in his life, namely his feeling that his work was not receiving the attention from others that was its due. Stumpf's reply asked Frege, presumably in response to this complaint, 'whether it would not be appropriate to explain your line of thought first in ordinary language and then – perhaps separately on another occasion … – in the Begriffsschrift: I should think that this would make for a more favourable reception of both accounts.'

GRUNDLAGEN

Frege took Stumpf's advice. His attempt to 'explain [his] line of thought in ordinary language' resulted in what many consider to be his masterpiece, *Die Grundlagen der Arithmetik* (*The Foundations of Arithmetic*). In the period when he was writing the *Grundlagen*, between 1882 and 1884, Frege was teaching part-time at the Pfeiffer Institute, a private school in Jena, and indeed he mentions a book by Grassmann that was intended for use in schools.

In the *Grundlagen* Frege criticizes various views that had been offered on the nature of numbers and of arithmetical truths, before sketching his own account. Chief among the views Frege criticizes is Kant's, that the truths of arithmetic are synthetic a priori. Frege's principal objection to Kant's view is that it does not explain the scope of arithmetic. If arithmetic were synthetic, it would depend on intuition, and all our intuitions, according to Kant, are ultimately dependent on the structure of space and time. So arithmetic, since derived from the spatio-temporal structure of reality, would be applicable only to it. Yet, Frege says, the scope of arithmetic is wider. In this respect Frege distinguished arithmetic from geometry.

The truths of geometry govern all that is spatially intuitable, whether actual or product of our fancy. The wildest visions of delirium, the boldest inventions of legend and poetry, where animals speak and stars stand still, where men are turned to stone and tress turn into men, where the drowning haul themselves up out of swamps by their own topknots – all these remain, so long as they remain intuitable, still subject to the axioms of geometry.[10]

In this respect, Frege believed, geometry differs from arithmetic. Here, he said,

we have only to try denying any one of our assumptions, and complete confusion ensues. Even to think at all seems no longer possible … The truths of arithmetic govern all that is numerable. This is the widest domain of all; for to it belongs not only the actual, not only the intuitable, but everything thinkable. Should not the laws of number, then, be connected very intimately with the laws of thought?[11]

Frege invites us to think, then, that, if arithmetic has the same range of applicability as logic itself (namely, everything thinkable), the explanation for this can only be that arithmetic is derivable from logic.

So far, though, what Frege had done was only to render this central claim plausible, not to prove it. Frege now turned to his attempt at a positive account of arithmetic as derived from logic. The first thing he did was to make the important observation that ascriptions of number do not apply to piles of stuff in the world. Before we can count, we need to know what it is we are counting. We need, that is to say, a concept. It is not the pack of cards itself that has the number fifty-two but the concept 'card in the pack'. The concept 'suit in the pack', by contrast, has the number four. This observation is no doubt obvious as soon as it is made, but to realize its importance one has only to read the confused writings of authors who failed to make it. Frege himself was probably helped to realize the point by a now-forgotten philosopher called Herbart (referred to by Frege briefly in a footnote), who said something similar, although rather less clearly, in 1825.[12]

Ascriptions of number are therefore on Frege's account second-level concepts. The first-level concept 'card in the pack' falls under the second-level concept 'having fifty-two instances'. More

[10] *Gl*, §14.

[11] *Ibid.*

[12] See D. Sullivan, 'Frege on the statement of number', *Philosophy and Phenomenological Research*, 50 (1990), pp. 595–603.

generally, the second-level concept 'having *n* instances' is called a *numerically definite quantifier*: '*F* has *n* instances' is abbreviated to $\exists^n xFx$. Frege now considers what at first sight looks like the promising proposal that we should define numbers implicitly by means of the numerically definite quantifiers, which can be defined recursively as follows:

$$\exists^0 xFx =_{Df} \sim\exists xFx;$$

$$\exists^{n+1} xFx =_{Df} \exists x(Fx \,\&\, \exists^n y(Fy \,\&\, y \neq x)).$$

From these definitions it is possible to prove various arithmetical laws using logic alone, and hence, it seems, to vindicate the logicist thesis. Indeed, Frege's presentation encourages the thought that he intends a treatment of arithmetic on something like these lines, since in the Introduction to the book he lays some stress on the injunction 'never to ask for the meaning of a word in isolation, but only in the context of a proposition'. This injunction, which is nowadays known as the Context Principle, seems on the face of it to be designed precisely to license implicit definitions such as the one just offered: the definition does not tell us explicitly what the numbers are, but allows us to eliminate numerals progressively from contexts in which they occur.

However, Frege's guiding principle in his search for an account of arithmetic was that numbers are self-subsistent objects. This principle places a constraint on the use of the Context Principle, since Frege took it as central to objecthood that there should be a principle of individuation that enables us to recognize the same object again. If we are to introduce a term to refer to an object, therefore, we must give its identity conditions: our definition must suffice to determine whether the object introduced is the same as or different from any other object already known to us.

And this creates a problem for the account in terms of numerically definite quantifiers, since the implicit definition does not suffice to determine whether the numbers are equal to other objects. For instance, it does not, to use Frege's 'crude example', determine whether Julius Caesar is a natural number. Hence, Frege thought, the account must be rejected.

Frege now turned instead to the consideration of another proposal, namely that we should derive the basic laws of arithmetic from what is often now called Hume's Principle, i.e. the principle that the number of *F*s equals the number of *G*s if and only if the *F*s and the *G*s are equinumerous. Frege had more sympathy with this proposal than with the previous one. Indeed he devoted a considerable part of the *Grundlagen* to sketching how the derivation of the basic laws of arithmetic from Hume's Principle can be achieved.

In the end, however, Frege rejected this proposal for the same reason that he rejected the earlier proposal involving numerically definite quantifers, namely that Hume's Principle, thought of as an implicit definition, fails to determine the truth conditions of mixed identity sentences, i.e. equations which have a number term on one side and a singular term of another kind on the other. It fails, that is to say, to solve the Julius Caesar problem.

Frege did not regard Hume's Principle as an entirely blind alley, however. Instead of regarding it as an implicit definition of numbers, he now suggested that we should treat it as a specification of the fundamental property of numbers that we need to recognize. Having identified this property, our task is simply to find an explicit definition of numbers which has Hume's Principle as a consequence.

Frege's suggestion[13] was that we could define the number of *F*s to be the extension of the concept of being a concept equinumerous with the concept *F*. It is easy to see that, if we define numbers in this way, Hume's Principle becomes a trivial logical consequence.

There are two obvious problems with Frege's proposal, though. The first is that it solves the Julius Caesar problem at the expense of arbitrariness. The definition has various consequences which we would surely not have regarded as properties of numbers. The second problem is that the proposal does not so much solve our problem as push it one stage back. We started by wanting to know what numbers are. Now we want to know what extensions are. Frege casually says (in a footnote), 'I assume it is known what the extension of a concept is.'[14] But this is plainly no more than a placeholder for the account of extensions of concepts which he now owes us.

'ON SENSE AND REFERENCE'

Frege's stipend was increased from 300 to 400 marks in 1885 and to 1,300 marks the following year. At this as at other points in his

[13] *Gl*, §68.
[14] *Gl*, §68n.

career, Frege had to thank Ernst Abbe, a senior member of the mathematics department who had taught him when he was a student and encouraged him thereafter. In 1887 Frege married Margarete Lieseberg, eight years his junior, who had also grown up in Wismar. They moved into a new house paid for by Frege's mother, who shared the house with them until she moved into a retirement home.

Between 1890 and 1892 Frege published three articles which marked the beginning of a new phase in his thinking about logic by outlining a more refined semantic theory. The most famous of these articles is 'On sense and reference'. Philosophy students are nowadays taught that this is where Frege introduced the important distinction between sense and reference. But if that is all they are taught, they miss the significance of the issue Frege was tackling. After all, it does not take great sophistication to distinguish between what a word means and what it refers to. Such a distinction was familiar to Sextus Empiricus, who pointed out that we may understand a word that the barbarians do not, even though they hear the word and see the object referred to.[15] Frege's concern was not with whether such a distinction can be drawn but with whether it is of relevance to logic. What he argued in 'On sense and reference' was that an account of the structure of thoughts requires us to recognize different ways in which an object may be presented to us: different senses contribute to different thoughts, and hence to different inferences that may be drawn from them. Logic, according to Frege, must distinguish between the thought that $a = b$ and the thought that $a = a$ since the former licenses inferences that the latter does not.

The issue is one that Frege had already addressed more than a decade earlier in the *Begriffsschrift*. What he had suggested there was that it is the difference between the *signs* 'a' and 'b' that explains the difference in content between '$a = a$' and '$a = b$'. But this account makes '$a = b$' express something linguistic, namely that the signs 'a' and 'b' refer to the same object. By 1892 Frege had come to think that this gets the subject matter of the thought wrong: if I learn that Hesperus is Phosphorus, what I learn is a fact about astronomy, not about language.

[15] 'Against the Professors', 8.11f, in A. A. Long and D. N. Sedley (eds.), *The Hellenistic Philosophers* (Cambridge: Cambridge University Press, 1987), pp. 195–6.

If Frege is right that an appeal to the difference between the signs does not explain the difference in content, the obvious alternative would be to appeal to a difference in the images which we attach to the signs in our minds. But that, again, would be at the wrong level: it would be in danger once more of making logic psychological. The content that we are aiming to explain – the content that is relevant to logic – is what is capable of being true or false, and this is something that can be communicated from me to you. So, on Frege's view, it cannot be anything mental, an idea, because ideas are private to whoever has them in mind. I cannot have your ideas, however hard I try, but I can grasp the same thoughts as you. When you tell me that Hesperus is Phosphorus, there is something true – a thought – that you have communicated to me. So thoughts (and, by extension, their components) cannot be mental.

SATURATED AND UNSATURATED

What we have so far is an argument, hotly contested in the modern literature on the subject, that the contribution a name makes to determining the truth conditions of declarative sentences in which it occurs cannot be explained solely by appealing to the object which the name refers to. Whether or not this argument is correct, it says nothing about the contribution that parts of the sentence other than names make to this task. In discussing *Begriffsschrift* we noted how Frege's explanation of quantification led him to the idea of removing a name from a sentence in order to obtain an expression which has various sentences, including this one, as possible instantiations. Frege there calls such an expression a function, but in his later writings he preferred to call it a *function symbol*, in order to leave the word 'function' free for whatever the symbol refers to. Frege called the function symbol *unsaturated* because it has an argument place; by contrast he called a name *saturated* because it has no argument place. (He probably borrowed the terminology from chemistry.)

Frege conceived of this distinction between saturated and unsaturated components of the sentence as correlated with a similar distinction between saturated and unsaturated components of the thought which the sentence expresses. So the sentence 'John is mortal' can be seen as made up of the name 'John' and the predicate '*x* is mortal'. Correspondingly, the thought which this sentence expresses is made up of the sense of the name 'John' and the sense of the predicate '*x* is mortal'.

The notion of composition which is in play here is highly problematic, however. On the one hand, Frege was drawn to the idea that the structure of the sentence is a guide to the structure of the thought. On the other, as we have seen, he wanted to insist that some features of the structure of the sentence may be irrelevant to logic and hence not correspond to anything in the structure of the corresponding thought. This is a tension which he never satisfactorily resolved.

But it is not just the relationship between a sentence and its sense that is problematic. There is in any case a difficulty about saying what the structure of a thought is. For if the thought is intended to capture only that part of what the sentence expresses that is relevant to logic, one might imagine that logically equivalent sentences would express the same thought. But in that case how can I fully grasp two logically equivalent sentences without realizing that they are equivalent? The notion of sense is in danger of rendering logic trivial. Wittgenstein was happy to adapt Frege's notion so as to have just that conclusion, but Frege himself was hesitant about going down this path, perhaps because what impressed him most about the polyadic logic he had invented (in contrast to the syllogistic logic of his predecessors) is precisely its lack of triviality.

Since Frege conceived of the sense of a term as the mode of presentation of what it refers to, it was natural, once he held that predicates have sense, for him to suppose that they are also capable of referring to something. The reference of a predicate Frege called a *concept*. He held that the distinction between saturated and unsaturated applies also at this level. For this reason, concepts are not objects: objects are saturated, whereas concepts are unsaturated. But concepts, although they are not objects, are nonetheless *objective*: they exist independently of us.

Concepts, for Frege, are extensional, so that, for instance, the predicates '*x* is a round square' and '*x* is a golden mountain' refer to the same concept (namely the empty one). However, he was curiously reluctant to call this relation equality, preferring to reserve this term for the relation that holds between objects.

ON CONCEPT AND OBJECT

Suppose now that we grant Frege his desire to treat concept-words as referring expressions, and consider the sentence 'Bucephalus is a horse'. The name 'Bucephalus' here refers to a particular object, namely the horse Bucephalus. And the predicate '*x* is a horse' is a referring expression as well. But what does it refer to? One might think that it refers to the concept *horse*. But it cannot. For the phrase 'the concept *horse*' is saturated: it contains no argument-place. So on Frege's account whatever this phrase refers to must be saturated too. But concepts are unsaturated. So whatever the phrase 'the *concept* horse' refers to is not a concept.

It is hard to know what to make of the bizarre conclusion that the concept *horse* is not a concept. Frege himself was surprisingly (because uncharacteristically) relaxed about it. Having derived the result, he immediately tried to downplay its significance. 'By a kind of necessity of language,' he wrote,

> my expressions, taken literally, sometimes miss my thought; I mention an object, when what I intend is a concept. I fully realize that in such cases I was relying upon a reader who would be ready to meet me halfway – who does not begrudge a pinch of salt.[16]

It was Wittgenstein, some years later, who by begrudging a pinch of salt confronted the notion that there are some things that are unsayable, not for the trivial reason that our language lacks the right words, but because they are not of the right shape to constitute thoughts that *any* language could aspire to express.

ON FUNCTION AND CONCEPT

As we have seen, there is in Frege's semantics a uniformity of treatment between names and predicates. It is a three-level view, according to which the realm of sense stands as an intermediary between language (where names and predicates reside) and the objective world (where objects and concepts do). Frege took one further step towards uniformity of treatment when he decided to treat sentences as a kind of name. The sense of a sentence he took to be the thought

[16] *CP*, p. 193.

which the sentence expresses; and the reference of the sentence he took to be its truth-value. Assimilating sentences to names in this fashion entailed also that Frege assimilated predicates to function symbols and concepts to functions: a concept is now simply a function whose possible values are truth-values.

This uniformity of treatment gave Frege's system a simplicity which he no doubt found tempting. However, it was a mistake, as Wittgenstein was the first to realize. It is in principle possible to name an object by blind pointing, but nothing analogous is available in the case of truth-values: one cannot formulate a declarative sentence without having, at least implicitly, a grasp of the distinction between truth and falsity. There is therefore a fundamental difference between a sentence, on the one hand, and a name of a truth-value, on the other. This point was, ironically, one that Wittgenstein could have got from Frege himself, since Frege had already noted that thoughts are essentially complex.

GRUNDGESETZE I

As we have seen, Frege had almost completed by the summer of 1882 a book in which he took himself to have derived the laws of arithmetic from definitions by means of logical laws alone. However, the first volume of *Grundgesetze* did not appear until 1893. The preface offers as part of the explanation for the ten-year delay changes which led him to put aside an almost completed version, but it is difficult to be sure just what these were. Perhaps it was partly the revisions to his semantic theory that caused the delay. Another reason, though, may have been the (understandable) difficulty Frege had in finding a publisher willing to subsidize the cost.

The purpose of *Grundgesetze* is twofold. First, Frege had to extend the formal system he had provided in *Begriffsschrift* so as to give an account of the notion of the extension of a concept. Second, having done this, he had to make good his claim that the basic laws of arithmetic can be derived formally from the explicit definition of numbers offered in *Grundlagen*, making use at every stage of nothing other than pure logic.

In the second of these stages, Frege's treatment was a success. The impact of its publication was muted, however, not only because of the unfamiliarity of his Begriffsschrift but also because by the time he published it Dedekind had already published his masterly treatment of arithmetic in *Was sind und was sollen die Zahlen?*

The first stage of Frege's project, in which he extended his Begriffsschrift to include the notion of an extension, also bears the marks of the revision of his semantic theory mentioned earlier. Frege's formal system now made no fundamental syntactic distinction between sentences and names. Instead, he drew that distinction only at the level of reference, by assuming that the two truth-values, True and False, are already known. He then hoped by a step-by-step procedure to secure a reference for every other well-formed term of the theory.

Since sentences are, for Frege, just a kind of name, concepts are similarly just a kind of function, and the extension of a concept is a particular instance of the general notion of the value-range of a function. That is to say, Frege now envisaged that every function has associated with it an object called its value-range. Frege assumed that this association is one-to-one: distinct functions are associated with distinct value-ranges. This assumption was the fifth of the basic laws Frege laid down in his formal system: hence it is nowadays known as Basic Law V. Although Frege stated it as a general law that applies to all functions, it is the case in which the function is a concept that is of the most significance. In this case, the law would be expressed in modern notation as

$$\{x{:}Fx\} = \{x{:}Gx\} \leftrightarrow \forall x Fx \leftrightarrow Gx.$$

Or, since Frege held that concepts are equal just in case they have the same objects falling under them, we could write it more simply as

$$\{x{:}Fx\} = \{x{:}Gx\} \leftrightarrow F = G$$

(although, as already noted, Frege himself preferred not to use the '=' sign in this way). This way of writing the law illustrates clearly that Frege's notion of the extension of a concept is no more than a level-changing operator, mapping each concept to an object. The reason Frege needed such an operator was, of course, his requirement that numbers should be objects.

CORRESPONDENCE WITH HILBERT

As we have noted, Frege made a clear distinction between arithmetic and geometry: he believed that arithmetic is derivable from logic, but he felt no temptation to disagree with Kant's view that geometry is synthetic a priori. Frege did not use the word 'axiom' to mean a hypothesis: for him, an axiom must be true. So if geometry is presented, as has been common since the time of Euclid, as an axiomatic system, it is to be understood that the axioms are being asserted as true.

In 1899 Hilbert published a book on the foundations of geometry. His purpose was to study the logical relationships between the axioms of geometry. He showed in this book that some of the axioms are independent of others. The standpoint he took throughout was that he was studying logical interrelationships between groups of propositions: whether those propositions are in fact true is, for mathematical purposes, irrelevant. His view was therefore directly opposed to Frege's.

In an exchange of letters that lasted almost a year Frege and Hilbert argued about this issue. In the course of their correspondence, Frege gradually disentangled two distinct issues, which he had at first intertwined. The first is the question of the logical relationships between propositions. Frege was slow to see that Hilbert's approach provides a way to show that one of the propositions in an axiom set is independent of the others by providing a non-standard interpretation in which this axiom is false and the others are true.

The second question is whether the terms of a theory can acquire a meaning simply through the role they play in it. Frege insisted that they cannot. In his view we must have in mind an intended interpretation which makes the axioms of the theory true. If we do not, then all we have is empty formalism, not mathematics. For Hilbert, on the other hand, the only constraint on mathematical invention is that the axioms should be consistent. In mathematics, according to him, consistency entails existence. As Frege pointed out, however, this principle does not apply outside mathematics. The notion of an omnipotent, omniscient, benevolent being may be consistent, but we do not imagine that that on its own entails the existence of one. What Hilbert owed but did not offer was an argument why his principle should apply in mathematics when it does not apply elsewhere.

The correspondence between Frege and Hilbert tailed off without a resolution when Hilbert suggested that it would be easier for Frege to come and visit Göttingen so as to continue it face to face.

(Frege did not take up the offer.) Perhaps, though, Frege's criticisms eventually had an effect. For Hilbert's later philosophy of mathematics was very different. The view he later proposed made elementary arithmetic contentful, and although he proposed a formalist attitude to more abstract parts of mathematics, that formalism was constrained by the requirement that the abstract methods be legitimated by a conservativeness proof demonstrating (in principle, at least) their eliminability.

GRUNDGESETZE II

The second volume of the *Grundgesetze* went to the typesetter in 1902, although Frege had probably written most of it before the first volume was published a decade earlier.[17] In the end, he resigned himself to paying the printing costs himself. In any case, the division into volumes does not correspond to the subject matter: part II of the book is split awkwardly between the two volumes, and part III tails off at the end of the second volume, evidently to have been continued in a third volume that never appeared.

Part III aims to extend Frege's logicism from the arithmetic of part II so as to encompass the theory of the real numbers as well. Frege's positive treatment of real numbers, which is strongly reminiscent of the account by Eudoxus which Euclid made use of in the *Elements*, is not without interest. Its principal novelty in comparison with the treatments by Dedekind and Cantor that are much more familiar to the modern reader is that Frege defines the real numbers all at once without going via a construction of the rational numbers as an intermediate step. However, Frege's account was almost entirely ignored by logicians until quite recently.

As he had done earlier in the case of arithmetic, Frege started with a critique of other accounts that had already appeared, thus clearing the way for his own. The most important part of this critique is undoubtedly the one devoted to Thomae and Heine, not because their treatments of real numbers were of special interest in themselves but because Frege saw their views as underpinned by a crude version of formalism. The fact that this version of formalism is not seriously advanced by philosophers of mathematics

[17] See M. Dummett, *Frege: Philosophy of Mathematics* (Duckworth, 1991), p. 241.

nowadays is largely due to Frege's critique. This critique has two main components.

First, Frege points out that, even if mathematics were a game like chess, with no subject matter beyond the signs themselves, there would be things one could say about the game: in chess, for instance, it can be shown that one cannot checkmate with only a king and two knights. But that is a fact about chess, not about signs. So even if formalism were correct in relation to some part of mathematics, it could not be correct in relation to the whole of it.

Second, formalism cannot offer an adequate explanation of any part of mathematics, such as arithmetic or the theory of real numbers, which purports to be applicable in reasoning about the world. This is because a position in a game does not express a thought and therefore cannot be used in reasoning at all.

Frege's criticisms of Thomae are important because formalism is a position of independent philosophical interest. When Frege turns his attention to other accounts of the real numbers that had been developed independently by Cantor, Dedekind, Weierstrass and others, the result is less happy. The tone of Frege's critique is unremittingly negative, and he is quite unable to distinguish criticisms of real substance from mere pedantic carping. Dummett is surely correct to say that

> Frege is anxious to direct at his competitors any criticism to which they lay themselves open, regardless of whether it advances his argument or not … The Frege who wrote volume II of *Grundgesetze* was a very different man from the Frege who had written *Grundlagen*: an embittered man whose concern to give a convincing exposition of his theory of the foundations of analysis was repeatedly overpowered by his desire for revenge on those who had ignored or failed to understand his work.[18]

RUSSELL'S PARADOX

The second volume of *Grundgesetze* is largely ignored by modern logicians. The only part that is referred to much is the appendix which Frege added after the volume had gone to press, in order to respond to the contradiction in his formal system which had only

[18] *Ibid.*, p. 243.

just been pointed out to him by Russell. The proof of contradiction Russell had discovered is brutally short. Suppose that we define

$$x \in y =_{\mathrm{Df}} \exists X(Xx \prod y = \{z{:}Xz\}).$$

If we let

$$a = \{x{:}\ x \notin x\},$$

then

$$\forall x\ x \in a \leftrightarrow x \notin x,$$

and so

$$a \in a \leftrightarrow a \notin a,$$

which is a contradiction. What this demonstrates is that a formal system is contradictory if it allows that every concept has an extension which is an object falling within the range of the quantifiers of the system.

Frege was forced to add an appendix to volume II in which he suggested a way of getting round the paradox. The proposal he made was to restrict Basic Law V in such a way as to block Russell's proof. But this is plainly ad hoc, and it is no great surprise that Frege's way out turns out to be contradictory as well.

What is perhaps more interesting about Frege's appendix is his refusal to countenance the obvious solution of treating extensions of concepts as objects of a different kind from the objects already referred to. Frege's response here is strongly reminiscent of his response to the Julius Caesar problem twenty years earlier. He evidently had a strong attachment to a univocal notion of object: just as in the *Grundlagen* he had been driven by a conviction that numbers are objects in this univocal sense, now he was driven by a conviction that extensions of concepts are objects too.

Frege evidently realized by 1906 that his attempt to fix his formal system in the appendix to volume II of *Grundgesetze* was a failure. 'Set theory in ruins,' he noted laconically.[19] After that, he continued to give annually at Jena his lecture course presenting the formal system of *Grundgesetze*, but now he omitted all mention

[19] *PW*, p. 176.

of the notion of the value-range of a function (or, in particular, the extension of a concept).[20]

WIDOWERHOOD

It has been popular to portray Frege after this as a broken man. But what caused his depression over the ensuing years may well have been the death of his wife in 1904 after a long illness rather than the disappointment brought on by Russell's paradox. In 1905 he took time off work for ill health; over the next decade he published hardly anything, and he corresponded with other philosophers noticeably less than he had done previously.

In 1908 Frege learned through a distant relative of the case of a brother and sister, Alfred and Toni Fuchs, then aged five and three, who had ended up in a care home as a result of parental neglect. Frege took over financial responsibility for the girl and himself fostered the boy.

Frege's health remained poor. In 1913 he once again took time off because of a 'nervous illness'. (Rather oddly, the application for leave of absence also mentions that Frege wanted to complete some of his philosophical work, which somewhat undermines the other ground for the application.) From then until his retirement he rented the upper floor of his house to a Dr Schön and his family: Frau Schön's later recollection was that Frege welcomed the sound of children about the house.

It was Frege's habit to go on holiday to Brunshaupten, a small seaside resort in Mecklenburg on the Baltic coast. It is said that he usually walked the whole way (about 200 miles) with his dog, staying at the same inns each time. On at least one occasion (in 1914) he was even there during January, when the Baltic coast is not generally at its warmest.

In his later life Frege was evidently not an outgoing person. Frau Schön described him as 'word-shy' and Carnap, who attended his lectures, memorably described how he addressed them almost entirely to the blackboard. When he taught a more advanced course, on the other hand, things were more intimate.

[20] See Erich H. Reck and Steve Awodey (eds.), *Frege's Lectures on Logic: Carnap's Student Notes, 1910–14* (Open Court, 2004).

In this small group Frege felt more at ease and thawed out a bit more. There were still no questions or discussions. But Frege occasionally made critical remarks about other conceptions, sometimes with irony and even sarcasm.[21]

Gershom Scholem, who later became the preeminent modern scholar of Jewish mysticism, attended the University of Jena in 1917–18 and became an enthusiast for Frege's conception of logic in preference to Lotze's. 'I enjoyed his unpompous manner,' he said. 'But in Jena hardly anyone took Frege seriously.'[22]

The most remarkable contact Frege made during this period, however, was with Wittgenstein. How many visits Wittgenstein made to Frege is not certain, but there were definitely at least three, one in 1911 or early 1912, one at Christmas 1912 and one at Christmas 1913; and at least one of the visits took place in Brunshaupten rather than Jena. What is clear is that they formed a bond, although it was stronger, probably, on Wittgenstein's side than on Frege's. Wittgenstein revered Frege all his life, and was enormously influenced by him.

LOGICAL INVESTIGATIONS

Frege retired at the end of 1918 on a pension of 5,000 marks. He had been in poor health again and had been given leave not to teach any courses in the preceding two years. In 1918 he received a donation from Wittgenstein, which may have helped to make his retirement possible: he sold his house in Jena (to the family to whom he had been renting the upper floor) and bought a retirement home in Bad Kleinen, a village on the edge of a lake a few miles south of his birthplace. (Before the war, Wittgenstein had made a similar donation to the Cambridge philosopher W. E. Johnson, with the aim of giving him the leisure to write up his work on logic, and it may well be that his gift to Frege had similar intentions.) At any rate Frege did indeed use his retirement to write up his *Logical Investigations* for publication. This is a series of three articles, 'Thoughts', 'Negation'

[21] Rudolf Carnap, 'Intellectual biography', in P. A. Schilpp (ed.), *The Philosophy of Rudolf Carnap* (Chicago: Open Court, 1963), p. 5.

[22] Gershom Scholem, *Walter Benjamin: The Story of a Friendship* (London: Faber, 1982), p. 66.

and 'Compound thoughts', which he published between 1918 and 1923. He also began but did not complete a fourth article in the series called 'Logical generality'. He evidently intended them to form a kind of textbook on logic. In the main, however, the *Logical Investigations* do not represent a new phase in Frege's thinking about logic. 'Thoughts' is based on a draft which he had written in 1897, and indeed the overall shape of the project is one that he had conceived even earlier.

Moreover, much of the article merely lays out views he had already expressed in his published writings. He states again, for instance, his conception of thoughts as objective, publicly graspable items capable of truth or falsity. (It is here that he introduces the phrase 'a third realm' to refer to the abstract domain, neither physical nor mental, that thoughts inhabit.) He stresses again his conception of logic as the study of the laws not of thought but of truth. He emphasizes, too, the indefinability of truth, and indeed comes very close to claiming that truth is altogether redundant. 'The sentence "I smell the scent of violets",' he observes, 'has just the same content as the sentence "It is true that I smell the scent of violets". So it seems, then, that nothing is added to the thought by my ascribing to it the property of truth.'[23]

But in 'Thoughts' Frege was clear, as he had not been before (not explicity, at least), about the fundamental difference in structure between thoughts and complexes. 'Do we not see,' he asks rhetorically, 'that the sun has risen? And do we not then also see that this is true?' No.

> That the sun has risen is not an object emitting rays that reach my eyes; it is not a visible thing like the sun itself. That the sun has risen is recognized to be true on the basis of sense-impressions. But being true is not a sensible, perceptible property.[24]

This distinction between the rising sun and the fact that the sun is rising made a great impression on Wittgenstein: it was central to his conception of the world as the totality of facts, not of things, in the *Tractatus*.

[23] *CP*, p. 354.
[24] *Ibid.*

Something else that we find in 'Thoughts' is Frege's argument against the correspondence theory of truth. Correspondence, he says, comes in degrees; truth does not. 'What is half true is untrue.' And if we were to say that truth is merely correspondence of our ideas with reality in a certain respect, we would simply have pushed the question back to whether it is *true* that our idea corresponds with reality in the specified respect. To overcome this difficulty, Frege claims, we should regard truth not as correspondence with the facts but as identity: a fact does not correspond with, but simply is, a thought that is true.

Frege then considers a difficulty for his conception of thoughts, and the senses which are their components, as intrinsically public, capable of being grasped by more than one person. If this is so, he wonders, what should we make of the sense of the word 'I'? Frege takes it that I have a direct, unmediated access to myself that is not available to others. 'Everyone,' he says, 'is presented to himself in a special and primary way, in which he is presented to no one else.' Yet if I say something about myself, I use the word 'I' in a way that *is* comprehensible to others. So the thought I have about myself is not the thought I express: the former is intrinsically incommunicable as the latter is not.

This is a highly uncomfortable conclusion for Frege to have reached. Wittgenstein's later response to the difficulty was to reject the Cartesian conception of the ego on which it is based. Frege's response was to attempt a refutation of idealism on Cartesian principles. But, as Wittgenstein pointed out, Frege's refutation is poor because it could at best refute a very simplistic kind of idealist.

The second of Frege's *Logical Investigations* is on 'Negation'. Here Frege pursues his distinction between thoughts and complexes in order to show that a false thought cannot be thought of as the absence of something. The main thing that is new here is a long response to the views of Bruno Bauch on negation.[25] Frege also offers an argument against regarding denial as a distinct linguistic act on a par with assertion. If we did this, he says, it would not enable us to do without negation, since we may wish to entertain, without either asserting or denying, a thought which we can only express by

[25] See Sven Schlotter, 'Frege's anonymous opponent in Die Verneinung', *History and Philosophy of Logic*, 27 (2006), pp. 43–58.

means of negation. On the other hand, if we have both assertion and negation, we do not need a distinct notion of denial, since denying something is tantamount to asserting its negation.

The third of the *Logical Investigations* extends Frege's account of thoughts in a fairly predictable manner to those formed by applying propositional connectives such as *and* or *or*.

LAST YEARS

In 1924 and 1925 Frege returned to the philosophy of mathematics and began to draft some work which, if he had completed it, would have amounted to a radical new approach to the philosophy of arithmetic. What prompted Frege to think about the foundations of arithmetic again is not known. His new idea was that, since logicism had wholly failed to supply a foundation for arithmetic, we should seek to base it instead on geometry. With this proposal Frege was, of course, reverting to the view of the ancient Greeks: in the *Elements* Euclid presents number theory as a branch of geometry. It is unfortunate that what has survived of these last thoughts on mathematics is not enough to amount to a worked-out view.

In mid-life Frege had been a supporter of the National Liberal Party, a traditionalist but mainstream political party which flourished in Prussia up to the First World War. In his old age, however, Frege veered to the right. At any rate he wrote shortly before he died a set of notes on his political views which have become notorious after Dummett commented on the editors' decision to omit them from the published version of his *Posthumous Writings* (on the ground that they have no bearing on his philosophy). The notes have now been published, but for many years most Frege scholars knew them only from Dummett's thumbnail sketch of them, and there was a danger that their notoriety would obscure the serious point Dummett was making, namely that by reading Frege's remarks he had learned something about human beings which, as he said, he 'should be sorry not to know'.

What little we know suggests in any case that Frege's last years were not happy. The value of his savings, which he had invested in German government bonds, had been severely damaged by the post-war hyperinflation and he was financially dependent on his pension. He fell out with his brother over whether he was responsible for paying an annuity that derived from their mother's estate. He suffered from a stomach complaint and was sometimes too ill to answer letters himself. When eventually he died, during the night of 25/26 July 1925, he was buried in the old cemetery at Wismar next to his parents.[26]

THE NACHLASS

Frege left his unpublished papers to Alfred, whom he had by then formally adopted as his son. 'They are not all gold,' he said, 'but there is gold in them. I believe that some of it will one day be held in much greater regard than now. See to it that nothing gets lost … It is a large part of myself that I here bequeath to you.'

In 1935 Alfred was approached by Heinrich Scholz of the University of Münster and agreed to hand the *Nachlass* over to Scholz with the intention that it should eventually be deposited at the university library there. Scholz, along with his assistant Hermann Schweitzer, set about preparing an edition of Frege's posthumous writings. During the war Scholz did indeed deposit the *Nachlass* for safekeeping with the manuscript department of the university library. The manuscript collection was evacuated to Salzuflen for safekeeping, but according to Heinrich Jansen, who was at the time head of the manuscript department, Frege's *Nachlass* was not included in the evacuation and was therefore destroyed when the library was set on fire by British bombers on 25 March 1945. What remained, and were eventually published many years later, were only the carbon copies of the typescripts Scholz had been preparing for publication before the war.

As Jansen was on sick leave at the time of the evacuation, and in Salzuflen at the time of the bombing, there is perhaps room to speculate on how authoritative his testimony is on the fate of Frege's papers. Nonetheless, it remains the case that no trace of the *Nachlass* has since been found.[27]

[26] Thus the official register: his tombstone mistakenly gives the date as 28 July.

[27] For a full discussion see Kai Wehmeier and Hans-Christoph Schmidt am Busch, 'The quest for Frege's Nachlass', in Beaney and Reck (eds.), *Gottlob Frege*, vol. I, pp. 55–68.

INFLUENCE

It is commonly said that Frege died a forgotten man. That is not quite true, but his influence on philosophy generally was largely indirect, via those he had influenced – Russell, Wittgenstein, Carnap. This method did not always transmit Frege's views with great accuracy. Russell, for example, did his best to bring Frege's works to the attention of an English readership when he summarized them in an appendix to the *Principles of Mathematics*, but thereafter he often proceeded as if he had not really understood them: Frege's long list of complaints,[28] when he tried to read the first few pages of *Principia*, may have been a little pedantic, but it was also just.

Wittgenstein's *Tractatus*, on the other hand, shows the deep influence of 'the great works of Frege' on almost every page. But studying the *Tractatus* is perhaps not a very good way of coming to understand Frege. It is, rather, the modern study of Frege that has given us a better chance of understanding the *Tractatus*.

Nonetheless, although Frege's influence on the major figures of the first half of the twentieth century is unquestionable, he was not at that time widely read. Max Black's 1934 book *The Nature of Mathematics*, for instance, contains extensive discussions of Ramsey, Russell and Wittgenstein, but treats Frege merely as an historical figure who 'unfortunately … did not use Boole's calculus of logic, preferring an elaborate but clumsy symbolism of his own, whose intricacy prevented his work receiving the recognition it deserved'. And in the 1930s the philosophy of logic, such as it was, displayed little sign of direct engagement with Frege's ideas.

One channel of influence which helped to nudge Frege into the mainstream was Carnap, his former student. Alonzo Church, who was Carnap's pupil and also studied at Göttingen in 1929, resurrected Frege's theory of sense from 'On sense and reference' in the 1940s. In Britain and America, however, Frege's influence was still limited by the lack of English translations. This began to change shortly after the war, when Max Black (who had spent some time at Göttingen) decided to translate 'On sense and reference' and the critique of formalism from volume II of *Grundgesetze*: these appeared in 1948.

[28] Letter to Jourdain, 28 January 1914, in *PMC*, pp. 81–4.

About the same time, J. L. Austin decided to offer the *Grundlagen* as a set text for a course he was teaching at Oxford. The translation of it which he produced for that purpose was published in 1950. Two years after that, Geach and Black published translations of a selection of Frege's other writings. The choice of what to include in their selection was strongly influenced by Wittgenstein.[29] This explains the omission from the collection of 'Thoughts', of which Wittgenstein had a low opinion, as well, no doubt, as the inclusion of the account of quantification from *Begriffsschrift* which had influenced Wittgenstein greatly.

Over the next two decades interest in Frege's works increased. One important step in persuading analytic philosophers of his central place in their subject was the publication of Kneale and Kneale's *The Development of Logic* in 1962. But the decisive moment was surely the publication of Michael Dummett's *Frege: Philosophy of Language* in 1973. Indeed Dummett's book did several things at once. It corrected some of the misconceptions that had built up around what Frege actually said. It convinced philosophers of the central importance of Frege's writings to many contemporary discussions in the philosophy of language. But it also showed a whole generation of philosophers a way of studying historical texts that was new to them. Dummett was concerned not so much to tell us where Frege had got his ideas from, but to treat Frege as a modern philosopher and evaluate his positions as if from the inside. 'It is impossible,' Dummett noted, 'adequately to evaluate Frege's doctrines without forming opinions about topics of which he did not treat … Some sections of the book are therefore hardly about Frege at all, but about matters which must be considered if one is to judge whether Frege spoke the truth.'

Even after Dummett's book it remained the case, however, that Frege's philosophy of mathematics was paid little attention. (Dummett's own book on the subject was delayed by more than a decade.) What led to Frege's re-entry to the mainstream of philosophy of mathematics was Crispin Wright's book, *Frege's Conception of Numbers as Objects*, in 1983. This book poses a deceptively simple question. What if Frege had not been persuaded by the Julius Caesar

[29] See P. Geach, Preface to Gottlob Frege, *Logical Investigations* (Oxford: Blackwell, 1977).

problem to drop the idea that number terms are implicit defined by means of Hume's Principle? If he had retained it, he would at a stroke have avoided the problems he got into by defining numbers in terms of extensions of concepts, because, as Wright conjectured and others proved soon afterwards, Hume's Principle (unlike Basic Law V) is consistent in second-order logic. (Geach had in fact pointed this out in print as early as 1976.)[30]

Wright's work has led to a renewed interest in abstraction principles, i.e. principles of the form

$$\Sigma(F) = \Sigma(G) \equiv F{\sim}G,$$

where $\sim$ is some suitably chosen equivalence relation on concepts. Hume's Principle is an example, being the abstraction principle obtained by taking the equivalence relation to be equinumerosity. Wright's proposal, which has come to be known as neo-Fregean logicism, is that abstraction principles may, in appropriate cases, provide us with a route to knowledge about some range of objects, and that approaching these objects via the abstraction principle is advantageous because it enables us to sidestep an epistemological debt which we would otherwise owe. If, for instance, we based arithmetic on the Dedekind-Peano axioms, we would owe an account of how we come to know that these axioms are true about the natural numbers. Neo-Fregeans, by contrast, have hoped that Hume's Principle might be seen, because it is an abstraction principle, as analytic (or something close to analytic) of the concept *number*. This hope has been encouraged by the metaphor of recarving employed by Frege himself. In an abstraction principle, he suggests, we replace the symbol $\sim$ 'by the more generic symbol = . … We carve up the content in a way different from the original way, and this yields us a new concept.' But we have already noted the difficulty Frege had in coming to a stable conception of the structure of propositional content, and the neo-Fregeans have not found it easy to cash out Frege's metaphor. A particular difficulty has been what is nowadays known as the Bad Company objection, namely that Basic Law V is also an abstraction principle, but an inconsistent one. Neo-Fregeans therefore have to find a principled way of distinguishing between

30 P. T. Geach, 'Critical notice of Dummett's *Frege: Philosophy of Language*', *Mind*, 85 (1976) pp. 436–49, at pp. 446–7.

consistent and inconsistent abstraction principles that does not simply appeal (as Geach's proof of the consistency of Hume's Principle did) to our prior knowledge of the domain our grasp of which we are trying to explain.

FURTHER READING

For a lucid and straightforward introduction to Frege's writings at a very introductory level Kenny[31] could scarcely be bettered. The remaining chapters of this Companion discuss most aspects of Frege's work in depth. In addition, readers may wish to consult Sullivan[32] for a good discussion of the logic of *Begriffsschrift*, and Dummett[33] for an excellent exposition of *Grundlagen* and *Grundgesetze*.[34]

[31] Anthony Kenny, *Frege* (London: Penguin, 1995).
[32] Sullivan, 'Frege's logic'.
[33] Michael Dummett, *Frege: Philosophy of Mathematics* (Duckworth, 1991).
[34] I am grateful to Michael Beaney and Peter Sullivan for comments on an earlier draft of this chapter.

JOAN WEINER

2 Understanding Frege's project

Frege begins *Die Grundlagen der Arithmetik*, the work that introduces the project which was to occupy him for most of his professional career, with the question, 'What is the number one?' It is a question to which even mathematicians, he says, have no satisfactory answer. And given this scandalous situation, he adds, there is small hope that we shall be able to say what number is. Frege intends to rectify the situation by providing definitions of the number one and the concept of number. But what, exactly, is required of a definition? Surely it will not do to stipulate that the number one is Julius Caesar – that would change the subject. It seems reasonable to suppose that an acceptable definition must be a true statement containing a description that picks out the object to which the numeral '1' already refers. And, similarly, that an acceptable definition of the concept of number must contain a description that picks out precisely those objects that are numbers – those objects to which our numerals refer.

Yet, while Frege writes a great deal about what criteria his definitions must satisfy, the above criteria are not among those he mentions. Nor does he attempt to convince us that his definitions of '1' and the other numerals are correct by arguing that these definitions pick out objects to which these numerals have always referred. There is, as we shall see shortly, a great deal of evidence that Frege's definitions are not intended to pick out objects to which our numerals already refer. But if this is so, how can these definitions teach us anything about our science of arithmetic? And what criteria must these definitions satisfy? To answer these questions, we need to understand what it is that Frege thinks we need to learn about the science of arithmetic.

WHY DEFINE THE NUMBER ONE?

Definitions of the number one and the concept of number are necessary, Frege thinks, if we are to prove the truths of arithmetic from primitive truths. What are primitive truths? And why should we prove the truths of arithmetic from primitive truths? In the early sections of *Grundlagen*, Frege offers two motivations for attempting to provide such proofs. The first, which he characterizes as mathematical, is a desire for increased rigour – proof wherever proof is possible. The second, which he characterizes as philosophical, is a desire to show whether the truths of arithmetic are a priori or a posteriori, synthetic or analytic.

For Frege, the classification of a provable truth as analytic or synthetic, a priori or a posteriori, is determined by its most economical (most general) proof – by the proof requiring the fewest specific assumptions. The least economical (least general) sort of proof is one that requires an appeal to facts, that is, unprovable truths about particular objects; unprovable truths that are not general. Appeals to facts are required by any proof of an a posteriori truth.[1] Truths of empirical science are examples of a posteriori truths. A truth that can be proved without appeal to facts is a priori and can be either synthetic or analytic. This classification, again, depends on what sort of proof is available. An a priori truth is synthetic if it cannot be proved 'without making use of truths which are not of a general logical nature, but belong to the sphere of some special science'.[2] Truths of Euclidean geometry are examples of synthetic a priori truths. For the axioms from which they are derived are not of a general logical nature (they govern a limited domain: that of spatial configurations) but are general (they are not truths about particular objects). Finally, an analytic truth can be proved using only 'general logical laws and definitions'. This is the most economical (or general) sort of proof – it requires no appeals to facts or to truths of a special science.

To find the most economical proof of some truth, we need a method for recognizing gapless proofs. Otherwise, we cannot rule out the possibility that a proof that apparently has only general

[1] *Gl*, §3, p. 4. All page references to Frege's works in this chapter are to the original German edition, unless otherwise stated.
[2] *Ibid.*

logical laws and definitions among its premises actually contains an implicit appeal to something that is neither a logical law nor a definition. The task is 'that of finding the proof of the proposition, and of following it up right back to the primitive truths'.[3] In the process, Frege says,

> we very soon come to propositions which cannot be proved so long as we do not succeed in analysing concepts which occur in them into simpler concepts or in reducing them to something of greater generality. Now here it is above all Number which has to be either defined or recognized as indefinable.[4]

One might suppose that, in this process, the concept of number *will* be recognized as indefinable. Yet Frege insists on defining the concept of number and the numerals. Why?

If the point of proving truths of arithmetic from primitive laws is to enable us to determine the correct classification of these truths, there will be eligibility conditions that determine what can be taken as a primitive law. One obvious eligibility condition is that its truth be evident without proof.[5] Another is that there be some means, other than examining a proof, of determining its classification (i.e., of determining whether it is a fact about particular objects, a primitive general truth of some special science, or a general logical law). For if there is to be a definite answer to the question about the correct classification of the truths of arithmetic, then there must be some means of classifying the primitive laws on which the truths of arithmetic depend.

To see how this might work, consider an example: the claim that every object is identical to itself.[6] Since its truth is self-evident, it satisfies the first eligibility requirement for primitive laws. Supposing this to be a primitive law, is it analytic? In *Grundlagen*, Frege mentions two features of analytic truths. One is maximal

[3] *Ibid.*

[4] *Ibid.*, §4, p. 5.

[5] Although Frege does not explicitly discuss this, it is obvious that, if the proofs based on an unproved primitive law are to establish the truth of their conclusion, the truth of the primitive law must be evident without proof.

[6] This is primitive-eligible, but not a basic *Begriffsschrift* law. Since Frege wants to minimize the number of primitive laws (see, e.g., *Gg*, vol. I, p. vi), he derives many laws that are primitive-eligible.

generality. Analytic truths govern 'not only the actual, not only the intuitable, but everything thinkable'.[7] Another is that we cannot deny them in conceptual thought. That is, we cannot deny them 'without involving ourselves in any contradictions when we proceed to our deductions'.[8] Fundamental truths of arithmetic seem to be analytic because if we try to deny any one of them 'complete confusion ensues. Even to think at all seems no longer possible.'[9] The law that every object is identical to itself exemplifies both of these features. First, this law surely tells us, not just about every actual (spatio-temporal) object or every intuitable object, but about *every* object. Second, it seems that we cannot deny it without involving ourselves in contradictions. Given these criteria, the law in question is analytic. The axioms of geometry, in contrast, are synthetic because we *can* assume the contrary of an axiom of geometry without involving ourselves in contradictions.[10]

What of basic truths about numbers? Frege suggests, without argument, that the fundamental propositions of the science of number have the same status as logical laws – that denying them will involve us in contradictions.[11] He also states, again without argument, that the truths of arithmetic govern the widest domain of all (*das umfassendste*). Thus these truths seem to be logical laws. But there are also reasons for thinking that they are not primitive-eligible logical laws. Truths about the number one do not seem to have the requisite maximal generality of logical truths. The number one, after all, is a particular object. Nor do laws about numbers seem maximally general. They seem to govern, not the widest domain of all, but the peculiar domain of numbers. Inferences by mathematical induction appear to be 'peculiar to mathematics'.[12] To substantiate his conviction that the truths of arithmetic are analytic, Frege needs to define the number one and the concept of number from recognizably logical notions and to prove the truths of arithmetic using only these definitions and logical laws.

[7] *Gl*, §14, p. 21.
[8] *Ibid.*, §14, p. 20.
[9] *Ibid.*, §14, p. 21.
[10] *Ibid.*, §14, pp. 20–1.
[11] *Ibid.*, §14, p. 21.
[12] *Ibid.*, §14, p. iv.

DEFINITIONS AND CONTENT

We now have one criterion that Frege's definitions must satisfy. They must enable him to provide gapless proofs of the truths of arithmetic from primitive truths – from primitive logical laws, if he is to show that they are analytic. Since the proofs must be of truths of *arithmetic*, Frege's definitions must not transform arithmetic into some new and foreign science. One might suppose, then, that Frege's aim is to give descriptions that pick out the objects that we are already talking about when we use the numerals and the term 'number'. Why, then, does he not say so?

The explanation, one might suspect, is simply that Frege expected this to be obvious to his readers. But a problem remains. If this is right, Frege's defence of his definitions should include an attempt to show that his definitions pick out the objects to which our numerals already refer. But Frege's *Grundlagen* defence of his definitions includes no argument that they pick out the objects to which our numerals already refer. What is Frege's defence?

The defence in *Grundlagen* appears in a group of sections labelled 'the completion and testing (*Ergänzung und Bewährung*) of our definition'. He first defines the concept *number which belongs to the concept F* and shows that this definition passes several tests. He then turns to the task of completing his definitions – defining the individual numbers – which is followed with tests of these definitions. The tests are tests of whether the definitions allow us to derive 'the well known properties of numbers'.[13] What are these properties? Although Frege is renowned for claiming that numbers are non-spatio-temporal objects, this is not the sort of property that must be derivable from the definitions. Rather, the properties in question are those that seem to underlie the uses we make of arithmetic, both in science and in everyday life. For example, we must be able to prove, using his definitions, that 0 is the number belonging to a concept if and only if no object falls under it (the number that belongs to a particular concept is the number of objects that fall under the concept); that if 1 is the number which belongs to a concept, then there exists an object which falls under that concept.[14]

[13] *Ibid.*, §70, p. 81.
[14] *Ibid.*, §75, p. 88; §78, p. 91.

The definitions must provide a basis for an arithmetic that meets the demand 'that its numbers should be adapted for use in every application made of number'.[15] Thus the definitions must be responsible to pure arithmetic. For the applications of arithmetic are applications of pure arithmetic. If Venus has 0 moons and the Earth has 1 moon, we ought to be able to infer that the Earth has more moons than Venus – something that would be blocked were it a truth of Fregean arithmetic that 0 = 1. What we take to be simple truths and applications of our arithmetic must be reproducible in an arithmetic based on Frege's definitions. No acceptable definitions of '0' and '1' will make it true that 0 = 1 or false (failing new astronomical events) that the Earth has 1 moon. Moreover, it will not suffice that it be *true*, given Frege's definition of '1', that if 1 is the number which belongs to a concept, then some object falls under the concept. It must be *derivable*. The definitions must not only preserve what we regard as the truths of pre-systematic arithmetic, they must also provide support for its inferences. That is, the introduction of these definitions should enable us to replace our original, enthymematic arguments about, say, the numbers of moons of Venus and the Earth, with gapless arguments.[16]

Definitions satisfying these constraints clearly preserve some pre-systematic content associated with the numerals and the term 'number'. This content seems very like the kind of content Frege introduces in *Begriffsschrift*:[17] conceptual content (*begriffliche Inhalt*), content that has 'significance for the inferential sequence'. And Begriffsschrift, the language in which he wants to carry out the proofs that will establish the analyticity of arithmetic, is designed to be a language that expresses conceptual content. This suggests that the criterion that must be met, if we are legitimately to regard Frege's definitions as faithful to arithmetic, is that they preserve whatever conceptual content is inherent in our pre-systematic views about arithmetic.

[15] *Ibid.*, §19, p. 26.

[16] One might suspect, as Patricia Blanchette argues in 'Frege's Reduction', *History and Philosophy of Logic*, 15 (1994), pp. 85–103, that Frege requires statements in the systematic science of arithmetic to be logically equivalent to claims of ordinary pre-systematic arithmetic. I argue below that this interpretation conflicts with Frege's statements about the roles played by Begriffsschrift and natural language.

[17] I use the word 'Begriffsschrift' italicized to refer to Frege's monograph, unitalicized to refer to his logical language.

Although Frege does not use the expression 'conceptual content' in *Grundlagen*, the link between the content he wants to preserve and significance for inference is evident. He wants to convince us that intellectual effort is needed if we are to understand the content of the expression 'number' and the numerals. And one sort of evidence he offers is that, while we routinely make everyday inferences from numerical formulae, these inferences do not seem immediately licensed by logical laws. Nor is it evident how these inferences can be made gapless. The link between content and inference is also apparent in other discussions of *Grundlagen*. For example, in his discussion of the content of the proposition 'All whales are mammals', he argues that the proposition is not about animals because, 'We cannot infer from it that the animal before us is a mammal without the additional premiss that it is a whale, as to which our proposition says nothing.'[18]

He also argues that the ideas we associate with an expression cannot constitute its content, because the associated ideas do not support our inferences. Thus Frege's explicit requirements on his definitions involve preservation of whatever conceptual content is already associated with the word 'number' and the numerals.

DEFINITIONS AND REFERENCE

But we might well have expected Frege to require that definitions of the numerals pick out whatever it is that we have been talking about all along. Or, to use a contemporary locution: that definitions of the numerals preserve pre-systematic reference. Yet Frege not only fails to articulate this requirement, he makes no attempt to show that his definitions satisfy it. One might suspect that he simply assumes that, to show that the definition picks out the object we have been talking about in our use of the term '1', it will suffice to show that the definition preserves the conceptual content of '1'. But Frege's actual remarks suggest something very different: that the terms to be defined do not actually have reference antecedent to his work.

Frege writes, in a criticism of a proposed definition of the concept of number, 'it must be noted that for us the concept of number has not yet been fixed, but is only due to be determined in the light

[18] *Gl*, §47, p. 60.

of our definition of numerical identity'.[19] This is an odd choice of words if each numeral *already* refers to a particular object and if to be a natural number is simply to be one of these objects. For Frege writes that all that can be demanded of a concept is that it should be determined, for each object, whether or not it falls under the concept.[20] If each numeral *already* refers to a particular object, and if the numbers are the objects to which the numerals refer, then the concept of number is already fixed. We may be lacking a definition that identifies this fixed concept of number. But it certainly does not follow that the concept of number is due to be determined in the light of our definition. Were this Frege's only remark of the sort, one might dismiss it as merely an odd choice of words. But it is not.

Most of the discussions of *Grundlagen* are about the natural numbers. However, Frege also discusses the complex numbers. He considers the possibility of stipulating that the time-interval of one second is the square root of –1, and adds, in a footnote, that we are entitled to choose any one of a number of objects to be the square root of –1. The reason is that 'the meaning [*Bedeutung*] of the square root of –1 is not something which was already unalterably fixed before we made these choices, but is decided for the first time by and along with them'.[21]

If this is so, our symbols for complex numbers do not already refer to particular objects – which he goes on to suggest in the next section. In this remark, unlike the earlier remark about the concept of number, there is no ambiguity. One might suspect that this marks a difference between the complex and natural numbers. But Frege gives us no indication that there is such a difference.

He goes on to suggest that there is a problem with defining the square root of –1 as a time-interval. This would import into arithmetic 'something quite foreign to it, namely time' and make arithmetic synthetic.[22] To show that arithmetic is analytic, Frege proposes using the same solution for complex numbers that he used for natural numbers: to define them as extensions of concepts. The notion of extension of concept is a logical notion, on Frege's view, and definitions of numbers as extensions of concepts should make

[19] *Ibid.*, §63, p. 74.
[20] *Ibid.*, §74, p. 87.
[21] *Ibid.*, §100, p. 110.
[22] *Ibid.*, §103, p. 112.

it possible to prove truths about numbers from logical laws. He ends *Grundlagen* with the following remark about offering such definitions:

Once suppose this everywhere accomplished, then numbers of every kind, whether negative, fractional, irrational or complex, are revealed as no more mysterious than the positive whole numbers, which in turn are no more real or more actual or more palpable than they.[23]

This would be an odd remark if, for example, '1' had *all along* referred to a particular extension of a concept while the symbol '*i*' refers to an extension of a concept only because of an arbitrary stipulation. But, again, this may be simply an odd choice of words. What other evidence is available?

Frege acknowledges that the correctness of his definitions is not evident. For we 'think of the extensions of concepts as something quite different from numbers'.[24] One might expect him to go on to argue that numbers really are extensions of concepts. But he does not. Rather, he claims that he attaches no decisive importance to bringing in the extensions of concepts.[25] This is completely mysterious if we assume that, when we use the numerals in our current pre-systematic language, we are talking about particular objects, and if we assume that Frege's task is to provide definitions that pick these objects out. Given these assumptions, either we are already talking about (our numerals already refer to) extensions of concepts (in which case it would be *essential* to bring in extensions) or we are already talking about (our numerals already refer to) objects other than extensions of concepts (in which case it would be *wrong* to bring in extensions). Frege's comments are simply not consistent with the assumption that his definitions are meant to pick out objects that we have been talking about all along. Unless we are prepared to engage in interpretive contortions, the appropriate conclusion is that, when Frege asks for a definition of the concept number, he is not asking for explicit descriptions of objects to which our numerals already refer. And, given this, it is implausible to attribute

[23] *Ibid.*, §109, p. 119.
[24] *Ibid.*, §69, p. 80.
[25] *Ibid.*, §107, p. 117. Later, Frege attached more importance to bringing in extensions. But his reason is that 'we just cannot get on without them' (*Gg*, vol. I, p. x), not that numbers really are extensions.

to Frege the view that there is a concept to which 'number' refers, and objects to which the numerals refer, antecedent to his introduction of his definitions. That is, antecedent to Frege's introduction of his definitions, the concept of number is not fixed.[26]

Of course, if Frege's explicit remarks are absurd, there may be a compelling reason to engage in interpretive contortions. But are they? There are many distinct set theoretic definitions of the numbers that fit our understanding – both everyday and scientific – of the numbers. Nothing in our understanding of the truths of arithmetic seems to offer grounds for deciding between alternative systems of set theoretic definitions or, for that matter, grounds for saying that numbers are (or are not) sets. Given this, Frege's explicit remarks do not seem absurd at all. There is every reason to believe that the numerals do not refer to particular objects and, consequently, that the content associated with the numerals can be captured by offering definitions that are at least partly stipulative.

REFERENCE AND TRUTH

There is a problem, however. Frege seems to assume, not just that such everyday sentences of arithmetic as '0 is not equal to 1' express truths but that they express truths about particular numbers. Otherwise, what would be the point of defining the numbers as objects? But now suppose '0 is not equal to 1' expresses a true claim about particular numbers. It seems that its truth must depend on the character of those numbers – i.e., the character of the objects to which '0' and '1' refer. If there are no objects to which '0' and '1' refer, it follows that '0 is not equal to 1' does not express a truth. It seems to follow that no statements of everyday arithmetic can express truths.

Frege never addresses this problem. The explanation, one might suspect, is that he simply did not notice this consequence of his views.[27] But this is not entirely convincing. For, he comes very close

[26] I concentrate here on reference. For a discussion of sense and of Frege's discussion of analytic vs. constructive definitions see 'What is a numeral? Frege's answer', *Mind*, 116 (2007), pp. 677–716.

[27] As Gary Kemp argues in 'Frege's sharpness requirement', in *Philosophical Quarterly*, 46 (1996), pp. 168–84. Many of the following arguments are responses to his objections.

to explicitly acknowledging this consequence. He writes 'would the sentence 'any square root of 9 is odd' have a comprehensible sense at all if *square root of 9* were not a concept with a sharp boundary?'[28] A sentence without comprehensible sense cannot have a truth-value. One might suspect that Frege takes it to be obvious that 'any square root of 9 is odd' does have a comprehensible sense and, hence, that the concept *square root of 9* does have a sharp boundary. However, a look at the context in which the question appears shows that this interpretation is incorrect. A concept has a sharp boundary just in case it determinately holds or not of each object. For example, in order for *greater than zero* (or *positive*) to be a proper concept, Frege says, 'it would have to be determinate whether, e.g., the Moon is greater than zero'.[29] He continues,

> We may indeed specify that only numbers can stand in our relation, and infer from this that the Moon, not being a number, is also not greater than zero. But with that there would have to go a complete definition of the word 'number', and that is just what is most lacking.

In the discussions of *Grundgesetze* that immediately follow, he suggests that such expressions as 'greater than' and '+' are used by mathematicians in such a way that they have no fixed meaning.[30]

It is difficult to imagine that Frege said all this without noticing the consequence that sentences in which 'greater than 0', 'greater than' and '+' appear have no truth-value.[31] Indeed, given his requirement that each predicate pick out a concept with a sharp boundary, few, if any, of our everyday sentences have comprehensible sense or truth-values. But, whether he noticed this or not, this creates a puzzle about Frege's conception of his project. Frege's avowed project is to show that the truths of arithmetic are analytic. Unless we already know some of these truths – unless our everyday sentences of arithmetic express them – what could be the point of this project? Although Frege does not address this problem explicitly, there are solutions to it to be found in his discussions of natural language, Begriffsschrift and science.

[28] *Gg*, vol. II, §56.
[29] *Ibid.*, §62.
[30] See the arguments in *ibid.*, §§56–67.
[31] See also the discussion of the universal generalization of '$(x>2) \supset (x^2>2)$' in 'Peano's conceptual notation'.

Frege characterizes Begriffsschrift, his logical language, as a tool that enables us to avoid some difficulties inherent in natural language. When we use natural language, he says, even careful use of logical laws will not prevent errors. Mistakes, he writes, 'easily escape the eye of the examiner, especially those which arise from subtle differences in the meanings of a word.'[32] He continues, 'That we nevertheless find our way about reasonably well in life as well as in science we owe to the manifold ways of checking that we have at our disposal. Experience and space perception protect us from many errors.' Frege does not suggest that there is anything wrong with relying on the manifold ways of checking or that the subtle differences in the meanings of a word should be eliminated from natural language. Rather, these features of natural language are rooted 'in a certain softness and instability of language which nevertheless is necessary for its versatility and potential for development'.

> In this respect, language can be compared to the hand, which despite its adaptability to the most diverse tasks is still inadequate. We build for ourselves artificial hands, tools for particular purposes, which work with more accuracy than the hand can provide. And how is this accuracy possible? Through the very stiffness and inflexibility of parts the lack of which makes the hands so dextrous. Word-language is inadequate in a similar way. We need a system of symbols from which every ambiguity is banned, which has a strict logical form from which the content cannot escape.

Neither natural language nor a logically perfect symbolic language is suitable for every purpose. Whether features of a language count as virtues or defects will depend on the purpose for which we want to use the language. Features of natural language that are defects, given Frege's specialized purposes, are desirable for other purposes. Begriffsschrift is not an ideal language. It is 'a device invented for certain scientific purposes and one must not condemn it because it is not suited to others'.[33]

Begriffsschrift is designed for the expression and evaluation of inferences. It must be capable of expressing all content of any statement that has significance for the inferences in which it can figure. Once an inference is expressed in Begriffsschrift, the employment of Frege's logical laws and rules are to make it a mechanical task to

[32] 'On the scientific justification of *Begriffsschrift*', p. 51/*CN*, p. 86.
[33] *Bs*, p. v.

determine whether it is correct and gapless, or whether it requires an unstated premise. Because the task is mechanical, no presupposition can sneak in unnoticed. We need such a language and logical system in order to produce identifiably gapless proofs of the truths of arithmetic. And only identifiably gapless proofs from primitive truths will enable us to determine whether the truths of arithmetic are correctly classified as analytic or synthetic.

In order to carry out this project, we must define all terms of arithmetic from primitive, undefinable terms and construct a list of axioms or primitive truths from which all truths of arithmetic can be proved by gapless logical inferences. To do this is (to use Frege's later expression) to provide a systematic science. And science, Frege claims, comes to fruition only in a system.[34] Arithmetic is a science in its early stages – a science whose sentences have not yet been associated with precise thought content. It is not that that arithmetic is less developed than other sciences. Although it is as highly developed as any science, arithmetic does not satisfy the standards for systematic science. In fact, there are no systematic sciences – Euclidean geometry comes closest, but its proofs are not gapless.[35] Frege's systematic science of logic, of which arithmetic is a part, will be the first.

NATURAL LANGUAGE AND BEGRIFFSSCHRIFT

We can now see why it is not absurd for Frege to say that the everyday sentences of natural language do not have truth-values. Frege's view seems to be that truth is what we get, not in everyday circumstances, but rather at some ideal end of inquiry. And the language for this ideal end, the language for systematic science, is not natural language but Begriffsschrift. But while natural language may not be a good vehicle for expressing truth, it is an essential tool in the early stages of our attempts to express truths. In a diary entry, Frege wrote, of his attempt to say what the numbers are,

> [O]ne might think that language would first have to be freed from all logical imperfections before it was employed in such investigations. But of

[34] 'Logic in mathematics', *NS*, p. 261. Also *PW*, p. 242.

[35] See, e.g., Frege's discussion, in 'On the scientific justification of *Begriffsschrift*', pp. 50–1/*CN*, pp. 84–5, of Euclid's tacit presuppositions.

course the work necessary to do this can itself only be done by using this tool, for all its imperfections. Fortunately as a result of our logical work we have acquired a yardstick by which we are apprised of these defects. Such a yardstick is at work even in language, obstructed though it may be by the many illogical features that are also at work in language.[36]

To systematize a science we begin with the everyday sentences that are regarded as its basic truths – such sentences as '0 is not equal to 1'. Our everyday view that this sentence expresses a truth is not quite right. The content associated with it is not yet precise enough; the science is not yet sufficiently well worked out. But much of the science of arithmetic is worked out. Many of the standards by which we judge sciences have been met. This sentence provides a guide for systematizing arithmetic. For it places constraints on our assignments of meaning to the terms '0' and '1'. On any acceptable assignment the sentence '0 is not equal to 1' must express a truth. Since it will help to have a label for this attitude in the discussion that follows, I will say that Frege *regards these sentences as true*. It is a consequence of Frege's view that few, if any, of our everyday sentences actually express truths. Nonetheless it is consistent with his view that we can regard some of these sentences – particularly the results of pre-systematic research – as expressing truths.

One might suspect that this view must conflict with Frege's statements in 'On sense and reference', which includes extensive discussion of natural language. Frege introduces his renowned *Sinn*/*Bedeutung* distinction by talking about words and sentences of everyday language. And the *Bedeutung* of an object expression is whatever object that expression designates. Yet it is difficult to find any actual inconsistency. Although Frege writes as if the terms of everyday language have *Bedeutung* and the sentences of everyday language have truth-values, he never actually *says* that they do. It is not because the subject never comes up. Although he raises the question of whether 'the Moon' has a *Bedeutung*, he does not go on to say that it does. He says only that we 'presuppose a meaning [*Bedeutung*]'.[37] Nor does he say that such presuppositions are always – or generally, or even *sometimes* – correct. He says only,

[36] September 1924. See *NS*, p. 285/*PW*, p. 266.
[37] 'On sense and reference', p. 31.

the question whether the presupposition is perhaps always mistaken need not be answered here; in order to justify mention of that which a sign means it is enough, at first, to point our intention in speaking or thinking. (We must then add the reservation: provided such a meaning exists.)[38]

We do, of course, presuppose that our terms have *Bedeutung*, that there is something we really are talking about and that our sentences really have truth-values.[39] As we have seen, Frege's comments about the numerals and 'number' indicate that he thinks there are scientific contexts in which this presupposition is incorrect. The incorrectness of this particular presupposition has not, however, impeded our everyday arithmetic. It has not even impeded such sophisticated mathematical uses as 'Weierstrass'.[40] It is the project of systematization that requires both that all presuppositions be eliminated and that the necessary work be done to guarantee that each term has *Bedeutung*.

This is not to say that *Bedeutung* is unimportant in our use of natural language. But it is not a prerequisite for our use of natural language – even in scientific contexts – that our terms have *Bedeutung*. But is this view plausible? Surely, one might think, it is essential that terms used in scientific contexts have *Bedeutung*. In fact, however, this apparently implausible view, at least in some cases, fits our conception of good scientific practice perfectly well. To see this, it will help to look at an example.

Today, as a result of a good deal of research, it is widely regarded as a well-established truth that obesity increases one's risk of heart disease. Yet 'obese' no more designates a fixed concept than 'number'. Although medical researchers studying obesity agree that obesity is some weight-related characteristic that is associated with increased morbidity and mortality, several distinct sorts of definitions are used in medical research. Most common, because of convenience, are definitions in terms of body mass index, an index calculated using measurements of height and weight.[41] The current general

[38] *Ibid.*, pp. 31–2.

[39] Frege also warns against apparent proper names without *Bedeutung*. But the imperfection in question is not that language *has* proper names with no *Bedeutung*, but rather that it is *possible to form* proper names with no *Bedeutung*. This possibility cannot be prohibited in natural language.

[40] See 'Logic in mathematics', *NS*, p. 239/*PW*, p. 221.

[41] Body mass index (or Quetelet index) is defined as: [weight in kg]/[height in meters]^2.

acceptance of the definition of obesity as BMI > 30 – by researchers, the World Health Organization, public health officials, newspaper reporters and their readers – is a fairly recent phenomenon. Only fifteen years ago the preferred definition was a two-part definition: for men the obesity began at BMI > 28.7, for women BMI > 28.3.[42] It was not that research revealed that the new cut-off was correct and the original incorrect. Rather, it was (and is) understood that any cut-off is to some extent arbitrary. This particular change was made, in part, for convenience. And it is also widely acknowledged that the current definition is not ideal. For almost everybody believes that obesity has something to do with body fat and some highly muscled athletes who do not have much body fat will be classified as obese, given this cut-off.[43]

The search for a good definition of obesity continues, along with the investigation of various hypotheses about obesity. Yet it would be unreasonable to halt all investigation of the effects of obesity on morbidity and mortality on the grounds that, since the concept has yet to be fixed, the hypotheses have no truth-values. It would be unreasonable to give up our view that it is *true* that obesity increases risk of heart disease. That is, an apparently absurd view that Frege seems to hold – that we are entitled to regard certain sentences as expressing truths, in spite of the fact that some of their terms do not have fixed meaning – is not absurd at all. It aptly describes perfectly unexceptionable views of researchers. But this is not to say that the issue of a term's having fixed meaning is of no concern to this sort of science. In fact, the problem with requiring that all terms used in scientific investigation already have fixed meaning is precisely that it can be part of the scientific enterprise to fix the meaning associated with a term already in use. The procedure, as we have noted already, involves a combination of research and stipulation.

What, then, is Frege's view of truth? He may seem to have two notions: the strict sort of truth that is the aim of science and a different sort of truth that applies to sentences of natural

[42] R. J. Kuczmarski and K. M. Flegal, 'Criteria for definition of overweight in transition: background and recommendations for the United States', *American Journal of Clinical Nutrition*, 72 (2000), p. 1077.

[43] See, e.g., the historical remarks in *ibid.* Myriad internet web pages give examples of athletes who count as obese on this definition. See, e.g., www.obesityscam.com/myth1.1.htm.

language – something very like the supervaluationist notion of truth.[44] After all, the significance of his regarding it as true that each number has a unique successor is that on every acceptable definition of the term 'number', it will be provable, hence true, that each number has a unique successor. There is, however, an important difference between Frege's and the supervaluationist's views of natural language. Although Frege shares the supervaluationist view that there is something right about many of our everyday sentences, he does not share the supervaluationist view that we are correct to presuppose that the constituents of these sentences have fixed meaning. For while there is something right about the sentences that we regard as setting out fundamental truths of pre-systematic arithmetic, the demands of truth, as Frege understands them, show us that there is also something wrong with these sentences. Frege wants to satisfy these demands, using what is right about pre-systematic arithmetic as a starting point.

Frege wants to replace imprecise pre-systematic sentences with precise systematic sentences – e.g. to introduce definitions of 'number' and 'successor', from which it can be proved that each number has unique successor. For Frege's interest in 'the sort of truth which it is the aim of science to discern' will not allow him to rest content with the standards of pre-systematic arithmetic.[45] To say that our statements do not now satisfy Frege's demand that all constituents have fixed meaning is merely to say that we are not finished. Our sciences have not yet reached fruition. The demands that Frege identifies as the demands of truth should be seen as part of a regulative ideal for science. But there is no reason to assume that any sentences of natural language actually satisfy the demand. Thus we can reconcile Frege's conception of his project with his statements about truth. To show that the truths of arithmetic are analytic is not to undertake a project external to the development of the science of

[44] Central to the supervaluationist approach is the notion of precisification or a sharpening of the bounds of a predicate. Given a particular precisification of, e.g., the term 'bald', each person is either bald or not bald. On the supervaluationist account of a sentence containing a vague predicate, the sentence is true just in case it is true given any admissible precisification. See, e.g., Kit Fine, 'Vagueness, truth and logic', *Synthese*, 30 (1975), pp. 265–300.

[45] 'Thoughts', p. 59.

arithmetic. It is a further – the final – step in bringing this science to fruition.

SEMANTIC DESCENT

There is, for Frege, only one notion of truth. It is what we obtain in systematic science. And Frege does not think this is an unobtainable ideal. His new logic is meant to be a systematic science. Moreover logic, he tells us, has a special relation to truth: the task of logic is to discern the laws of truth. Thus Frege's logic seems to give us some sort of theory of truth. But what is this theory like? It is widely supposed that Frege means to give a theory that tells us how the truth of a sentence is determined by semantic values of its subsentential constituents. Of course, no such theory is stated in Begriffsschrift. But the Begriffsschrift proofs, many think, are only one part of Frege's logic. His logic, on this view, is a familiar enterprise that involves a formal language and its interpretation; it is a science in which metatheory and metatheoretic proof play important roles. There are, however, a number of difficulties with this reading.[46]

One difficulty lies in the significance accorded to language. If a theory of truth tells us how the truth of a sentence is determined by semantic values of its subsentential constituents, then language would appear to be the subject of the theory of truth. Moreover, language appears to be the subject of most metatheoretic proof. Consider, for example, the sort of metatheoretic justification one might offer for *modus ponens*:

if A is true and A → B is true, then B is true.

This purports to state a general truth about sentences, with 'A' and 'B' used as metatheoretic variables that range over sentences. But as we saw earlier, Frege believes that laws of logic are distinguished by their universality. They hold, not just over the realm of some special science or the realm of the spatio-temporal, but over an unrestricted realm. How can the metatheoretic claim about *modus ponens* – which appears to be, not a statement about

[46] For a more thorough discussion, see J. Weiner, 'Semantic descent', *Mind*, 114 (2005), pp. 321–54.

everything thinkable, but a statement about the restricted realm of the linguistic – justify truths that hold over an unrestricted realm?

Quine, who also claims that language is not the subject of logic, offers us a familiar answer.[47] He points out that '"Wombat" is true of some creatures in Tasmania', which is about a linguistic expression, is also a paraphrase of 'There are wombats in Tasmania', which is not. Thus it is possible to use statements about language to express something whose subject is really not language at all. The strategy of talking about words when our actual interest is in something else, which Quine labels *semantic ascent*, is, he argues, necessary for logic. Logic 'can be expounded in a general way *only* by talking of forms of sentences'.[48]

The reason stems from the sorts of generalizations required by logic. Consider the clause 'time flies' in the sentence 'if time flies then time flies'. Quine writes,

> We want to say that this compound sentence continues true when the clause is supplanted by any other; and we can do no better than to say just that in so many words, including the word 'true'. We say 'All sentences of the form "if p then p" are true.' We could not generalize as in 'All men are mortal', because 'time flies' is not, like 'Socrates', a name of one of a range of objects (man) over which to generalize. We cleared this obstacle by *semantic ascent* by ascending to a level where there were indeed objects over which to generalize, namely linguistic objects, sentences.[49]

On Quine's account, semantic ascent solves a problem. Semantics is required for logic because the generalization needed for a general account of the logical laws is not generalization over objects.

To see how this works, consider a contemporary rendering of Frege's Basic Law I, '$(A \rightarrow (B \rightarrow A))$'. The contemporary rendering is

[47] That we need to talk of forms of sentences and truth predicates in order to make general claims (e.g., of infinitely many axioms of the form '$P \rightarrow P$', that they are logical truths) is not just Quine's view. See, e.g., Jason Stanley's claim, in 'Truth and metatheory in Frege', *Pacific Philosophical Quarterly*, 77 (1996), p. 53, that one reason that a truth predicate occurs ineliminably in discussions of the validity of rules of inference is that they are generalizations.

[48] W. V. O. Quine, *Word and Object* (Cambridge, Mass.: MIT Press, 1960), p. 273, my emphasis.

[49] W. V. O. Quine, *Pursuit of Truth* (Cambridge, Mass.: Harvard University Press, 1992), p. 81; see also *The Philosophy of Logic*, 2nd edn (Cambridge, Mass.: Harvard University Press, 1986), pp. 11–12.

neither an expression in contemporary logical notation nor a single logical law. It is, rather, a schema in which '*A*' and '*B*' are used as metalinguistic variables that range over sentences. A claim about the truth of Basic Law I is really a claim about the truth of infinitely many sentences in the formal language. Similarly, the statement that explains the justification of *modus ponens* is the statement of a general claim about sentences and truth: for any sentences A and B, if A is true and A → B is true, then B is true. One hallmark of contemporary logic, then, is the use of schemata. Michael Dummett writes, 'Logic can begin only when the idea is introduced of a schematic representation of a form of argument.'[50]

Another hallmark of contemporary logic is the use of the truth predicate, where truth is a property of sentences.[51] Quine writes that the truth predicate has its utility,

> in just those places where, though still concerned with reality, we are impelled by certain technical complications to mention sentences. Here the truth predicate serves, as it were, to point through the sentence to the reality; it serves as a reminder that though sentences are mentioned, reality is still the whole point.[52]

It makes sense, on this view, to talk of the laws of logic as the laws of truth and it makes sense to think that any general account of the logical laws must be metatheoretic. How close is this to Frege's view?

Some differences between the contemporary versions of the logical laws and rules and Frege's versions are purely notational, but others are not. To understand these differences and their significance, it will help to look at some of the discussions from the early sections of *Basic Laws* – the sections containing Frege's introduction and

[50] Dummett, *The Logical Basis of Metaphysics* (Cambridge, Mass.: Harvard University Press, 1991), p. 23.

[51] Or, for those who favor a model-theoretic view, a relation between sentences and interpretations. Since Frege objects to viewing Begriffsschrift expressions as subject to multiple interpretations (see, e.g., 'Foundations of geometry', vol. II, p. 384), a contemporary version of Frege's view would be one on which truth is a property. One might think that the view stated here is already far from Frege's since, especially in his later work, Frege characterizes truth as something that applies to thoughts rather than sentences. However, there is a relevant property of sentences – not that of truth, but that of expressing a truth. This issue does not affect the argument that follows.

[52] Quine, *Philosophy of Logic*, p. 11.

defence of the second version of his new logic.[53] As we have seen, a metatheoretic justification of *modus ponens* involves both the use of schemata to generalize over linguistic entities and a truth predicate. Frege's explanation, in its entirety, is:

> for if Γ were not the True, then since Δ is the True [*das Wahre ist*] Δ → Γ would be the False.[54]

Does Frege use a truth predicate? The only candidate for a truth predicate in the above passage is the expression 'is the True' (*das Wahre ist*). To see whether this is simply a peculiarly worded truth predicate, we need to look at Frege's use of the expressions 'the True' and 'the False'.[55]

Frege introduces the True and the False in order to make out his claim that a concept is a sort of function.[56] But why take concepts to be functions? To define a function is to indicate what values it has for each argument. And a concept definition does not seem to give values for arguments but, rather, an indication of what falls under the concept. However, we might think of a concept as a function that gives us something for each object – either the answer 'true' or the answer 'false'. Taking this a bit further, Frege writes, of concepts, 'I now say: "the value of our function is a truth-value", and distinguish between the truth-values of what is true and what is false. I call the first, for short, the True; and the second, the False.'[57] Concept expressions are predicates. Thus the expression for the value a concept has for a particular object will be a sentence. For example, '2 is

[53] I focus solely on the second (*Grundgesetze*) version of Frege's logic. No argument is needed about the first (*Begriffsschrift*) version, since there is no candidate for a truth predicate there. Frege does not use the term 'wahr' but, rather, 'bejaht' (affirms) and 'verneint' (denies) and, on occasion, a variety of other terms (e.g., 'stattfindet'). 'Bejahen' is not a truth predicate. It is not applied to sentential expressions but, rather, to sentential expressions prefixed by the judgement stroke.

[54] *Gg*, vol. I, p. 25. For convenience I use the arrow rather than Frege's actual symbols: the horizontal combined with the condition stroke.

[55] David Bell has pointed out to me, in conversation, that 'is the True' is indisputably a truth predicate in this sense: it is a predicate whose only topic is truth. However, what is at issue here is whether 'is the True' is the sort of predicate used in contemporary semantics or metatheory. I use the expression 'truth predicate' to describe a predicate that is meant to hold either of (all and only) true sentences or of (all and only) true thoughts.

[56] This, of course, is a long story. For an account see chapter 5 of J. Weiner, *Frege Explained* (Chicago: Open Court, 2004).

[57] 'Function and concept', p. 13.

a prime number' is an expression for the value the concept *prime number* has for 2. Since 2 *is* a prime number, '2 is a prime number' designates the True, as do all other true sentences. Similarly, all false sentences designate the False. Frege's strategy for assimilating concepts to functions is to assimilate sentences to proper names.

As we saw earlier, the technique of semantic ascent is needed in contemporary logic because there is an obstacle: we cannot generalize over slots occupied by sentences because sentences are not proper names. But sentences *are* proper names for Frege. And, consequently, there is no such obstacle for Frege. An upshot is that Frege has no need for one of the essential elements of a metatheoretic soundness proof: a truth predicate. And there is no truth predicate (that is, no predicate that holds of true sentences) in Frege's discussion of *modus ponens*.[58] For 'is the True' is not a predicate that holds of all true sentences.

To see why, consider Frege's statements about sentences and proper names. In 'On concept and object', Frege writes, 'a name of an object, a proper name, is quite incapable of being used as a grammatical predicate'.[59] It is not, Frege continues, that we cannot use predicates in which a proper name follows 'is' (for example, 'is Venus'). It is that in these predicates, 'is' is not the copula but, rather, the 'is' of identity. Since the True is an object, 'the True' is an object name. It follows that the 'is' in 'is the True' is the 'is' of identity. That is, the predicate 'is the True' means the same as 'is identical to the True'.

These views are repeated in *Grundgesetze*. Frege introduces the truth-values (the True and the False) as objects.[60] The view that the 'is' in the predicate 'is the True' is the 'is' of identity, comes out the use of the predicate 'is the True' to explain the horizontal. The horizontal is offered as a Begriffsschrift translation of 'is the True'. Frege writes,

> —Δ
> is the True [*das Wahre ist*] if Δ is the True; on the other hand it is the false if Δ is not the True [*nicht das Wahre ist*].[61]

Moreover, the 'is' in 'is the True' must be the 'is' of identity. For he continues,

[58] There is also, I shall argue (see footnote 69), no use of a predicate that holds of true thoughts.

[59] 'On concept and object', p. 193.

[60] *Gg*, vol. I, p. 7.

[61] *Ibid.*, p. 9.

Accordingly,

—ξ

is a function whose value is always a truth-value – or by our stipulation, a concept. Under this concept falls the True and only the True.[62]

That is, 'is the True' is a predicate that holds of one object (the True) and no other.[63] Thus, since there are distinct true sentences, the predicate 'is the True' cannot hold of all true sentences.[64] And it is, as we have seen, 'is the True' rather than 'means (or denotes) the True'(*bedeutet das Wahre*) that appears in Frege's discussion of *modus ponens*.[65]

What of the capital Greek letters that appear in Frege's defence of this rule? Are they used as metalinguistic variables? A moment's thought should show that they cannot be – since 'is the True' cannot hold of true sentences. But if the capital Greek letters that appear in Frege's discussion of *modus ponens* are not to be understood as generalizing over linguistic expressions, how are they to be understood? The quick answer is that the generalization involved is no

[62] *Ibid.*, pp. 9–10.

[63] Jamie Tappenden argues that this assertion about the meaning of 'is the True' is unjustified. See his 'Metatheory and mathematical practice in Frege', *Philosophical Topics*, 25 (1997). For a response to Tappenden, see Weiner, 'Semantic descent'.

[64] Thus, e.g., were the following correct:

'1 + 1=2' = the True

and

'2 < 5' = the True

we could infer that

'1 + 1 = 2' = '2 < 5'

– i.e., that the sentences are the same. Of course the statements set off above are not correct, on Frege's view. Rather, on his view,

(1 + 1 = 2) = the True

and

(2 < 5) = the True.

The consequence that (1 + 1 = 2) = (2 < 5) is one that Frege embraces. The same argument also shows that 'is the True' is not a predicate that holds of true thoughts (if it were, there would be only one true thought). This is not to say that there is never any use in *Grundgesetze* of a predicate that is meant to hold of true sentences. Frege does use such a predicate; it is *bedeutet das Wahre*. But this predicate does not appear in his discussions of his rules of inference and logical laws.

[65] One might suspect that Frege was simply not as careful as he might have been. After all, Frege *could have* used the predicate that is supposed to hold of all true sentences in his discussion of *modus ponens*. Perhaps Frege simply did not notice the difference between 'is the True' and 'means (or denotes) the True'. For an argument that this is not so, see Weiner, 'Semantic descent'.

different from any generalization over objects. To see this, consider, again, Frege's introduction of the horizontal. To define a first-level concept is to indicate, for each object, whether or not it holds of that object. By telling us that the horizontal names a concept that holds of the True and only the True, Frege does just that – both for objects named by sentences and for objects not named by sentences. He then goes on to say what the expressions '$—2$', '$—2^2 = 4$', and '$—2^2 = 5$' name. If Δ is not the True, $—\Delta$ is the False. Thus, given 2 is not the True, $—2$ is the False. Similarly, since the Moon is not the True, — (the Moon) is the False. That is, the capital Greek letters that appear in Frege's statements are not special metalinguistic variables. The generalization in the statements in which they appear is generalization over all objects.[66]

Thus, Frege's discussion of why *modus ponens* is a good rule, unlike the kind of metatheoretical justification that appears in contemporary soundness proofs, exploits no truth predicate and no metalinguistic variables. But the rule itself differs only notationally from the contemporary rule. Frege's statements of his laws, in contrast, are different from contemporary laws. Because his actual symbols are difficult to print, I will continue using the contemporary arrow, rather than Frege's condition stroke, in the discussion of this rendering, but I will now add some of the requisite horizontals. Frege's assertion of Basic Law I looks something like this:

$$\vdash (— a \rightarrow (— b \rightarrow — a)).$$[67]

But Frege's rendering, unlike the contemporary rendering, is not to be understood as a metatheoretic claim about infinitely many logical laws: a claim that '$— (— a \rightarrow (— b \rightarrow — a))$' *turns into* a true Begriffsschrift sentence whenever appropriate expressions are substituted for '*a*' and '*b*'. Frege's Basic Law I is a single law directly expressible

[66] However, sometimes Frege *does* use capital Greek letters as metalinguistic variables. Whenever he uses these symbols as metalinguistc variables (see, e.g., the *Grundgesetze* introduction of the identity sign), he uses quotation marks and a predicate, *bedeutet das Wahre*, which is meant to hold of linguistic expressions. In contrast, when he does not (e.g., in the passage quoted above), he uses no quotation marks and the predicate, 'is the True' is meant to hold of non-linguistic objects.

[67] The vertical line that begins this expression is Frege's judgement stroke. Although it would take us too far afield to discuss his use of this expression in detail here, one feature of its use is that expressions like this expression of Basic Law I that are preceded by a judgement stroke are universally quantified. The actual quantifiers inserted in the expressions below simply represent in more familiar form something that is actually in Frege's notation.

in Begriffsschrift. The law is simply a universal generalization 'for any a and b …'. A more revealing rendering of the content of Frege's law, using the peculiar notation I have introduced above, might be:

$$\vdash (a)\,(b)\,(\,—\,a \rightarrow (\,—\,b \rightarrow —\,a)).$$

Given this machinery, it should not be particularly surprising that no truth predicate appears in Frege's discussion of Basic Law I.

Given Frege's assimilation of sentences to proper names, there is no need for semantic ascent in the discussion of the justification of the basic logical laws and rules.[68]

WHY AVOID USING A TRUTH PREDICATE?

It would obviously be anachronistic to read Frege's work as offering a critique of the conception of logic on which metatheory plays a central role. But there are reasons to think that he was actively seeking a means to minimize use of the predicate 'is true' in the discussions of the justification of his logical laws and rules. Given the absence of a need for (or conception of) semantic ascent, if a truth predicate (which holds of sentences) plays a central role in an account of the justification of primitive logical laws, then logic would seem not to have the requisite generality. Its subject matter would appear to be, not everything thinkable, but only the limited domain of the linguistic. Of course, appearances can be misleading. After all, Frege thinks that the laws of arithmetic, which seem to express the peculiarities of a restricted domain, are logical laws. On the other hand, the fact that these laws seem to express peculiarities of a restricted domain is part of Frege's motivation for undertaking his project. One aim of Frege's proofs is to unmask the truths of arithmetic – to exhibit their true nature. Given the

[68] This is not to say that Frege is offering a strategy for eliminating the truth predicate from natural language. As I will argue shortly, a natural language truth predicate is useful for Frege's purposes. Nor is there reason to believe that Frege would object to the use of a truth predicate in a systematic science. But such a science would not be logic. In particular, as I argue in 'Semantic descent', the 'new science' Frege discusses in 'On the foundations of geometry' is not logic. Moreover, although Frege uses the expressions 'denotes the True' and 'denotes the False' throughout the early sections of *Grundgesetze*, with only two exceptions, he completely avoids them when he is talking about the justification of his basic laws or rules.

importance of this sort of unmasking, it should be no surprise that Frege would want to avoid justifications of the rules of inference and basic laws of Begriffsschrift that seem to involve truths about the specific domain of the linguistic. The laws of truth should be laws that clearly do hold everywhere.

What, then, are we to make of Frege's statements that our conception of the laws of logic is connected with how one understands the word 'true';[69] that the laws of logic are the laws of truth?[70] Taken in isolation, these statements suggest that the laws of logic either are, or are justified by, general statements or laws in which the predicate 'is true' appears. However, there is no such suggestion in the context in which these statements appear. When Frege says in *Grundgesetze* that the laws of logic are the laws of truth, he says this by way of warding off the interpretation of the laws of logic as psychological laws; the laws in accord with which we think. Thus he writes, 'I understand by "laws of logic" not psychological laws of takings-to-be-true, but laws of truth.'[71] They are 'guiding principles for thought in the attainment of truth'.[72] But this does not distinguish laws of logic from laws of other sciences. The laws of geometry and physics, Frege says, are also laws of thought in this sense.[73] They differ from laws of logic only in applying over more limited domain. The laws of geometry, for example, are guiding principles for thought in the attainment of truth about the peculiarities of what is spatial. The laws of logic, in contrast, are guiding principles for thought in all domains, they are laws that hold everywhere: or, simply, the laws of truth.[74]

But how, if not via metatheoretic proof, can he convince us that these laws are both true and universal? Consider, again, the law that every object is identical to itself. Frege says it is impossible for us to reject this law. And it is evident that this law holds, not just over the limited domain of some special science, but over everything. If so, to see that a Begriffsschrift proposition expresses this law is to see that it expresses a logical law. The same should hold for

[69] *Gg*, vol. I, pp. xiv–xv.
[70] *Ibid.*, p. xvi.
[71] *Ibid.*, p. xvi.
[72] *Ibid.*, p. xv.
[73] *Ibid.*, p. xv, see also, *PW*, pp. 145–6.
[74] See, e.g., *Gg*, vol. I, p. xv; see also *PW*, pp. 3, 128.

any primitive logical law we can identify. Similarly, one might suppose that anyone who understands Begriffsschrift will simply recognize the Begriffsschrift rules as correct rules of logic. Frege seems to have thought so when he wrote *Begriffsschrift*. For all he says there in defense of *modus ponens* is that its correctness is apparent from his explanation of the condition stroke. A lucid introduction of Begriffsschrift should suffice to convince the reader that the primitive Begriffsschrift laws and rules are logical laws and rules.

This is not to say that Frege has no metatheoretic perspective in any sense at all. It would be ridiculous to suggest that there is any way to introduce an artificial language such as Begriffsschrift without using everyday natural language to assign meanings to its symbols. Insofar as natural language discussion of Begriffsschrift belongs to metatheory, there can be no question that Frege's logic involves metatheory. Moreover, Frege certainly makes some arguments in natural language about the characteristics of his formal system.[75] These are metatheoretic arguments. But these arguments are no part of a foundation for logic – the foundation is simply the primitive logical laws and rules.

But, if this is so, what is the purpose of the (often elaborate) discussions of the truth of his basic laws in the early sections of *Grundgesetze*? Why would he not simply introduce the primitive terms, list the axioms and rules and get immediately to work on the Begriffsschrift proofs?

THEORY AND ELUCIDATION

Frege frequently remarks that the meaning of primitive terms can only be communicated via hints or elucidations. For example,

> Since definitions are not possible for primitive elements, something else must enter in. I call it elucidation. It is this, therefore, that serves the purpose of mutual understanding among investigators, as well as of the communication of the science to others. We may relegate it to a propaedeutic. It has no place in the system of a science; in the latter, no conclusions are based on it.[76]

[75] E.g., in his unpublished articles about Boole's logical notation and *Begriffsschrift*, Frege argues that Begriffsschrift is superior to Boole's notation. I am indebted to Ian Rumfitt for bringing up the issue of Frege's discussions of Boole.

[76] 'Foundations of geometry II', p. 301, see also *Gg*, vol. I, p. 4.

But it is not much help simply to provide a label. What is elucidation?

Definitions are statements in a theory that are designed to communicate the meanings of terms. There are rules for a properly formed definition, rules designed to guarantee that the definition fixes the meaning of a term. Thus it is tempting to suppose that there will also be rules designed to guarantee (or least make probable) the success of elucidation – albeit different and, perhaps, less reliable rules. But Frege also describes some elucidations as hints. After saying that we must rely on elucidation in our introduction of primitive terms, he also says 'we must be able to count on a little goodwill and cooperative understanding, even guessing'.[77] There are, then, no rules for successful elucidation.

Nor is elucidation a technique for effective communication of the meaning of primitive terms. There is no such technique – as Frege makes clear in some of his discussions of logically simple notions. An example is his discussion of the notions of concept and object. Frege appears unperturbed by his recognition of the apparently paradoxical nature of his remarks and goes on, notoriously, to claim that some of his statements must either be false or miss his thought and to ask his readers to grant him a grain of salt. This would be mysterious were we to interpret Frege as giving an argument designed to establish, as its conclusion, one of his general remarks about the nature of concepts. But these remarks are, rather, part of Frege's elucidatory attempt to communicate an understanding of the notions of concept and object. If we can accept Frege's characterization of elucidations as hints, his odd attitude is explicable.[78]

This is not to say that elucidation *must* consist of apparently paradoxical utterances or failed attempts to express the inexpressible. Indeed, most of Frege's elucidatory remarks are entirely unproblematic. There is nothing paradoxical about Frege's claim that the singular definite article indicates an object-name. Nor is this a failed attempt to express the inexpressible. But this claim, like the apparently paradoxical claims about what it is to be concept, is no part of Frege's systematic science. There is no Begriffsschrift expression for predicating objecthood, hence no logical law that

[77] 'Foundations of geometry II', p. 301.

[78] For an argument, see J. Weiner, *Frege in Perspective* (Ithaca, N.Y.: Cornell University Press, 1990, pp. 246–59; see also Weiner, *Frege Explained*, pp. 103–14.

tells us what it is to be an object. Frege's introduction of the term 'elucidation' is meant to highlight the difference between these attempts to communicate the meanings of terms and actual definitions. Statements that appear in discussions belonging to the propaedeutic of a theory are to be distinguished from actual propositions of the theory.

What does this understanding of the role of elucidation in Frege's project tell us about his discussions of the justification of his primitive laws and rules? Consider, first, the laws. Frege writes, 'The questions why and with what right we acknowledge a law of logic to be true, logic can answer only by reducing it to another law of logic. Where that is not possible, logic can give no answer.'[79] As we have seen, Frege indicates that we cannot doubt primitive logical laws; that all we need in order to see that the primitive laws on which he relies are true is an understanding of the Begriffsschrift terms used in their statement. He also claims, later in his career, that the denial of a logical law can appear, if not nonsensical, at least absurd.[80] In this case logic, it would seem, can (and need) give no answer. All we need, it seems, is elucidation.

But can Frege's actual discussions be elucidations? After all, they do not introduce primitive terms but, rather, are attempts to show that the complex expressions he uses to express his primitive laws express truths. Moreover, they seem to have the character of arguments, not hints.[81] A closer look at Frege's writings shows us that he restricts elucidation neither to the introduction of primitive terms nor to having the character of hints. There is at least one complex function term that Frege both defines and offers 'a

[79] *Gg*, vol. I, p. xvii.

[80] 'Compound thoughts', p. 50.

[81] Some of the discussion below is a response to these objections from Stanley, 'Truth and Metatheory in Frege'. I do not have space here to discuss Stanley's claim that the discussions of section 31 of *Grundgesetze* are meant as metatheoretic proofs. This much is clearly right: section 31 contains natural language discussions about Begriffsschrift that have the character of argument. But if that is all that is meant by 'metatheory', then metatheory includes what Frege calls 'elucidation' (see the remarks below about section 34) and need not satisfy the standards we apply to proofs. I argue in 'Section 31 revisited: Frege's elucidations', in E. Reck (ed.), *From Frege to Wittgenstein: Perspectives on Early Analytic Philosophy* (Oxford: Oxford University Press, 2002) that, while these discussions do make sense as elucidations, they do not make sense as metatheoretic proofs.

few elucidations' to help his readers understand the term.[82] These elucidations consist of perfectly straightforward natural language arguments – arguments that are indistinguishable from natural language proofs – that are meant to show us what value the defined function has for different arguments. What is characteristic of elucidation comes out in the next section, where Frege writes, '[O]ur elucidation could be wrong in other respects without placing the correctness of those proofs in question; for only the definition itself is the foundation for this edifice.'[83] The character of the discussions and arguments in Frege's writings that play elucidatory roles varies dramatically – from apparently paradoxical remarks that (Frege himself says) must either be false or miss his thought to elaborate arguments that might easily be (and sometimes ultimately are) expressed in Begriffsschrift. What marks a discussion as elucidatory is neither its form nor its content but, rather, its role in the project. The mark of elucidation is its contribution to the propaedeutic.

TRUTH

Let us return, finally, to the issue of the truth predicate. Frege makes at least one general statement about truth that seems to be neither a gloss on the meaning of Begriffsschrift terms nor assimilable to logical laws. This is the statement with which we have been concerned for most of this paper: that a sentence can have truth-value only if each of its constituents has *Bedeutung*. How is this to be understood? It purports to distinguish between sentences – those that do, and those that do not, have truth-values. Were this statement a part of a theory of truth (or a law of truth), one would expect it to be applicable to particular languages. And as we saw earlier, given the criteria a term must satisfy if it is to have *Bedeutung*, one upshot of this law is that virtually no sentence of natural language has truth-value. If our interest is in natural language, the purported distinction between sentences that do, and sentences that do not, have truth-values does not do any work. The same holds for logically perfect language. For a logically perfect language must be

[82] See *Gg*, vol. I, §34. This is somewhat obscured in Furth's translation, where 'mögen einige Erläuterungen nicht überflüssig sein' is rendered 'a few explanatory remarks are in order'.

[83] *Ibid.*, vol. I, §35.

constructed, Frege tells us, so that each of its terms has *Bedeutung*. All sentential expressions of a logically perfect language will have truth-values. Thus, as a statement of a theory about language and truth, it seems either platitudinous or wrong.

The point of this statement, I have been arguing, is that it is part of Frege's articulation of his standards for introducing a systematic science of logic; part of his articulation of a regulative ideal. And we can see from this why, in spite of the fact that there is no need for a truth predicate in *Grundgesetze*, Frege would not want to eliminate the truth predicate from natural language. For there is a continuing role for a truth predicate to play in natural language – even on Frege's view. Our everyday concerns and concepts play a continuing role in introducing new questions to be addressed by scientific research. It will always be the case that most science is in its early stages and not yet systematizable. Science in its early stages requires a language with logical defects – a language, for example, in which it is possible to use a predicate that has no *Bedeutung*. But on Frege's view the further development of science requires progressively more rigour and precisification. It requires, in particular, fixing the *Bedeutung* of the predicate in question. Thus to say that a sentence has a truth-value only if each of its constituents has *Bedeutung* is to state part of a regulative ideal that, Frege thinks, guides scientific research. The statement of, and attention to, this regulative ideal will continue to have importance for as long as our everyday concerns and concepts motivate the formulation and pursuit of new scientific projects.[84]

[84] My thanks to David Bell, Gary Ebbs, Mark Kaplan, Michael Liston and Thomas Ricketts for comments at various stages. Versions of parts of this paper were read to the conference on Truth in Frege at the University of London, the Society for Analytic Philosophy at Erlangen-Nürnberg and the philosophy departments at Friedrich-Schiller Universität Jena, Indiana University, Leipzig University, Notre Dame University, Sheffield University and the University of Illinois at Urbana-Champaign. I am indebted to members of these audiences. I am indebted to the American Philosophical Society and the Bogliasco foundation for support and to the Philosophy Programme, School of Advanced Study at the University of London. This paper contains short versions of arguments from 'What's in a numeral? Frege's answer' (*Mind*, April 2007) and 'Semantic descent' (*Mind*, April 2005). The arguments are reprinted with permission.

WARREN GOLDFARB

3 Frege's conception of logic

Frege is of course an important progenitor of modern logic. The technical advances he made were comprehensive. He clearly depicted polyadic predication, negation, the conditional and the quantifier as the bases of logic; and he gave an analysis of and a notation for the quantifier that enabled him to deal fully and perspicuously with multiple generality. Moreover, he argued that mathematical demonstrations, to be fully rigorous, must be carried out using only explicitly formulated rules, that is, syntactically specified axioms and rules of inference.

Less clear, however, is the philosophical and interpretive question of how Frege understands his formalism and its purposes. Upon examination, it appears that Frege had a rather different view of the subject he was creating than we do nowadays. In lectures and seminars as far back as the early 1960s, Burton Dreben called attention to differences between how Frege viewed the subject matter of logic and how we do. The point has been taken up by several commentators, beginning with Jean van Heijenoort.[1] The technical development historically required to get from a Fregean conception

[1] In Jean van Heijenoort, 'Logic as calculus and logic as language', *Synthese*, 17 (1967), pp. 324–30. Other discussions of this difference in viewpoint are contained in Burton Dreben and Jean van Heijenoort, 'Introductory note to 1929, 1930, and 1930a', in *Kurt Gödel, Collected Works*, vol. I, ed. S. Feferman *et al.* (New York: Oxford University Press, 1986), pp. 44–59; Jaakko Hintikka, 'On the development of the model-theoretic viewpoint in logical theory', *Synthese*, 77 (1988), pp. 1–36; and Thomas Ricketts, 'Objectivity and objecthood: Frege's metaphysics of judgement', in L. Haaparanta and J. Hintikka (eds.), *Frege Synthesized: Essays on the Philosophical and Foundational Work of Gottlob Frege* (Dordrecht: D. Reidel, 1986), pp. 65–95. Hintikka gives a variant version of what I call below the schematic conception of logic.

to our own was discussed in my 'Logic in the twenties: The nature of the quantifier'.[2] Yet there is currently little appreciation of the philosophical import of these differences, that is, the role in Frege's philosophy that his conception of logic, as opposed to ours, plays. Indeed, some downplay the differences and assign them no influence on or role in the philosophy. Thus Dummett says only that Frege was 'impeded' from having the modern view by a particular way of looking at the formulas of his Begriffsschrift.[3] I want to urge on the contrary that Frege's conception of logic is integral to his philosophical system; it cannot be replaced with a more modern conception without serious disruptions in that system. The reasons for this will, I hope, be instructive about the roots of Frege's philosophizing.

I

The first task is that of delineating the differences between Frege's conception of logic and the contemporary one. I shall start with the latter. Explicit elaborations of it are surprisingly uncommon. (In most writing on issues in philosophical logic, it is implicitly assumed; yet many textbooks gloss over it, for one pedagogical reason or another.) There are various versions; I will lay out the one formulated by Quine in his textbooks[4] as it seems to me the clearest.

On this conception, the subject matter of logic consists of logical properties of sentences and logical relations among sentences. Sentences have such properties and bear such relations to each other by dint of their having the logical forms they do. Hence, logical properties and relations are defined by way of the logical forms; logic deals with what is common to and can be abstracted from different sentences. Logical forms are not mysterious quasi-entities, à la Russell. Rather, they are simply schemata: representations of the composition of the sentences, constructed from the logical signs (quantifiers and truth-functional connectives, in the standard case)

[2] W. Goldfarb, 'Logic in the twenties: The nature of the quantifier', *Journal of Symbolic Logic*, 44 (1979), pp. 351–68.

[3] M. Dummett, *The Interpretation of Frege's Philosophy* (London: Duckworth, 1981), p. 151.

[4] W. V. Quine, *Elementary Logic* (Boston: Ginn, 1941) and *Methods of Logic* (New York: Holt, 1950).

using schematic letters of various sorts (predicate, sentence, and function letters). Schemata do not state anything and so are neither true nor false, but they can be interpreted: a universe of discourse is assigned to the quantifiers, predicate letters are replaced by predicates or assigned extensions (of the appropriate r-ities) over the universe, sentence letters can be replaced by sentences or assigned truth-values. Under interpretation, a schema will receive a truth-value. We may then define: a schema is *valid* if and only if it is true under every interpretation; one schema *implies* another, that is, the second schema is a *logical consequence* of the first, if and only if every interpretation that makes the first true also makes the second true. A more general notion of logical consequence, between sets of schemata and a schema, may be defined similarly. Finally, we may arrive at the logical properties or relations between sentences thus: a sentence is logically true if and only if it can be schematized by a schema that is valid; one sentence implies another if they can be schematized by schemata the first of which implies the second.

The notion of schematization is just the converse of interpretation: to say that a sentence can be schematized by a schema is just to say that there is an interpretation under which the schema becomes the sentence. Thus, a claim that a sentence R implies a sentence S, that is, that S is a logical consequence of R, has two parts, each of which uses the notion of interpretation: it is the assertion that there are schemata R^* and S^* such that

(1) R^* and S^*, under some interpretation, yield R and S; and

(2) under no interpretation is R^* true and S^* false.

This is often called the Tarski-Quine definition, or (in the Tarskian formulation) the model-theoretic definition, of logical consequence.[5] It is precise enough to allow the mathematical investigation of the notion. For example, using this notion of logical consequence, we can frame the question of whether a proposed formal system is sound and complete, and this question may then be treated with mathematical tools. Better put, though, we should say that the definition

[5] Tarski's formulation in 'On the concept of logical consequence', in his *Logic, Semantics, Metamathematics* (Oxford: Oxford University Press, 1956, originally published 1935), pp. 409–20, does not introduce schemata, but obtains the same effect for the formalized languages he treats by disinterpreting the non-logical vocabulary so as to allow for arbitrary reinterpretations.

is capable of being made precise. For the definition quantifies over all interpretations. This is a set-theoretic quantification; hence, complete precision would require a specification of the set theory in which the definition is to be understood. (However, it turns out that for implications between first-order schemata, the definition is rather insensitive to the choice of set theory. The same implications are obtained as long as the set theory is at least as strong as a weak second-order arithmetic that admits the arithmetically definable sets of natural numbers.)[6] (As an aside, let me note that this explication of logical consequence has recently come under attack in John Etchemendy's *The Concept of Logical Consequence*.[7] Etchemendy argues that, if *S* is a logical consequence of *R*, then there is a necessary connection between the truth of *R* and the truth of *S*, and the Tarski-Quine definition does not adequately capture this necessity. Of course, neither Tarski nor Quine would feel the force of such an attack, since they both reject the cogency of the philosophical modalities. Moreover, it is only the Tarski-Quine characterization of logical consequence in terms of various interpretations of a schematism that makes the notion of logical consequence amenable to definitive mathematical treatment.)

On this *schematic conception* of logic, the formal language of central concern is that of logical schemata. Pure logic aims at ascertaining logical properties and logical relations of these formulas, and also at demonstrating general laws about the properties and relations. Applied logic, we might say, then looks at sentences – of one or another formal language for mathematics or science or of (regimented versions of) everyday language – to see whether they may be schematized by schemata having this or that logical property or relation. Thus, there is a sharp distinction between logical *laws*, which are at the metalevel and are about schemata, and logical *truths*, which are particular sentences that can be schematized by valid

[6] That the arithmetical sets are enough for implications between schemata was shown in David Hilbert and Paul Bernays, *Grundlagen der Mathematik*, vol. II (Berlin: Springer, 1939), p. 252. The same proof shows that, for an infinite set of schemata, we need no more than the sets arithmetically definable from that set, but that we may need more than just the arithmetical sets themselves, however, was noted in George Boolos, 'On second-order logic', *Journal of Philosophy*, 72 (1975), pp. 509–27.

[7] John Etchemendy, *The Concept of Logical Consequence* (Cambridge, Mass.: Harvard University Press, 1990).

schemata. The pivotal role in this conception of schemata, that is, of uninterpreted formulas that represent logical forms, gives a specific cast to the generality of logic. Logic deals with logical forms, which schematize away the particular subject matter of sentences. Thus logic is tied to no particular subject matter because it deals with these 'empty' forms rather than with particular contents.

Such a schematic conception is foreign to Frege (as well as to Russell). This comes out early in his work, in the contrast he makes between his Begriffsschrift and the formulas of Boole: 'My intention was not to represent an abstract logic in formulas, but to express a content through written signs in a more precise and clear way than it is possible to do through words.'[8] And it comes out later in his career in his reaction to Hilbert's *Foundations of Geometry*: 'The word "interpretation" is objectionable, for when properly expressed, a thought leaves no room for different interpretations. We have seen that ambiguity [*Vieldeutigkeit*] simply has to be rejected.'[9] There are no parts of his logical formulas that await interpretation. There is no question of providing a universe of discourse. Quantifiers in Frege's system have fixed meaning: they range over all items of the appropriate logical type (objects, one place functions of objects, two place functions of objects, etc.) The letters that may figure in logical formulas, for example, in '$(p \& q \rightarrow p)$' are not schematic: they are not sentence letters.[10] Rather, Frege understands them as *variables*. Here they are free variables, and hence in accordance with Frege's general rule the formula is to be understood as a universal closure, that is, as the universally quantified statement '$(\forall p)(\forall q)(p \& q \rightarrow p)$.' Similarly, logical formulas containing one-place function signs are to be understood not schematically, but as generalizing over all functions.

On Frege's conception the business of logic is to articulate and demonstrate certain true general statements, the logical laws. '$(\forall p)(\forall q)(p \& q \rightarrow p)$' is one; it states a law, we might say, about all objects. Similarly, '$(\forall F)(\forall G)(\forall H)((\forall x)(Fx \rightarrow Gx) \& (\forall x)\ (Gx \rightarrow Hx) \rightarrow (\forall x)(Fx \rightarrow Hx))$'

[8] 'Über den Zweck der *Begriffsschrift*', *Jenaische Zeitschrift für Naturwissenschaft*, 16, supplement (1882), pp. 1–10, at p. 1. Throughout this chapter the page references for Frege's works refer to their original publication.

[9] 'Über die Grundlagen der Geometrie', *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 15 (1906), pp. 293–309, 377–403, 423–30 (hereafter cited as Frege 1906), at p. 384.

[10] Throughout this paper I use modern logical notation rather than Frege's.

is a law about all functions.[11] The business of pure logic is to arrive at such laws, just as the business of physics is to arrive at physical laws. Logical laws are as descriptive as physical laws,[12] but they are more general. Indeed, they are supremely general; for, aside from variables, all that figure in them are the all-sign, the conditional and other signs which are not specific to any discipline, but which figure in discourse on any topic whatsoever. Notions of the special sciences first appear when we apply logic. In applied logic, we infer claims that contain more specialized vocabulary on the basis of the laws of pure logic. For example, in applied logic we might demonstrate, 'If Cassius is lean and Cassius is hungry, then Cassius is lean'; or, 'If all whales are mammals and all mammals are vertebrates, then all whales are vertebrates.' These statements may be inferred from the logical laws given at the beginning of this paragraph. Here we also see a typical situation, that these specialized statements are inferred from the logical law by instantiation of universal quantifiers.

On Frege's *universalist conception*, then, the concern of logic is the articulation and proof of logical laws, which are universal truths. Since they are universal, they are applicable to any subject matter, as application is carried out by instantiation. For Frege, the laws of logic are general, not in being about nothing in particular (about forms), but in using topic-universal vocabulary to state truths about everything.

The question arises immediately of how different these conceptions actually are. They can look very close. Both take pure logic to be centrally concerned with generality. Generality is captured in the schematic conception by definitions that invoke all interpretations of the given schemata, and in the universalist conception by universal quantifiers with unrestricted ranges. In the schematic conception, logic is applied by passing from schemata to sentences that are particular interpretations of them; in the universalist conception, applications are made by instantiating the quantified variables of a general law. Given these close parallels, it is no wonder that

[11] Here, in using modern notation, I am eliding a nicety required by Frege's quantifying over all functions, not just all concepts, namely, his use of the horizontal.

[12] *Gg*, vol. I, p. xv.

many logicians and philosophers would be inclined to minimize the distinction between the two conceptions.

Parallels are not identities, however, and there are philosophically important ways that the conceptions differ. First and most obviously, the schematic conception is metalinguistic. The claims of logic are claims about schemata or about sentences, and thus logic concerns features of discourse. In contrast, on the universalist conception logic sits squarely at the object level, issuing laws that are simply statements about the world. What logical laws describe are not phenomena of language or of representation. As Russell put it, 'Logic is concerned with the real world just as truly as zoology, though with its more abstract and general features.'[13] This difference will have consequences for the philosophical characterization of logic. For example, the universalist conception leaves no room for the notion that logic is without content; the laws of logic, although very general, have to be seen as substantive. Indeed, in the *Tractatus*, Wittgenstein breaks with the universalist conception in order to arrive at a view in which the propositions of logic are empty. Even if Wittgenstein's characterization of logic is rejected, the metalinguistic conception will inevitably make the nature of discourse, or of our representations, the focus of any account of logic. A sharp sense of this can be obtained by contrasting the remark of Russell's just cited with this one of Dummett's, made unselfconsciously and with no argument at all, at the start of laying out his own metaphysics: 'Reality cannot be said to obey a law of logic; it is our thinking about reality that obeys such a law or flouts it.'[14] On Frege's view, as on Russell's, it is precisely reality that obeys the laws of logic.

Indeed, the universalistic conception is an essential background to many of Frege's ontological views. Frege took not just proper names but also sentences and predicates to be referring expressions, that is, to have a reference; in the latter case, the referents were of a different logical sort from those of proper names and sentences. From many contemporary viewpoints, it is odd to think of sentences as names at all; and if predicates are thought to refer, it would be to properties

[13] B. Russell, *Introduction to Mathematical Philosophy* (London: Allen & Unwin, 1919), p. 169.

[14] M. Dummett, *The Logical Basis of Metaphysics* (Cambridge, Mass.: Harvard University Press, 1991), p. 2.

or sets or some other entities that need not be sharply distinguished in logical character from the referents of singular terms.

It should be clear that the universalistic conception *demands* that sentences and predicates refer. As we have seen, for Frege the truth-functional laws look like '$(\forall p)(\forall q)(p \& q \rightarrow p)$' and will be applied by instantiating the quantifiers with sentences. For 'If Cassius is lean and Cassius is hungry then Cassius is lean' to count as a genuine instance of the law, the expressions which instantiate the quantified variables have to refer, to things that are values of the variables, just as to count as a genuine instance of '$(\forall x)(x$ is a prime number greater than $2 \rightarrow x$ is odd)' the name replacing 'x' has to refer, and what it refers to must be among the values of 'x'. (To be is to be the value of a variable as much for Frege as for Quine.) Similarly, since the laws of logic include many that generalize in predicate places, and their application requires instantiating those quantified variables with predicates, here too we are driven to take predicates as referring expressions.

In the case of sentences, it requires a further argument, based on intersubstitutivity phenomena, to conclude that what sentences refer to are their truth-values, and it requires yet other considerations to support taking the truth-values to be of the same logical type as ordinary objects. The former is pretty compelling; the latter has elicited heated objections.[15]

For predicates, however, support for the sharp distinction in logical type of the referent can come from the structure of applications of logic, on the universalist conception. If the position occupied by a predicate in a statement is taken to be generalized on directly, the distinction in logical type is apparent, since the predicate position has argument-places; and if an expression has an argument-place and so can be used in an instantiation of a quantified predicate variable, then it cannot be used to instantiate a singular term, without yielding expressions that violate the most basic rules of logical (and grammatical) syntax. Thus we see that the universalist conception demands second-order logic.[16] Indeed, it was one of Quine's avowed

[15] For example, see M. Dummett, *Frege: Philosophy of Language* (London: Duckworth, 1973), p. 180 and following. Dummett calls Frege's view on the matter a 'gratuitous blunder'.

[16] Charles Parsons canvasses an objection to Frege's conclusion in his 'Objects and logic', *Monist*, 65 (1982), pp. 491–516, at pp. 499–501. Suppose we take it that predicates are generalized not directly but only via 'nominalization,' that is, only

motivations, in developing the schematic conception, to show that logic did not require us to take there to be anything designated by the predicates in our statements.

Logic, as construed on the universalist conception, is also in back of a doctrine of Frege's that many have found puzzling, namely, that all functions be defined everywhere; for the special case of concepts, this is the requirement that concepts 'have sharp boundaries'.[17] For Frege, all quantified variables have unrestricted domain. Given this, and given that '$(\forall F)(\forall x)(Fx \vee -Fx)$' is a logical law, Frege's requirement follows at once. If something is a concept, then an expression for it can instantiate the quantifier in this law; thus we can logically derive that, for every object, either the concept holds of it or the concept does not. This is just what Frege means by 'sharp boundaries'.

II

A second important difference between the two conceptions concerns the role of a truth predicate. Clearly, the schematic conception employs a truth predicate: the definitions of validity and logical consequence talk of the truth under all interpretations of schemata.[18]

once they are transformed into names of qualities, properties, sets or the like: for example, only once '*x* is malleable' is transformed into '*x* has (the property of) malleability'. Since 'malleability' lacks argument-places, it may be taken as generalizable using a variable of the same logical type as those over objects. Parsons notes a problem with this: since the assertions with which one starts and ends will contain the unnominalized predicate, one needs a principle to underwrite the transformation, that is, a general principle a particular instance of which will be '$(\forall x)(x$ is malleable $\leftrightarrow x$ has the property of malleability)'. But any such principle, it seems clear, will have to contain a generalization directly in the predicate place occupied in this instance by 'is malleable.' Hence the Fregean conclusion stands. As Parsons notes ('Objects and logic', p. 503), there is one way around this conclusion. That is to use the method of semantic ascent and understand the principle underwriting the nominalization not as a direct generalization over what predicates refer to, but as an assertion that any assertion of a certain form may be transformed into another form *salva veritate*. To adopt this strategy is simply to give up the universalist conception, as it requires a metalinguistic principle that makes ineliminable use of a truth predicate.

[17] *Funktion und Begriff*, p. 20.

[18] The truth predicate needed is a predicate of sentences. For Frege, it was not sentences but rather thoughts (senses of sentences) that were true or false. Consequently, those who take Frege to have a form of the schematic conception treat implication as a relation among thoughts. In this version, the definition would require a truth predicate of thoughts.

Since the predicate is applied to an infinite range of sentences, it cannot be eliminated by disquotation. On the universalist conception, in contrast, no truth predicate is needed either to frame the laws of logic or to apply them. Moreover, although Frege sometimes calls logical laws the 'laws of truth', he does not envisage using a truth predicate to characterize the nature of those laws.

On the schematic conception, logic starts with the definitions of validity and consequence and goes on to pronounce that a given schema is valid or is a consequence of other schemata. Formal systems may be introduced as a means to establish such facts, but this then requires a demonstration of soundness to show that what the system produces are, in fact, validities and consequences. The introduction of a formal system also raises the (less urgent) question of completeness, of whether all validities and all implications can be obtained by means of the system. Thus it is the overarching notions of validity and consequence that set the logical agenda and provide sense to the question of how well a system for inference captures logic. On this conception, the notion of logical inference rule is posterior to that of consequence: a logical inference rule is one whose premises imply its conclusion or, in the context of a system for establishing validities only, is one that always leads from valid premises to a valid conclusion.

In Frege's universalist conception, there is no analogous characterization of what is a logical law or what follows logically from what. Frege's conception of logic is retail, not wholesale. He simply presents various laws of logic and logical inference rules, and then demonstrates other logical laws on the basis of these. He frames no overarching characteristic that demarcates the logical laws from others.[19] Consequently, the only sense that the question has of whether the laws and rules Frege presents are complete is an 'experimental' one – whether they suffice to derive all the particular results that we have set ourselves to derive. For example, at one point, Frege entertains the possibility that a failure to obtain

[19] In this Frege differs from Russell, who from the first tries to formulate necessary and sufficient conditions for a truth to be a logical truth. Russell tries to use generality as the key element of the characterization. Wittgenstein takes up the problem, but criticizes Russell's invocation of generality. He takes himself to have solved the problem with the notion of tautology; that notion gives him a general conception of the logical.

established results while developing an area of mathematics axiomatically could lead us to recognize a new logical inferential principle.[20] The closest Frege comes to providing a notion of logical consequence occurs in 'On the foundations of geometry', where he defines one truth's being logically dependent on another. The definition is: when the one can be obtained by logical laws and inferences from the other (Frege 1906, p. 423). No further characterization of logical laws and inferences is made. Thus, in direct contrast to the situation in the schematic conception, Frege's notion here rests on the provision of the logical laws and inference rules.

Now Frege does say, 'Logic is the science of the most general laws of truth.'[21] But he does not intend this as a demarcation of logic, only as a 'rough indication of the goal of logic'. As we have seen, generality and absence of vocabulary from any specialized science are, on the universalist conception, features of the logical. Frege does not attempt to give any specification of the vocabulary allowable in logic; moreover, there is no reason to think that he would take truth and absence of specialized vocabulary as sufficient for logical status.[22] Yet there is a deeper reason that his phrase gives only a 'rough indication', and that has to do with the anomalous status of 'true' when used as a predicate.

Frege repeatedly calls attention to that anomalous status. In 'Der Gedanke', he presents a regress argument to show that any attempt to define truth must fail, and concludes that 'the content of the word "true" is *sui generis* and indefinable'.[23] Both the argument and his subsequent considerations show that he does not mean simply that the notion of truth is a primitive notion, not to be defined in terms of anything more basic. After reflecting that 'I smell the scent of violets' and 'It is true that I smell the scent of violets' have

[20] 'Über die Begriffsschrift des Herrn Peano und meine eigene', p. 363.

[21] 'Logic', in *PW*, pp. 126–151, p. 128.

[22] Richard Heck suggests that Frege may have formulated the principle of countable choice to himself and found himself unable to derive it. 'The finite and the infinite in Frege's *Grundgesetze der Arithmetik*', in M. Schirn (ed.), *Philosophy of Mathematics Today* (Oxford: Oxford University Press, 1998). If so, Frege may have wondered whether that principle is an example of a truth statable without using non-logical vocabulary, but not itself a truth of logic.

[23] 'Der Gedanke', *Beiträge zur Philosophie des deutschen Idealismus*, 1 (1918), pp. 58–77 (hereafter cited as Frege 1918), at p. 60.

the same content, so that the ascription of truth adds nothing, he concludes: 'The meaning of the word "true" seems to be altogether *sui generis*. May we not be dealing here with something which cannot be called a property in the ordinary sense at all?' (Frege 1918, p. 61). In ' Introduction to Logic',[24] Frege goes farthest in suggesting that truth is not a property at all: 'If we say "the thought is true" we seem to be ascribing truth to the thought as a property. If that were so, we should have a case of subsumption. The thought as an object would be subsumed under the concept of the true. But here we are misled by language. We don't have the relation of an object to a property' (*PW*, p. 194). In 'My basic logical insights',[25] he connects the use of 'true' in characterizing logic with the idea that the ascription of truth to a thought adds nothing:

> So the sense of the word 'true' is such that it does not make any essential contribution to the thought. If I assert 'it is true that sea-water is salty', I assert the same thing as if I assert 'sea-water is salty'. This enables us to recognize that the assertion is not to be found in the word 'true' but in the assertoric force with which the sentence is uttered ... '[T]rue' makes only an abortive attempt to indicate the essence of logic, since what logic is really concerned with is not contained in the word 'true' at all but in the assertoric force with which a sentence is uttered.

Thus, rubrics like 'general laws of truth' cannot serve to give a real characterization of logic or a demarcation of the realm of the logical. The notion of truth is unavailable for the role of setting the agenda for logic. Moreover, if we take Frege's scruples seriously, it follows that the schematic conception of logic is simply unavailable to him. To formulate it, as we have seen, use has to be made of a truth predicate. That predicate figures not as a suggestive way of talking, nor as a term whose usefulness arises only from the 'imperfection of language', as Frege put it in *1915, but as a scientific term in the definitions of the most basic concepts of the discipline. Clearly, Frege would not think that legitimate.

The question then arises of whether Frege's scruples are well placed, or whether they can be dismissed as merely peripheral phenomena, with no deep systematic connections. Addressing this question requires a careful examination of the arguments Frege

[24] In Frege, *PW*, pp. 185–96.

[25] In Frege, *PW*, pp. 251–2 (hereafter cited as Frege *1915).

adduces. I shall not attempt this here; for a detailed treatment, see Thomas Ricketts' 'Logic and truth in Frege'.[26] I limit myself to mentioning the philosophical outlook which I take to be expressed in Frege's scruples about a truth predicate. It is that objective truth is not to be explained or secured by an ontological account. Such an account would take us to have a conception of things 'out there' and of their behaviours or configurations that exist independent of our knowledge, and it would depict those behaviours or configurations as being that which renders our thoughts true or false. Such an account is often ascribed to Frege, for it is just what is involved in ascribing a truth-conditional semantics to him. But this ascription is incompatible with Frege's remarks on truth. To take Frege's scruples seriously is to appreciate that there is no general notion of something's making a truth true – that is, that there is no theory of how the thoughts expressed by sentences are determined to be true or false by the items referred to in them. It is thus to put us in a position to appreciate the extraordinarily subtle view Frege can be read as unfolding. On this reading, Frege is not a realist, on the usual philosophical characterizations of that position. He is committed to the objectivity of truth and its independence of anyone's recognition of that truth, but the conception of truth here is immanent within our making of judgements and inferences, our recognitions of truth.

III

Earlier I noted that the most obvious difference between the universalist and the schematic conceptions is that in the former logic operates at the object level, whereas in the latter it operates at the meta-level. Even this by itself has consequences, and it can be used to get at an important role the universalist conception has in Frege's system.

Of course, it is important to avoid anachronism here. At the time Frege was writing, a distinction between object level and meta-level could hardly have been drawn; in fact, it was not to become clear until the 1920s. Nonetheless, we can see a precursor of the

[26] T. Ricketts, 'Logic and truth in Frege', *Proceedings of the Aristotelian Society*, supplementary volume 70 (1996), pp. 121–40.

distinction as being at issue. Many traditional logicians spoke of logic as being about the forms of judgement, which were to be obtained by abstraction from judgements. Although this conception was far from precise and traditional logic lacked the machinery to work it out, it seems clear that forms of judgement were invoked as a way of capturing the generality of logic and lack of tie to the content of individual judgements. Thus we can see here a proto-schematic conception. (This is particularly visible in Bolzano.) Frege rarely speaks of forms of judgement. It is not hard to surmise some reasons.

First, talk of forms of judgement and of abstracting from individual judgements has a dangerously psychological ring to it. The very locution 'forms of judgement' suggests that the forms are of mental acts and so are prime material for psychologistic treatment. Moreover, Frege argued vigorously against any notion of abstraction as needed to get from particulars to general notions.[27] Indeed, elimination of any role for abstraction is central not just to Frege's antipsychologism, but also to his anti-Kantianism. To eliminate abstraction is to eliminate the question, How do we attain the general? Frege replaces it with the question of the relation between the (already given) general and the particular, a question to be answered by logic.

This leads us to the second reason Frege has for discarding talk of the forms of judgement. He has no need of such talk, precisely because his devising of the quantifier gives him a rigorous tool to capture the generality that 'forms of judgement' gestures toward. The generality is directly expressed by the quantifier. The relation of general to particular is given by the logical rule of instantiation from former to latter, not by some imprecise, psychological notion of abstraction from the latter to the former.

These two reasons are not relevant to the modern schematic conception, which has found precise non-psychologistic notions to replace 'forms of judgement' and 'abstraction' and which uses quantification (in the metalanguage) to capture the desired generality. There is, however, another consideration at work in Frege that is not simply obsolete.

[27] See, for example, 'Rezension von Dr. E. G. Husserl: *Philosophie der Arithmetik I*', *Zeitschrift für Philosophie und philosophische Kritik*, 103 (1894), pp. 313–32.

Frege's conception of logic is intertwined with his notion of justification. A cornerstone of Frege's thinking is the sharp distinction between the rational basis of a claim – the truths that it presupposes or depends upon – and what we might call concomitants of thinking or making the claim: the psychological phenomena that occur when a person thinks of the claim, or believes it, or comes to accept it, the empirical conditions someone must satisfy in order to know the claim, the history of the discovery of the claim, and so on. The distinction is emphasized throughout Frege's writings, and particularly vividly in the Introduction to *Die Grundlagen der Arithmetik*. Remarks like these abound: 'Never let us take a description of the origin of an idea for a definition, or an account of the mental and physical conditions on which we become conscious of a proposition for a proof of it.'[28] The point is more general than antipsychologism, or a distinction between objective and subjective, as the following shows:

> A delightful example of the way in which even mathematicians can confuse the grounds of proof with the mental or physical conditions to be satisfied if the proof is to be given is to be found in E. Schröder. Under the heading 'Special Axiom' he produces the following: 'The principle I have in mind might well be called the Axiom of Symbolic Stability. It guarantees us that throughout all our arguments and deductions the symbols remain constant in our memory – or preferably on paper,' and so on. (*Gl*, p. viii)

Clearly, we would not be able to arrive at correct mathematical arguments if our inkblots were constantly to change. Yet that does not imply that mathematics presupposes the physical laws of inkblots, that those laws would figure in the justifications of mathematical laws.

It is important to note that something must give content to the distinction between rational basis and mere concomitant; something must provide a means for saying what counts as showing that one proposition is the rational basis for another, and showing when one proposition presupposes another. It is Frege's logic that plays this role. Logic tells us when one claim is a ground for another, namely, when the latter can be inferred, using logical laws, from the

[28] *Gl*, p. vi.

former. Explanation and justification are matters of giving grounds. For Frege, then, the explanation of a truth is a logical proof of that truth from more basic truths; the justification of a truth is a logical proof of that claim from whatever first principles are its ultimate basis. Thus the laws of logic are explicatory of explanation and justification; on this rests their claim to the honorific title 'logic'.

Given this role for logic, it should occasion no surprise that Frege's conception of logic and the demands he puts on the notion of justification are closely linked. Now the notion of justification plays a philosophically very important role for Frege, as it is key to his argument for the logicist project. Although we might start off thinking that arithmetical discourse is completely understood, transparent, and poses no problem, Frege urges that we lack knowledge of the ultimate justification of the truths of arithmetic. In order to 'afford us insight into the dependence of truths upon one another', we must analyse the seemingly simple concept of number and find the 'primitive truths to which we reduce everything' (*Gl*, p. 2). Frege also brings up 'philosophical motives' for the logicist project, asking what looks to be the traditional philosophical question of whether arithmetic is analytic or synthetic. But actually he redefines these notions (as well as those of a priori and a posteriori) so that they concern 'not the content of the judgment but the justification for making the judgement' (*Gl*, p. 3). Here too it is the notion of justification that is doing the work.

Essential to the role of this notion of justification in supporting the logicist project, and to the plausibility of Frege's redefinitions of traditional philosophical terminology, is the applicability to all knowledge of the standards of justification. The canons of justification must be universal in their purview: 'Thought is in essentials the same everywhere: it is not true that there are different kinds of laws of thought to suit the different kinds of objects thought about' (*Gl*, p. iii). Another important feature of justification is explicitness: a justification must display everything on which the truth of the claim being justified depends. To insure that 'some other type of premise is not involved at some point without our noticing it', a justification must provide 'a chain of inferences with no link missing, such that no step in it is taken which does not conform to some one of a small number of principles of inference recognized as purely logical' (*Gl*, p. 102).

Obviously, these demands are met when logic, as invoked in Frege's notion of justification, is taken on the universalist conception. That the canons of justification must extend to all areas of knowledge requires utmost generality and universal applicability of the logical principles. Explicitness is vouchsafed by the direct applicability of logic: there are no presuppositions, no implicit steps, in the application of logical laws. To illustrate this, let us examine how, on Frege's picture, logic would be used to justify the conclusion that all whales are vertebrates on the basis of the claims that all whales are mammals and that all mammals are vertebrates. We start with the assertions:

(1) All whales are mammals.

(2) All mammals are vertebrates.

We then provide a logical demonstration from first principles that ends with:

(3) $(\forall F)(\forall G)(\forall H)[(\forall x)(Fx \rightarrow Gx) \rightarrow ((\forall x)(Gx \rightarrow Hx) \rightarrow (\forall x)(Fx \rightarrow Hx))]$.[29]

Three instantiation steps then license us in the assertion of:

$(\forall x)(x$ is a whale $\rightarrow x$ is a mammal$) \rightarrow ((\forall x)(x$ is a mammal $\rightarrow x$ is a vertebrate$) \rightarrow (\forall x)(x$ is a whale $\rightarrow x$ is a vertebrate$))$.

Or, in ordinary English:

(4) If all whales are mammals, then if all mammals are vertebrates then all whales are vertebrates.

By *modus ponens* from (4) and (1) we obtain:

(5) If all mammals are vertebrates then all whales are vertebrates.

Finally, by *modus ponens* from (5) and (2), we arrive at:

(6) All whales are vertebrates.

Taken together, all these assertions, including those in the logical proof of (3), constitute the justification of the assertion of 'All

[29] See *Bs*, §23.

whales are vertebrates' on the basis of the assertions of 'All whales are mammals' and 'All mammals are vertebrates'.

The requirement of explicitness and the need for the logical laws to be directly applicable can be highlighted by consideration of an argument against logicism devised by Henri Poincaré.[30] The version I summarize here is formulated by Charles Parsons.[31] In order to show that arithmetic is logic, one must devise a formal system of logic and show how the theorems of arithmetic can be obtained in that formal system. Now, to give a formal system is to specify, first, the class of formulas and, second, the class of derivable formulas. The usual form of specification is this: certain basic expressions are stipulated to be formulas; other formulas are specified as those and only those expressions obtained from the basic expressions by finitely many applications of certain syntactic operations. Similarly, certain formulas are stipulated to be axioms; the derivable formulas are specified as those and only those formulas obtained from the axioms by finitely many applications of certain inference rules. Thus these specifications are inductive in nature: the notion of a finite number of applications of given operations is essential to them. Therefore, number is presupposed in the logicist foundation for arithmetic. This is a *petitio principii*. Thus there is a logical circle in the logicist reduction.

I believe Poincaré's objection fails, and it is important to see why. The objection would succeed if Frege construed the justification of arithmetic to involve, for one or another arithmetical claim, the following assertion: 'This claim is provable in such-and-such formal system.' That assertion is a metatheoretic one. It is about the formal system; since Poincaré is quite right that inductive definitions are used to specify the formal system, it follows that the assertion relies on number theory. That is not, however, how Frege conceives of justification. To give a justification of an arithmetical claim is to give the claim with its grounds. It is not to assert that the claim is

[30] Henri Poincaré, 'La logique de l'infini', *Revue de Métaphysique et de Morale*, 17 (1909), pp. 461–82. Poincaré was responding to Bertrand Russell's 'Mathematical logic as based on the theory of types', *American Journal of Mathematics*, 30 (1908), pp. 222–62.

[31] Charles Parsons, 'Frege's theory of number', in his *Mathematics in Philosophy: Selected Essays* (Ithaca, N.Y.: Cornell University Press, 1983), pp. 150–75, at p. 168. See also Mark Steiner, *Mathematical Knowledge* (Ithaca, N.Y.: Cornell University Press, 1975), p. 28ff.

provable; it is to give the proof. Now, of course, one might want to verify that what has been given is, in fact, a proof by the lights of the formal system. Such a verification would proceed by syntactic means, and does presuppose the specification of the system. The verification is not *constitutive* of the argument's being a justification; it is just a means for us to ascertain that it is. In order for us to be psychologically sure that what we are giving are justifications, we have to use our knowledge of the formal system, that is, our metatheoretic knowledge which is of an inductive nature. But that is different from what the justification of the claim actually is.

Here Frege is relying precisely on the distinction between what we might have to do, in fact, by our natures, in order to be in a position to do mathematics, and what the justification of mathematics is. That we need to set out a formal system to be sure of our justifications is no more relevant to the rational grounds of mathematics than our need to write down proofs because otherwise we will not remember them.

The Fregean rebuttal to Poincaré requires that in what Frege would call a justification, say of an arithmetical truth, everything that is presupposed by the truth does play a role. This lies in back of his demand for 'gap-free' deductions.

To gain an appreciation of the role of the universalist conception of logic in this, it is instructive to contrast how a justification abiding by the Fregean requirement of explicitness would have to proceed if logic were taken on the schematic conception. Let us once again undertake a justification of 'All whales are vertebrates' on the basis of 'All whales are mammals' and 'All mammals are vertebrates'. We can't simply pass from the latter to the former, with a note ('off to the side', so to speak) that the latter two jointly imply the former, since this does not make explicit what is involved in the inference. Rather, matters have to be laid out as follows. As before, we start by asserting:

(1) All whales are mammals.

(2) All mammals are vertebrates.

We then assert, along with whatever grounds needed to show it from first principles:

(3) There is an interpretation of '$(\forall x)(Fx \to Gx)$,' '$(\forall x)(Gx \to Hx)$,' and '$(\forall x)(Fx \to Hx)$' under which these schemata

become (regimented versions of) the sentences 'All whales are mammals', 'All mammals are vertebrates', and 'All whales are vertebrates', respectively.

We now adduce a mathematical proof culminating in:

(4) Any interpretation that makes '$(\forall x)(Fx \rightarrow Gx)$' and '$(\forall x)(Gx \rightarrow Hx)$' true also makes '$(\forall x)(Fx \rightarrow Hx)$' true.

Using some logical laws and intermediate steps for making the transition, we can assert on the basis of (3) and (4):

(5) If 'All whales are mammals' and 'All mammals are vertebrates' are true, then 'All whales are vertebrates' is true.

To apply (5), we must adduce the Tarski paradigms:

(6) 'All whales are mammals' is true if and only if all whales are mammals.

(7) 'All mammals are vertebrates' is true if and only if all mammals are vertebrates.

(1), (2), (5), (6), (7), and truth-functional laws will allow us to obtain:

(8) 'All whales are vertebrates' is true.

Finally, adducing

(9) 'All whales are vertebrates' is true if and only if all whales are vertebrates,

we obtain:

(10) All whales are vertebrates.

Needless to say, from Frege's point of view this outline already looks terribly circuitous, and the amount that has to be filled in to provide justifications for (3) and (4) will make matters worse. Even ignoring Frege's scruples about a truth predicate, the status of the disquotational biconditionals is also troublesome, for, in what is outlined, those biconditionals figure among the grounds of 'All whales are mammals' as much as do assertions (1) and (2). If, for example, they are meant to be consequences of a substantial semantic theory, then we are in the position of requiring that theory in the justification of 'All whales are vertebrates' on the basis of 'All whales are mammals' and 'All mammals are vertebrates.' Matters look less peculiar if the truth predicate is meant to come merely from a Tarski-style definition; but even here an oddly large body of mathematics must figure in order to justify what is, after all, a rather simple logical inference. All this is to say that the schematic conception of logic fits poorly with the Fregean picture of justification.[32]

This lack of fit comes out in another difficulty as well. In the justification as just outlined, various transitions, like that from (3) and (4) to (5), will be made by applying logical rules. On the schematic conception, logical rules are justified only on the basis of their soundness, that is, their yielding logical consequences. But then it looks like the justification we have presented is not fully explicit; there is something left unsaid that it presupposes.

It might be objected, however, that there is a similar problem in the justification given on the universalist conception. In it, inferences are made in accord with certain inference rules. Shouldn't the demand of explicitness be invoked further, to require that whatever principles lie behind the correctness of the inference rules be made explicit and considered part of the justification? In general, the only way of stating these principles are as the soundness or truth-preservingness of the rules and involve semantic ascent and a truth predicate. Thus the 'directness' alleged for the universalist conception papers over an elision.[33]

Now I believe Frege would reject the idea that inference rules rest on or presuppose the principles expressing their soundness. Rather, our appreciation of the validity of the rules is not the recognition of the truth of any judgement at all; it is manifested in our use of the rule, in our making one assertion on the basis of another in accordance with the inference rule.[34] There is nothing more to be made explicit, although of course individual instances of the inference rule can always be conditionalized and asserted as logical truths.

[32] Not that it was meant to: Tarski's and Quine's views of justification are different from Frege's.

[33] This line of objection is suggested by a remark by Parsons in 'Objects and logic', p. 503.

[34] Here I draw on Thomas Ricketts, 'Frege, the *Tractatus*, and the logocentric predicament', *Noûs*, 19 (1985), pp. 3–14, at p. 7. See also Ricketts, 'Logic and truth', §2.

To some this may appear to be an evasion. But let us investigate the question we left hanging with respect to the schematic conception. There, the justification looked inexplicit because it omitted a demonstration of the soundness of the logical rules it employed, and, on the schematic conception, logical rules are justified only on the basis of their soundness. Of course, one could adjoin a demonstration of soundness. Naturally, that demonstration will use logical rules. Usually the soundness of those rules will not be vouchsafed by the adjoined demonstration, because the quantified variables in the demonstration will have to range over a larger class than any of the universes of discourses of the interpretations covered by the soundness proof. For example, an everyday soundness proof shows that the usual logical rules are sound with respect to all interpretations whose universes of discourse are sets. The reasoning in that proof involves variables ranging over all sets; hence, the universe of discourse of that reasoning is a proper class. A soundness proof for the logical rules used in the everyday proof would therefore have to show something stronger than everyday soundness, namely, that the rules were sound with respect to interpretations whose universes of discourse were proper classes. This would require a stronger set-theoretic language yet, in which collections of proper classes existed, and the reasoning in the stronger soundness proof would involve variables ranging over such collections. This process continues with no end. To avoid a vicious regress, we have to be able to take the logical rules used in the justification for granted. Yet, on this conception, it has to be admitted that a fuller justification, one amplified by a further soundness proof, is always possible. In passing to that fuller justification, we also pass to a larger universe of discourse. The upshot is that at no level can one think of the quantifier as ranging over everything; there is no absolutely unrestricted quantifier. All the while, though, in enunciating the claims at any level, one is not (yet) in a position to specify how the quantifiers are restricted: they range over everything that at that point one can have. This is a curious position, one which goes far more against Frege's demand for explicitness than our acceptance of a rule of inference without an explicit semantic principle to back it up.[35]

[35] The position that there can be no such thing as a truly unrestricted quantifier is due to Charles Parsons in 'The Liar Paradox', in his *Mathematics in Philosophy*, pp. 221–50. Parsons's argument is based on phenomena associated with the Liar Paradox. The use of the idea in the current context is due to Ricketts.

This last argument has brought us rather far afield. My central aims in this paper have been to delineate Frege's universalist conception of logic and contrast it with a more familiar one, to show that this conception connects with many other points in Frege's philosophy, and to suggest that the conception is a well-motivated one, given the nature of Frege's project. Of course, today most of us would find the schematic conception (or some variant of it) far more natural, if not unavoidable. But I hope to have caused us to reflect on how much else has to shift in order to make it so.[36]

[36] Reproduced from 'Frege's conception of logic', in J. Floyd and S. Shieh (eds.), *Future Pasts: The Analytic Tradition in Twentieth-Century Philosophy* (Oxford University Press, 2001), pp. 25–41. By permission of Oxford University Press, Inc. I am greatly indebted to Thomas Ricketts for countless conversations and comments, as well as for access to his unpublished works. Needless to say, he does not agree with all the formulations given in this paper.

4 Dummett's Frege

1 INTRODUCTION

It has become standard for commentators to note sadly how little impact Frege's work had amongst his contemporaries, but then to temper this observation by claiming an enormous indirect influence for his ideas through the work of those few who did pay serious attention to them, perhaps most notably Russell, Wittgenstein and Carnap. How effective or transparent those conduits were is still a matter of scholarly debate.[1] For myself, I am increasingly persuaded that much of what we would now judge to be most centrally important in Frege was at best imperfectly transmitted.

That we can now attempt judgement on what is thus central is owed, in the first place, to the republication and translation of Frege's work that effectively began with Austin's version of *Grundlagen* in 1950. Austin had translated the work so as to be able to set it for an Oxford finals paper. Michael Dummett took the course, and was, he reports, 'bowled over by the *Grundlagen*', so much so that during the following year he 'settled down to read everything that Frege had written'[2]. Soon, though not at first, Geach and Black's *Translations*

[1] The route via Carnap is considered by Erich Reck and Steve Awodey (Erich Reck and Steve Awodey (eds.), *Frege's Lectures on Logic: Carnap's Student Notes, 1910–14* (Chicago: Open Court, 2004), Introduction). Recently Michael Potter has added greatly to our understanding of the route through Cambridge, demonstrating how certain of Frege's very general insights, for instance about the unique status of logic and the special character of its defining concern with the notion of truth, shaped Wittgenstein's earliest reflections (M. Potter, *Wittgenstein's Notes on Logic* (Oxford: Oxford University Press, 2009), ch. 29).

[2] M. Dummett 'Intellectual autobiography', in R. E. Auxier and L. E. Hahn (eds.), *The Philosophy of Michael Dummett* (Chicago: Open Court, 2007), pp. 9–10.

(1952) would help in this, but before long the work would take Dummett to Munster to examine Frege's unpublished work: the first result of this study is the 1956 'Postscript' to his 1955 'Frege on functions', itself an important early step in dispelling bizarre misconceptions of Frege's doctrines which seem then to have been prevalent.[3] Dummett began to form plans for a comprehensive book on Frege. This required further sustained study of the *Nachlass*, including a visit in 1957 when, its editors acknowledge, Dummett provided an important stimulus and essential 'spadework' (*PW*, p. xii) towards its publication. *Frege: Philosophy of Language*, a rather different book from that first planned, eventually appeared in 1973. Dummett modestly remarks of it, 'I believe that the book helped to revive interest in Frege.'[4] Peter Geach, with whom he had been in discussion about Frege virtually from the beginning of the work, more nearly conveys its importance:

> As a guide to Frege's thought [Dummett] is absolutely unrivalled. Many books about Frege have now appeared; none even distantly approaches this in its wide learning, deep sympathy with Frege, and clarity and patience of exposition.[5]

Of course, many more books have since appeared, but a similar judgement today would be no less clearly true. The development of the interpretation the book offers played so vital a role in the wider rediscovery of Frege that the two cannot be disentangled. Its insight and force are such that it can fairly be said to be the second factor that now informs our understanding of what Frege was centrally about. (This much is acknowledged – in a back-handed kind of way that has understandably irritated Dummett[6] – by those who contest his interpretation, while nonetheless describing it as the 'orthodox' or 'standard' view.)

To attempt to sketch even the highlights of Dummett's Frege in one chapter would be foolish. The approach I have chosen instead is to confine attention to the opening three chapters of *Frege: Philosophy*

[3] E. D. Klemke's anthology, *Essays* on *Frege* (Urbana: University of Illinois Press, 1968), bears witness to many of these.

[4] Dummett 'Intellectual autobiography', p. 24.

[5] P. T. Geach, 'Critical notice of FPL', *Mind*, 85 (1976), p. 449.

[6] M. Dummett, *The Interpretation of Frege's Philosophy* (London: Duckworth, 1981), p. xiv. Hereafter *IFP*.

of Language, because in these chapters I think we can most clearly see the central reason for the unique authority and depth of his interpretation. This reason is that Dummett's interpretation begins where Frege himself began, with the discovery, in *Begriffsschrift*, of quantification, 'the deepest single technical advance ever made in logic'.[7] As he says, at the opening of chapter 2,

> The discovery by Frege, at the outset of his career, of the notation of quantifiers and variables for the expression of generality dominated his entire subsequent outlook upon logic. By means of it, he resolved, for the first time in the whole history of logic, the problem which had foiled the most penetrating minds that had given their attention to the subject. It is not surprising that Frege's approach was ever afterwards governed by the lessons which he regarded as being taught by this discovery. (*FPL*, p. 8)

This discovery and those lessons will be our concern. We will approach them, as Dummett does in chapter 1, through an account of the philosophical context in which the discovery took place.

2 THE FOUNDATION OF A THEORY OF MEANING

'Precision and rigour,' Frege says, 'are the prime aim of the concept script' (*PW*, p. 32). He tells us that the idea of the concept script first arose in the course of investigations to determine 'how far one could get in arithmetic by means of logical deductions alone, supported only by the laws of thought' (*Bs*, Preface; *CN*, p. 104). The aim immediately requires some means of being assured that no other, unacknowledged support is being tacitly relied on. It would be essential, then, to lay out his deductions *in full*, 'to keep the chain of reasoning free of gaps' (*ibid.*). But what does it *mean* to do that, and how will we know whether we have managed it? On both questions Frege found that the natural language in which he at first tried to formulate these deductions let him down. Aiming for explicitness produced only unwieldiness and complexity, thus increasing, rather than reducing, the resources presupposed of anyone in his receiving, understanding and endorsing the reasoning set out. This situation seemed to put, not just the achievement of the goal, but even any understanding of it as an approachable ideal, further

[7] M. Dummett, *Frege: Philosophy of Language*, 2nd edn (London: Duckworth, 1981), p. xxxiii. Hereafter *FPL*.

from view. Dwelling on it, one could easily be led to think that full explicitness in reasoning, like full understanding of another human being, is a whimsical fantasy.

This is the problem from which Dummett's exposition sets out. 'The original task that Frege set himself', he says, 'was to bring to mathematics the means to achieve absolute rigour in the process of proof … What Frege wanted was a framework within which all mathematical proofs might be presented and which would offer a guarantee against incorrect argumentation: of a proof so set out, it would be possible to be certain that it was not erroneous, or valid only within certain restrictions not made explicit, or dependent on unstated assumptions' (*FPL*, p. 1). In the three paragraphs that begin with this observation Dummett sets out a train of thought that gradually expands upon what the realization of Frege's goal requires. Each step is of the form: to achieve Y it was necessary to provide X. The striking conclusion soon reached is this: that 'Frege had …to provide the foundation of a theory of meaning' (*FPL*, p. 2).

This remarkable train of thought is what justifies Dummett in choosing 'Frege: Philosophy of Language' as his title. It starts out from what looks at first like a narrow methodological requirement of a specific and characteristically nineteenth-century programme in scientific foundations, and develops from it a fundamental reconception of the philosophical enterprise of understanding thought and its relation to reality that would dominate the twentieth century. It takes the seeming whimsy of fully explicit understanding and creates out of it a crystalline formal model towards which that enterprise has been directed. So to appreciate this train of thought would in large measure be to understand, not just how analytic philosophy, a movement with its roots in a particular foundational project in mathematics, should have accorded such centrality within philosophy generally to a concern with language and the theory of meaning, but the particular formal and aprioristic character of that concern.

The first steps of this train of thought specify the framework Frege wanted as necessarily including a formalized language supplemented by formal rules of proofs. As Frege typically presented them, these points are clearer in reverse order.

The aim of gapless proof is to 'expose each presupposition which tends to creep in unnoticed, so that its source can be investigated'

(*Bs*, Preface; *CN*, p. 104). Such presuppositions may be unstated premises, but equally they may be implicit in modes of reasoning special to the subject matter. (Frege illustrates this (*CN*, p. 85) by showing how assumptions of the asymmetry and transitivity of 'larger than' have entered into geometrical reasoning.) His initial remedy is to require that every special assumption be declared in the form of a premise, restricting allowable inferences to those sanctioned by entirely general laws of thought. The first and most obvious requirement of this is that the allowed inference modes must be strictly circumscribed. The second is that there must be no opening left for 'something intuitive … to squeeze in' (*CN*, p. 104) in determining whether a given inference exemplifies one of the specified modes. The allowable modes of inference must therefore be formally specified, so that correct inference can be confirmed 'on the basis of linguistic form' (*CN*, p. 85). It is just here, though, that the inadequacy of natural language shows itself. Ordinary linguistic practice of course offers no determinate set of allowable inference modes. But the more fundamental point is that its looseness and vagueness cannot be rectified by specifying such a set. 'Language is not governed by logical laws in such a way that mere adherence to grammar would guarantee the formal correctness of thought processes' (*ibid.*). No formally specified rules could therefore be sound in direct application to inference conducted in natural language. At best we could devise rules to which ordinary-language inferences would be compared and held responsible: these laws would be 'applied externally, like a plumb-line' (*ibid.*). Clearly, though, that process of comparison would be another entry point for 'something intuitive' to creep in.

The conclusion thus far, then, is that gapless proof must be formal proof, which must be proof conducted in a formalized language. Approaching the conclusion in Frege's way has made prominent points about what a formalized language must be that Dummett's third paragraph elaborates. But it is worth dwelling briefly on his summary of these initial steps. 'Frege was,' he says, 'proposing [*A*] to take the step from the axiomatization of mathematical theories … to their actual formalization'. The axiomatic method had been employed to isolate and elucidate the basic notions of various mathematical theories. But 'what Frege wanted was [*B*] to subject the process of proof to an equally exact analysis' (*FPL*, pp. 1–2). How is it that *A*, the step to formalization, amounts to *B*, the analysis of proof itself?

A straightforwardly mathematical answer presents itself. The formalization of proof constitutes it as a precisely delineated phenomenon susceptible to mathematical investigation, just as the axiomatization of spatial properties and relations in geometry had done for them. Dummett is clear, though, that, while this was part of Frege's achievement, it was not his aim (*FPL*, p. xxxv): making formal systems into 'the objects of mathematical investigation' is *one* way in which 'light is thrown on the nature of mathematical proof' (*FPL*, p. xxxiv), but it is not the way Dummett has in mind here. His idea depends instead, I think, on a rather different sense in which, in the formalized proofs Frege envisaged, correct movements of thought will be fully laid out, and everything involved in them rendered fully explicit. According to this idea everything that makes the proof into an instance of correct reasoning, and so what it is in this instance for the reasoning to be correct, will be there on the page and open to reflective review. Of course, recognizing it as such will depend on one's understanding of it, but it will not rely on any *further* understanding, not expressed in it, in the light of which its correctness is to be appreciated. In that way the role of the receiving mind shrinks, as it were, to a pure receptivity, since the configuration it is to adopt is the configuration exemplified in the proof itself. The thought, as Wittgenstein would later put it, *is* the proposition.[8] This idea has, I think, been enormously influential. It is at work, for instance, in Davidson's striking methodological pronouncement that he 'will not think of [languages] as separable from souls'.[9] More relevantly here, it informs Dummett's own contention that the analysis of thought consists in – is to be achieved through, and only through – the analysis of language. He is right to find the roots of this idea in Frege's formalizing project.

Dummett now brings these points about formalization into connection with truisms about inference: that the validity of a proof depends on the meanings of the statements figuring as premises and conclusion; and that the meaning of a statement is fixed by the

[8] L. Wittgenstein, *Tractatus Logico-Philosophicus* (London: Kegan Paul, 1922), p. 4.

[9] D. Davidson, 'On the very idea of a conceptual scheme', reprinted in his *Inquiries into Truth and Interpretation* (Oxford: Clarendon Press, 1984), p. 185.

meanings of the words occurring in it and their manner of combination (*FPL*, p. 2). If, in a formalized proof, validity is to be determinable by grammatical form alone, then the grammatical forms involved must offer a consistent and complete reflection of the 'logical relations' (*CN*, p. 85) between concepts on which the meanings of its constituent statements, and thus the validity of the proof, depend. As Leibniz had foreseen, the idea of a *calculus ratiocinator*, a method of proof in which validity is effectively decideable by reference to syntactic criteria,[10] 'has the closest possible links' with that of a *lingua characteristica*, 'a script which compound[s] a concept out of its constituents' (*PW*, p. 9): the first is feasible only as a supplement to the second. 'An analysis of proof', then, must rest on 'an analysis ...of the structure of the statements that make up the proof' (*FPL*, p. 2), since that analysis must be embodied in the grammatical structures of the sentences of the language in which the proof is conducted. What this calls for is a semantic analysis, one that explains 'how the meaning of each sentence [is] determined from its internal structure'. This is the requirement that Dummett summarizes by saying that 'Frege had, in other words, to provide the foundation of a theory of meaning' (*FPL*, p. 2).

Why is this an appropriate summary? One might be misled on this by a formulation earlier in the paragraph, which represents it as Frege's task 'to give an analysis of the structure of the statements of our language' (*FPL*, p. 2). The semantic analysis the project demands relates rather to the sentences of the formalized language, those actually occurring in proofs. There are familiar views according to which such an analysis sheds a more or less direct light on the means by which the same propositions are expressed by our own, natural-language sentences; such views would then cast Frege's analysis as at least an important first step in explaining how our sentences bear the meanings they do – perhaps, then, as a 'foundation of a theory of meaning' for them. But no such view is entailed in the train of thought we have reviewed, and none is plausibly attributed to Frege (*CN*, p. 85; *PW*, p. 13). Dummett, moreover, would clearly agree with these points. He repeatedly emphasizes that Frege did

[10] 'Calculemus' suggests a stronger notion of decidability, where the production, and not just the checking, of proofs could be mechanized; this was not part of what Frege took from Leibniz.

not aim to explain 'how our language works' (*FPL*, p. 36) – the complex and disorderly means by which it manages, to the extent that it does, to express the propositions that interested him – and that his key achievement, the analysis of propositions involving multiple generality, was reached 'by ignoring natural language' (*FPL*, p. 20). Of course, if natural language has the means to express these propositions, then Frege's analysis of their structure will stand as a constraint on any account of what those means might be; but 'if not', Dummett says, nicely capturing Frege's own attitude, 'so much the worse for natural language' (*ibid.*). Frege need take no stance on the issue, and neither is any stance taken in the conclusion of Dummett's that we are considering.

This conclusion should instead be understood in the light of Dummett's later explanation that a theory of meaning is 'a general account of the workings of language' (*FPL*, p. 83), or 'an account of how language functions, that is, not only of how it does what it does, but of what it is that it does' (*FPL*, p. 92). Formalization demanded of Frege's language complete perspicuousness in how it does what it does; the language thus furnishes us with a transparent model of what it is that it does. The philosophical questions that we look to a theory of meaning to answer are focused on language as such, not on this language or that, but the demands of explicitness placed on a formalized language are such that it cannot but present us with answers to some of the most basic of those questions. In this sense Frege had indeed to provide 'the foundation of a theory of meaning'.

As yet this is, as Dummett might say, entirely programmatic. Frege's insights into the demands and yields of formalization are of fundamental importance for understanding the course taken by the philosophy of language and thought in the twentieth century. But it seems scarcely possible either that he should have arrived at these general insights, or that they should have had such an influence, except in connection with the specific analyses proposed in *Begriffsschrift*, to which we now turn.

3 QUANTIFICATION

No texts exist by which we might try to reconstruct Frege's discovery of quantificational logic. What we can do, and what Dummett brilliantly does in chapter 2 of *FPL*, is to identify the insights underlying the discovery and to make vivid the force they would have for someone attaining to them for the first time (*FPL*, p. 9).

The first of these insights is that the idea of step-by-step construction, the conception of a complex expression as put together from its basic vocabulary in a series of steps, should be invoked in understanding how expressions of generality function in sentences (*FPL*, p. 10). The step-by-step conception was of course not new with Frege: children would have been drilled in it then as they are now, learning how differently to evaluate $(2+3)\times 4$ and $2+(3\times 4)$. Nor was its application in logic: using the signs to mean logical addition and multiplication, $(p+q)\times r$ and $p+(q\times r)$ are differently evaluated in a precisely parallel way. Nor again was its application to the logic of generality: with a different understanding of the letters, as standing now for classes, parallels between propositional and syllogistic logic had been mapped. But here we run into the limitations of this approach. In the first place, the model borrowed from arithmetic supplies us with a treatment either of propositional logic or of syllogistic logic, but, because different understandings of the operators are involved, no combination of the two (*PW*, pp. 14–18). In the second place, its treatment of the logic of generality is restricted to inferences turning on the occurrence of only one expression of generality (*PW*, pp. 18–20). Frege's insight was that these are effectively the same limitation, to be overcome at once by extending propositional logic to include operations of generalization, operations that could be iterated as (it was already understood) can the propositional operators.

The apparent role of an expression of generality, such as 'everyone' or 'someone', is to combine with a predicate to specify what things the predicate is true of, as the role of a name is to combine with a predicate to specify one particular thing the predicate is true of. But the iterable operations now sought would necessarily lead from sentences to sentences. So the implementation of Frege's first insight called for 'a second fundamental idea' (*FPL*, p. 15), which would reconcile these seemingly incompatible roles. This was to conceive the required operations as involving two steps. The first, which we can call 'predicate abstraction', takes a sentence as input and forms an incomplete expression, a predicate (or 'function' in Frege's terminology), by omitting, or conceiving as replaceable (at one or more of its occurrences), some component of the sentence (*Bs*, §§9–10). The second consists in the application of a quantifier, e.g. 'everything', to a predicate thus abstracted, forming again a sentence (*Bs*, §11). This sentence will be true just in case the predicate to which the quantifier is attached is true of everything; and this predicate will be true of a given thing just in case a sentence from which it might have been formed, by omitting a designation of that thing, is true. Because these two-step operations are iterable they can account uniformly for the significance of sentences containing any number of expressions of generality. And because they lead from sentences to sentences they can interleave, in the step-by-step construction, with the operations of the propositional logic they extend. These two features together explain the immense power of Frege's conception.

4 UNIQUE STRUCTURE

In everything said so far one idea has been dominant: namely, that the understanding yielded through Frege's logical insights is an understanding of structure. The real benefit of formalizing inference emerges only through the requirement it imposes on the language in which inference is conducted, to render its content more exactly (*PW*, p. 12), to 'spell out [logical relations] in full' (*PW*, p. 13; *CN*, p. 85), to construct a content out of its constituents (*PW*, p. 35). The primary insight of Frege's account of quantification yields such benefits precisely because it incorporates the expression of generality within the only feasible model for that construction. So, unless all of the above has been very wide of the mark, it is an utterly central commitment of Frege's that propositions have a unique structure which is reflected in their expressions and in virtue of which they stand in inferential relations one to another. Yet what Dummett identifies as the second fundamental idea of Frege's account is that of predicate abstraction, and some of what Frege said in elaboration of this idea has led others to attribute to him a contrary commitment.

To introduce the issue we can contrast two rather different notions often in play in philosophical discussion of 'the structure of propositions' (or 'the structure of thoughts'). The first invokes the idea of each proposition as a node in a space of propositions. The geometry of the space is given by internal relations holding amongst propositions, relations which determine any given proposition's location in the space (and so, presumably, sufficient to identify it). According to the second notion the structure of a proposition is instead a matter of its inner complexity, the way in which it is constructed out of its ingredient elements. This potential ambiguity is often tolerated, it seems, because of a presumption that the two notions will somehow run along in step. But even if that is so we can still ask which leads and which follows. Does the inner complexity of a proposition ground and explain its internal relations to others, or is it rather that the structure ascribed to any particular proposition is merely a reflection of these relations? If the first of these views is correct, we will expect there to be, for each proposition, a single determinate account to be given of how it is constructed from its constituent elements. The second view allows, on the other hand, that there will be different, equally legitimate ways of representing a proposition as structured, each highlighting some amongst the internal relations it bears to others.

The stance adopted here is that Frege adheres firmly to the first of these views. Dummett's discussion in chapter 2 of *FPL* demonstrates the correctness of this stance, and goes on to show how, within it, we should explain the most important of those features of Frege's account that have led others to ascribe to him instead the second view. The immediately following section will consider these matters. Dummett goes beyond this, however, in recognizing and exploring a different challenge to the first view. This further challenge takes off from the observation that, where something *X* is held to stand in any such relation as grounding, entailing or explaining to something *Y*, we should not in general suppose that *X* will fully reveal itself in *Y*, or that *X* will be adequately characterizable by attention to *Y*. A thing's physical make-up explains why it will tip the balance against some other things while being outweighed by still others, but there is only so much you can learn about the physical make-up of a thing with a set of scales. If the structure of propositions grounds their inferential connections, might it not similarly be the case that logic, which attends to and encodes these connections, can give us only a partial view of the grounding structures? Wittgenstein for one, and Ramsey for another, certainly thought so. Their views represent a deep challenge to Frege's conception of the relations between logic, analysis and ontology. The challenge lies well beyond the scope of this chapter. But in the final section I will consider Dummett's treatment of an issue that is, in my view, essential for evaluating it.

5 ANALYSIS AND DECOMPOSITION

Traditionally the philosophical analysis of a proposition or thought, like the grammatical analysis of a sentence, begins with its division into subject and predicate. Frege, however, allows this distinction no place in his way of representing a judgement (*Bs*, §3), and proposes to replace the concepts of *subject* and *predicate* with those of *argument* and *function* (*Bs*, Preface; *CN*, p. 107). Using these simply as replacement terms we might then describe 'carbon-dioxide is heavier than oxygen' (*cHo*) as dividing into the argument 'carbon-dioxide' and the function '... is heavier than oxygen' (i.e. as dividing *c* | *Ho*). 'This distinction,' Frege however remarks, 'has nothing to do with the conceptual content, but only with our way of viewing it' (*Bs*, §9). To reinforce the point he introduces for comparison the proposition 'carbon-dioxide is heavier than hydrogen' (*cHh*). Suppose we think of this second as obtained from the first by replacing 'oxygen' by 'hydrogen', leaving the remainder constant; then we see the pair as dividing

(1) *cH* | *o*
cH | *h*

and so have the same function with different arguments. But now suppose instead that carbon-dioxide represents the focus of our interest; then we will view the second, as initially we viewed the first, as telling us something about it, so that the pair now divide

(2) *c* | *Ho*
c | *Hh*

giving us different functions with the same argument. Each of these divisions of the pair groups them along with further propositions. The catalogue begun in (1) continues, for instance, with 'carbon-dioxide is heavier than | nitrogen'; that begun in (2) with 'carbon-dioxide|is inert'. Of the interconnections between propositions highlighted in this way different ones will be relevant in different connections – most importantly, in connection with different inferences. But neither of these divisions could claim to take us closer than the other to how any of the propositions divides 'of itself'.

What Frege says here is plainly right. Equally plainly, it has no tendency to tell against the view that there is a way the proposition is articulated 'in itself'. To state the obvious, *cHo* can be divided after its first constituent (*c*|*Ho*) or before its third (*cH*|*o*) because it has three constituents. Neither of these divisions provides an account of how the proposition is constructed out of these three constituents; instead, each of them *depends on* such an account.

This case presents us with probably the simplest example of a contrast which Dummett introduces in *FPL* as one between two different sorts of analysis, or two different understandings of 'analysis' (*FPL*, p. 28), but which he later elaborates in the distinction between *analysis* and *decomposition*. Analysis aims at an account of how a proposition is constructed, typically in several stages, from its simple constituents. Decomposition is then the process by which a proposition so constructed may be regarded as dividing into constant and variable components.[11] The importance of this distinction emerges more clearly in connection with more complex examples. But before we turn to such examples it is worth countering a misunderstanding that finds a challenge to the idea that propositions have a unique structure even in what Frege has to say about this simplest of cases.

In a natural language sentence the principal focus of interest will typically be presented by the grammatical subject of the sentence. So one way of encouraging us to view 'carbon-dioxide is heavier than oxygen' as dividing *cH*|*o* would be to recast it so that the argument 'oxygen' appears as subject: 'oxygen is lighter than carbon-dioxide'. This recasting leaves unaltered the propositional content expressed (*Bs*, §3); and this content has, we said, three basic constituents. But now, it might be asked, which three constituents are these? 'Heavier' and 'lighter' are not synonyms, but indicate relations that

[11] *IFP*, ch. 15; the terminology of 'analysis' and 'decomposition' is introduced at p. 271. We will also follow Dummett's terminological distinction between the immediate or ultimate 'constituents' of a proposition exposed by its analysis, and the 'components' into which it divides on any one decomposition.

are converse to each other. Which of them occurs in the supposed single content expressed by our two sentences?[12]

If this were a good question, we should have to admit that there is no good answer to it; and that would be one route to the conclusion that a propositional content has 'in itself' no unique analysis into its constituents. But it is plainly not a good question. For Frege, a relation is 'incomplete', which means (at least) that to mention a relation is to speak of one thing's standing in the relation to another. So the only sensible question to ask in this case is whether our proposition speaks of something's being heavier than another, or of something's being lighter than another. And the only sensible answer to that question is that the proposition does both, simultaneously. To introduce into Frege's language a symbol allowing us to say that something x is heavier than something y is automatically to introduce a symbol allowing us to say of y that x is heavier than it, i.e. that y is lighter than x. So we do not have to choose whether our single content involves *heavier* or *lighter* as constituent: for it to include the one *is* for it to include the other.[13]

[12] Both J. Levine ('Analysis and decomposition in Frege and Russell', *Philosophical Quarterly*, 52 (2002), p. 205) and T. Ricketts ('Pictures, logic, and the limits of sense in Wittgenstein's *Tractatus*', in H. Sluga and D. Stern (eds.), *The Cambridge Companion to Wittgenstein* (Cambridge: Cambridge University Press, 1996), p. 67) discuss Russell's response to this question, which was to accept that, since *heavier* and *lighter* are obviously different, '*A* is heavier than *B*' and '*B* is lighter than *A*' must express different (though equivalent) propositions, the first having *heavier* as constituent, the second *lighter*. Surprisingly both go on to suggest, quite wrongly, that Frege would have difficulties answering the same question (Levine, pp. 205–6; Ricketts, fn. 31). This is to overlook the essential role played in Russell's response by his commitment that 'these two words ['heavier' and 'lighter'] have certainly each a meaning, *even when no terms are mentioned as related by them*' (B. Russell, *The Principles of Mathematics* (London: Allen and Unwin, 1903), §219, my emphasis), a commitment Frege does not share.

[13] This is *not* to say, as Russell contemplated saying (B. Russell, 'On the notion of order', in his *Collected Papers*, vol. III, ed. G. H. Moore (London: Routledge, 1993), p. 300), and later did say (*Theory of Knowledge*, in his *Collected Papers*, vol. VII, ed. E. R. Eames (London: Routledge, 1983), pp. 87–8), that a relation and its converse are 'the same'. On Frege's account 'sameness' of relations amounts to their relating the same things, so that a relation R will be 'the same' as its converse just in case $Rxy \equiv Ryx$, that is, just in case it is symmetrical, which *heavier than* is clearly not. What the above claims to be 'the same' are rather what it is for a proposition *to express* a relation and what it is for it to express the converse of that relation.

Our counter to that challenge highlights a further principle governing Frege's analysis which Dummett makes central to his exposition (*FPL*, pp. 35ff.), namely, that to identify any element of a proposition, or of the sentence that expresses it, is to identify the role that element plays in determining what is required for the truth of the proposition. Merely pointing to the supposed constituent *heavier*, or *lighter*, falls short of that, since it does not settle which thing is to be the heavier, and which the lighter, if the proposition is true. This point applies to the constituents of a proposition revealed by its analysis, because what analysis aims at is precisely an account of how it is settled, by the constituents of the proposition and the manner in which they are combined, how things must be for the proposition to be true. And since the various possible ways in which the proposition may then be subsequently decomposed are fixed by its analysis the point holds equally of its components under such a decomposition.

Frege indeed makes this point immediately on introducing the notion of decomposition in *Bs*, §9, where he warns that the possibility of dividing the natural language sentences

> The number 20|can be represented as the sum of four squares

and

> Every even number|can be represented as the sum of four squares

in the way indicated does show that we have here the same function of different arguments. To have two values of the same function requires more than the recurrence of the same phrase, and more than, as we surely have here, the recurrence of the same phrase carrying, by ordinary standards, the same meaning. What is required is that the phrase play the same role, in relation to its putative argument expression, in settling what is required for the truth of the whole. In the first of our sentences the role of the phrase is to formulate a proposition that will be true if the condition it expresses is satisfied by the object designated by its argument expression. Since the suggested argument expression in the second sentence, 'every positive integer', does not designate any particular object, that cannot be the role of the common phrase in the second sentence: as Frege summarizes the point, 'what is asserted of the number 20 cannot be asserted in the same sense of [the concept] "every positive integer"'; the concepts are, he says, of different 'rank', or level (*Bs*, §9).

This is the crucial point for deflecting a serious misunderstanding of Frege's remark in the following section that, because Φ appears at a place in the symbol $\Phi(A)$ (which Frege offers as schematic for '*A* has the property Φ'), 'we can consider $\Phi(A)$ as a function of the argument Φ' (*Bs*, §10). Consider alternative pairings of the propositions,

Sk (Kevin snores)
Sh (Herbert snores)
Yh (Herbert yawns)
/ \
(1) *Sk* *Sh* (2)
Sh *Yh*

In the first pair we find the same function (*S* ...) of different arguments (*k* and *h*); in the second the same function (... *h*) of different arguments (*S* and *Y*). Frege holds that 'the different ways in which the same conceptual content can be considered as a function of this or that argument have no importance', or that they have 'nothing to do with the conceptual content' (*Bs*, §9). The misreading concludes that whether *S*, or anything else, is a function has to do only with our way of viewing it – whether we view it in line with pairing (1) or pairing (2) – and so purports to discover a far more radical rejection of the concepts of subject and predicate. But now if the *S* ... that occurs as function in pairing (1) is to be recognized again in pairing (2), its role there must be the same. So conceiving *S* ... as argument to a certain function ... *h*, as pairing (2) encourages, is not an *alternative* to conceiving it as the function we recognize in pairing (1), but must be another instance of that. Thus the view we have of *Sh* under pairing (2), in which *S* ... appears as argument to a certain function ... *h*, is likewise not merely an alternative to the view we took of it under (1), where *S* ... takes *h* as argument; instead it *depends on* that prior conception of it. It follows that the function ... *h* we recognize under pairing (2), which Frege represents '$\varphi(h)$' (*Gg*, §22), cannot be identified with *h*. Characterization of the role played under (2) by $\varphi(h)$ depends on a characterization of its argument *S* ... under (1), represented by Frege as '$S(\xi)$', which in turn depends on that of *h*. In Dummett's terminology, the decomposition of *Sh* into the components $\varphi(h)$ and $S(\xi)$ rests on its prior analysis as constituted from the function $S(\xi)$ and *h*.

This application of the analysis-decomposition contrast illustrates the basic ground of Frege's stratification of expressions and the entities they refer to into hierarchies of names, first-level predicates, second-level predicates, and so on, and of objects, first-level concepts, second-level concepts, and so on. Two principles drive this. The first is that the identification of a propositional element includes the identification of its role in fixing the truth-condition of the proposition. (This first principle is virtually a restatement of Frege's 'context principle', that an expression has meaning only in the context of a proposition.) The second, which is distinctively responsible for the hierarchy, is that some such roles can be characterized only by reference to others, which are therefore presupposed as already determinate. But now what is the root reason for attributing to Frege this whole interlocking conception, in preference to the view on which the structures of propositions are merely a reflection of their interconnectedness?

As befits such a basic feature of his interpretation, Dummett locates this reason in Frege's account of quantification. As we saw, this account explains the move from an instance '$A(c)$' to its generalization '$\forall x\, A(x)$' as proceeding through the formation, by decomposition, of a predicate '$A(\xi)$', to which the quantifier is then attached. The account thus assumes that an understanding of the instance is sufficient to ensure understanding of the subsequent steps, and we therefore need to ask what guarantees this assumption. Dummett distinguishes three versions of it. He formulates an intermediate version as follows.

> [Frege] is making the assumption that, whenever we understand the truth-conditions for any sentence containing (one or more occurrences of) a proper name, we likewise understand what it is for any arbitrary object to satisfy the predicate which results from removing (those occurrences of) that proper name from the sentence ... (*FPL*, p. 17)

An immodest version of the assumption would be got by replacing 'any arbitrary object' by 'all objects'. This immodest version incorporates what in Dummett's view is the additional assumption that grasp of the instance is sufficient to ensure grasp of the range of generalization as a determinate totality (*FPL*, p. 19).[14] This additional assumption – or more accurately, whether it is an additional assumption – need not concern us; I have identified it only because doing so helps to isolate Dummett's treatment of what for current purposes is the core issue. A variant on the other side gives the modest assumption, 'that, from our understanding of "*A*(*c*)", … we can derive the truth-conditions of another sentence "*A*(*d*)", given the sense of "*d*"' (*FPL*, p. 19). We can begin with it.

On the view adopted here, according to which propositions have an intrinsic structure reflected in their expressions, this modest assumption is guaranteed 'just because we understand the sentence by understanding the senses of its constituent expressions' (*FPL*, p. 19). We grasp the proposition '*A*(*c*)' on the basis of our understanding of the name '*c*' and of whatever other simple expressions occur in the context '*A*(ξ)', together with the operations of combination that would be detailed in the analysis of '*A*(*c*)'. Those other simple expressions and operations will be replicated in the analysis of '*A*(*d*)' which will run precisely in parallel save for the replacement of '*c*' by '*d*', which we are also assumed to understand. Our basis for understanding '*A*(*c*)' therefore provides intrinsically for the understanding of '*A*(*d*)', thus guaranteeing Frege's assumption in its modest version. To advance from there we need only add that the understanding of '*c*' contributes to determining the truth-condition of '*A*(*c*)' by fixing an object as its referent, the sentence being understood to express a condition on that object. The basis of our understanding of '*A*(*c*)', that is, intrinsically provides for understanding the abstracted predicate '*A*(ξ)' as expressing a condition that any object will satisfy or not, thus guaranteeing Frege's assumption in its intermediate version.

No such guarantee is forthcoming on the opposed view. On that conception our representing two propositions through the expressions '*A*(*c*)' and '*A*(*d*)' reflects some connection between them, but the conception offers no account of how grasp of one of the propositions puts one in a position to understand the other, or of how, if one happens to understand both, one is thereby placed to discern this particular connection between them. These negative contentions do

[14] Cf. M. Dummett, *Frege: Philosophy of Mathematics* (London: Duckworth, 1991), ch. 24. Hereafter *FPM*.

not depend on assuming an extreme version of the opposed view, as committed to the dubious notion that propositions are grasped initially as wholes, to which a structure is only subsequently attributed in consequence of some connection recognized as holding amongst them. The complaints will still hold against a modest version of the opposed view which ties the grasp of a proposition to linguistic understanding, and so holds that any way of grasping a proposition will discern some structure in it. Frege's assumptions depend on the stronger contention, that for any given proposition there is some structure that must be apprehended by anyone who grasps that proposition, and this must be denied by any variant of the opposed view.

Frege's commitment to the view that each proposition has a unique analysis is an intrinsic part of his account of quantification. When put into the balance against that, such remarks as there are elsewhere in his writings suggestive of the opposed view[15] can weigh very little.

6 SIMPLE AND COMPLEX PREDICATES

Frege's parallel hierarchies, of expressions and the entities they mean, were mentioned briefly in the previous section. They form the topic of chapter 3 of *FPL*, where the principles governing them are clearly set out (*FPL*, p. 45). Because the two hierarchies run in parallel it is customary to speak of 'the' hierarchy, indicating that our interest is in their common structure. I will do that here and, in keeping with the same interest, I will use terminology (and quotation marks) in ways that reflect indifference between expressions and their meanings.

In that vein, then, we can say that Frege's hierarchy is grounded on the two categories of 'complete' expressions, propositions and names.[16] Every other category is conceived as a type of functions,

[15] There certainly are such remarks. Many of the relevant texts are gathered and thoughtfully discussed in Levine, 'Analysis and decomposition'; D. Bell, 'Thoughts', *Notre Dame Journal of Formal Logic*, 28 (1987), pp. 36–50; and H. Hodes, 'The composition of Fregean thoughts', *Philosophical Studies*, 41 (1982), pp. 161–78.

[16] In Frege's later work propositions are merely a kind of names – names of truth-values – so that there is only one basic category; this was, as Dummett says, a 'retrograde step' (*FPL*, p. 7), and we will ignore it.

yielding things of one or the other of these basic categories as values, and taking as argument(s) either things of one of these categories or functions of a type already specified by reference to them.

Of many equivalent notations for these types the one adopted here uses 'P' and 'N' for the basic categories of propositions and names, and specifies a function from *n* things of types $\alpha_1,\ldots, \alpha_n$ to something of type β as being of type $n\beta(\alpha_1,\ldots, \alpha_n)$. Thus in 'Herbert snores unless Keith is a liar' the connective 'unless' takes two propositions to form another, and so is of type 2P(P,P). Its first constituent proposition, 'Herbert snores', contains the name 'Herbert'; its predicate, 'ξ snores', is thus of type 1P(N). In 'Everyone snores' this predicate occurs as the sole argument to a quantifier, 'Everyone φ', which is thus of type 1P(1P(N)). In 'Only Herbert snores' we find the complex predicate 'Only ξ snores', again of type 1P(N), making 'Only ξ φ' of type 2P(N, 1P(N)).[17]

These types fall into levels: N and P are of level 0; and the level of any function is one higher than that of its highest-level argument.[18]

When it comes to pronouncing these type-symbols two readings suggest themselves. We might describe an $n\beta(\alpha_1,\ldots,\alpha_n)$ as something which, in conjunction with *n* items of types $\alpha_1,\ldots,\alpha_n$, will form a β; call this the 'construction' reading. Alternatively, we might describe it as what remains of a β when items of types $\alpha_1,\ldots,\alpha_n$ are removed; call this the 'remainder' reading. Neither reading happily covers all cases.

Consider, for instance, 'Only Herbert snores' with respect to the occurrence in it of the predicate 'ξ snores'. This predicate is of type 1PN, and the containing context 'Only Herbert φ' is thus of type 1P(1PN). But it would be at best misleading, adopting the 'construction' reading, to say of 'Only Herbert φ', i.e. of '$\forall x\ (\varphi x \equiv x =$ Herbert)', that it is something which, in conjunction with 'ξ snores', will form the proposition 'Only Herbert snores'. The division into 'Only Herbert φ' and 'ξ snores' is one possible decomposition of that

[17] This type-notation is designed, like 'Polish' logical notations, to be unambiguous without brackets or other punctuation ('2PN1PN' has only one possible parsing); but, again like 'Polish' notations, it is more readable if some redundant punctuation is included, and from now on I will do that.

[18] If all brackets are included, counting them outwards gives the level.

proposition, but the proposition is *formed*, that is constructed, in the quite different way given by its analysis, which would trace some such[19] constructional history as this:

Snores (Alan)   Alan = Herbert

\   /

Snores (Alan) ≡ Alan = Herbert

|

Snores (ξ) ≡ ξ = Herbert

|

∀*x* (Snores (*x*) ≡ *x* = Herbert)

Now consider on the other hand 'Everyone snores', with its analysis,

Snores (Alan)

|

Snores (ξ)

|

∀*x* Snores (*x*)

Here it would be almost as misleading to say, in line with the remainder reading, that 'Everyone φ', or '∀*x* φ(*x*)', is what remains of this proposition when the predicate 'ξ snores' is removed. The quantifier 'Everyone φ' is no mere remainder: it is that by the application of which the proposition is constructed.

The type-notation we have presented is insensitive to such differences.[20] In Frege's later terminology, these types carve things adequately at the level of reference, where 'Only Herbert φ' and 'Everyone φ' are just two second-level functions. At the level we have been working at, though, which concerns itself with how the

[19] 'Some such', because the occurrence of 'Alan' in the atomic propositions is an arbitrary choice: the complex predicate 'snores (ξ) ≡ ξ = Herbert' is conceived as arrived at by the omission of some name, but it doesn't matter which. In this limited respect Frege's conception does not assign the proposition a unique constructional history (cf. *FPL*, p. 14).

[20] So, if type-symbols are to be pronounced, it might be sensible to choose a form of words that slurs over the difference. '1P(N)', for instance, might be read: '1 item short of a proposition, viz. a name'.

significance of propositions is determined through their composition, there is an important distinction to be made. One way of making it is this. The simplest propositions in which an $n\beta(\alpha_1,\ldots, \alpha_n)$ occurs are those in which it is combined with items of types $\alpha_1, \ldots, \alpha_n$, these latter determinately occupying its argument places. For some expressions of the type, such a proposition will be analysable as constructed by the application of that expression to those arguments, so that it, as well as they, are genuinely constituents of the proposition. For other examples of the type, this will never be so: though they may of course be constituents of other, more complex propositions, in relation to these simplest propositions in which they occur these expressions can only ever be components disclosed through decomposition. The first kind will be the primitive or logically simple examples of the type; the second will be called complex. The quantifier 'Everyone φ' is, then, a logically simple second-level predicate, while 'Only Herbert φ' is a complex instance of the same type.

This is (an obvious generalization of) the distinction between simple and complex predicates which Dummett introduces in chapter 2 of *FPL* (pp. 27ff.). The fact that the classifications it makes are in many cases obvious – that it counts '$\forall x\ \varphi x$' as simple and '$\forall x\ (\varphi x \equiv x = h)$' as complex, or (seemingly) '$S(\xi)$' as simple and '$S(\xi) \supset T(\xi)$' as complex – might make one impatient with the roundabout way I have approached it. But not all cases are quite so obvious. Recall, for instance, what was said in section 5 in distinguishing 'Herbert' from the second-level predicate 'φ(Herbert)'. Of type 1P(1PN), the simplest propositions in which this predicate occurs will be those in which it is combined with a first-level predicate (a 1PN), most straightforwardly such a proposition as 'Herbert snores'. Analysis will not represent such a proposition as constructed by the application of this second-level predicate. So, by the above, 'φ(Herbert)' is complex – even if it does not immediately look to be. It is an example of what Dummett calls a 'degenerate' complex predicate (*FPL*, p. 30).

Degenerate examples bring out especially clearly that the core of Dummett's distinction between simple and complex predicates lies in the different theoretical roles that the two notions are required to fill (*FPL*, p. 27). Simple predicates serve the needs of analysis. They are among the building blocks of propositions, or, speaking linguistically, the primitive vocabulary that defines the expressive resources of a language; in this respect, at least, simple predicates belong with the names of the language. Complex predicates do not add to these resources, but represent ways in which the logical machinery of a language like Frege's exploits them to codify inferential connections. The notion of a complex predicate is needed, in the first instance, for Frege's account of quantification. There, as we saw, it is conceived as abstracted from, or formed through decomposition of, an instance of the desired generalization, and represents a pattern shared by all of these instances. More specifically, it represents a common pattern in the route by which the content or truth-condition of any instance of the generalization would be determined by its composition, as revealed in the analysis of that instance (*FPL*, p. 29). The complex predicate 'ξ snores and ξ snuffles', for instance, captures the pattern that each of its instances requires for its truth the conjoint truth of a pair of propositions, '*n* snores' and '*n* snuffles', ascribing snoring and snuffling to the same thing. Each of the instances, that is, is a conjunction, but the complex predicate is not itself a conjunction of any kind, nor is it to be thought of as put together in any other way from any ingredients. It is instead a compendious representation of a range of conjunctions, and the effect of attaching a universal quantifier to it is to assert that any conjunction formed in accordance with the captured pattern will be true.

It is just this notion of a complex predicate that Frege introduces, though not by this name, in *Bs*, §§ 9–10, because it is this notion that is needed, in accordance with the two-step model described in section 3 above, to prepare for the account of generalization that follows in *Bs*, §11. What Frege then says about predicates (or 'functions') is offered as applying to predicates so conceived, and I think we can agree with Dummett that the same holds good generally in Frege's writings (*FPL*, p. 31). This is most clearly true of the contention discussed in section 5, that the discernment of a certain predicate as occurring in a proposition has to do with our way of viewing the proposition, rather than with how the proposition is constituted in itself. It is also true of the hierarchical stratification of predicates, since it is of complex predicates that it is most clearly true that their propositional role can be characterized only by reference to the already determinate role of the arguments omitted in abstracting them; indeed, as Frege describes things, we can come to recognize a complex predicate as occurring in a proposition only given that we already recognize its argument as belonging, along with others that might replace it there, to an already determinate range of generalization. Yet while such examples show that Frege's attention was primarily directed to complex predicates, he can hardly have forgotten that the patterns they represent must be patterns *in something*. Whether we say they represent patterns in sentences, in the propositions these sentences express, or in the constructional histories by which the significance of these propositions is determined, we are directed equally to the building blocks of these constructions, simple predicates among them. So we should say something to explain why it is that, although Frege does on occasion clearly recognize this fact (*PW*, 17), he does not much highlight it (cf. *FPL*, p. 30; *IFP*, p. 292).

The first and easiest point to make here is no more than a reminder of something we saw early on. Frege introduced his conception of a predicate in connection with examples (like *c|Ho* and *cH|o* from section 5 above) where the prior construction from names and a simple predicate, and the way this prior construction makes possible the decompositions Frege illustrates, is simply too obvious to need mentioning. A second point is equally straightforward. The general notion of a complex predicate, conceived as abstracted from a proposition, is key to his account of quantification, and therefore constitutes Frege's distinctive contribution to logic; it is hardly surprising that his discussions give it top billing (*FPL*, p. 32). A final point, though, is slightly more delicate. As Frege says, if a proposition is to be decomposable in accordance with his new conception, 'it must already be itself articulated' (*PW*, p. 17). But the demands this conception places on just *how* it is articulated are minimal. Frege's account presumes that certain basic elements, which we have called names, can be distinguished in propositions, and understood as filling a role there that would be differently filled by other names which might replace them. He presumes, for instance, that in the proposition 'hydrogen is lighter than carbon-dioxide' we can recognize the element 'hydrogen', and understand it to be replaceable by 'oxygen' or 'nitrogen'. On such replacement '"oxygen" or "nitrogen" enters into the relations in which "hydrogen" stood before' (*Bs*, §9), but just what those relations are Frege need not say. We know of them only what is implied in the above, namely, that by whatever means they enable the original proposition to be understood as expressing a condition on hydrogen that can be considered as holding also of oxygen or nitrogen. From this minimal basis Frege's new conception takes over and generates of itself, in the way described at the start of this section, the whole hierarchy of dependent roles. Moreover, the minimalism of the basis is reflected in a certain thinness in the way these roles are characterized. The basis dictates of a 1PN only that it expresses a condition on an N, hence of a 1P(1PN) only that it expresses a condition on such conditions; and this is something that the simple quantifier 'Everyone φ' and the complex 'Only Herbert φ' both do. In other words, without assuming any more than we have said about the functioning of the basic, simple constituents of propositions, Frege's conception supplied him with a system of roles for them to fill. He was thus able to construe them, and subsequently did construe them, as filling (some of) these roles. In this sense Dummett is clearly right to represent Frege as having 'tacitly assimilated simple predicates to complex ones' (*FPL*, p. 30).

Is the assimilation legitimate? It certainly does, as we observed, suppress important differences, differences that emerged in our uncertainty between the 'construction' and 'remainder' readings of type-symbols. But overlooking differences, e.g. between Austrians and other German-speakers, is not wrong; for given purposes, the broader category might be the important one. Is the case comparable to that one, or is it more like claiming Austrians to be Germans, which they are not?

It is not immediately clear whether Dummett has a consistent answer to this question. For several reasons I opted above to introduce the distinction of simple and complex predicates in connection with second-level, rather than first-level examples. This made it easier to separate the core of the distinction, which we saw has to do with the different theoretical roles of the notions, from further theses Dummett has maintained about first-level cases which are either dubious or which have prompted misguided criticism. For an example of the first kind, Dummett claimed that, whereas a complex predicate must be viewed as a feature of a sentence, or a pattern in it, a simple first-level predicate is an isolable part of a sentence, 'as much a linguistic entity capable of standing on its own as are … proper names' (*FPL*, p. 28). In later discussion he retreated somewhat from this claim (*IFP*, p. 318), in response to points made by Geach;[21] but in any case the claim is not central to the distinction, and would not have been made about the simple second-level predicate 'Everyone φ'. Other criticisms from Geach provide examples of the second kind. Dummett had observed, quite rightly, that complex predicates form 'the prototype for Frege's general notion of an incomplete expression' (*FPL*, p. 31); and he had gone on to observe that the clearest sense in which complex predicates are incomplete – namely, that they are formed by abstraction, as what remains when some component of a proposition is omitted, and so are reasonably counted incomplete propositions – does not hold of simple predicates. Geach protested that this 'false and unFregean doctrine' casts both the sense and the reference of a simple predicate as a species of object, and so undermines Frege's account of propositional unity.[22] In the first place, though, Dummett's distinction is clearly compatible with acknowledging that simple and complex predicates are not distinguished at the level of reference (see p. 106 above), and in consequence that the sense of a simple predicate, being the way in which an incomplete referent is determined, is itself incomplete (*IFP*, p. 319). And in the second place, Frege's resolution of the problem of propositional unity does not in any case lie in his notion of incompleteness, but is entirely contained in his context principle: it is only *because* Frege allowed complex predicates to subsume simple ones that he so often presented this resolution in terms more suited to the former than the latter. I hoped, then, to avoid such tangles by introducing the distinction by second-level examples. But the most straightforward reason for this choice was just to emphasize that the distinction is a general one, applicable in principle at any level, and applying in fact at any level where the language includes simple predicates. (Frege's language includes simple third-level predicates.) This would lead one to expect that, if the assimilation of simple to complex is objectionable, it will be equally objectionable at whatever level. But Dummett's description of Frege's hierarchy suggests otherwise. At the first level he holds that 'strictly speaking, Frege ought to have treated separately of simple … predicates', though 'with a

[21] P. T. Geach, 'Names and identity', in S. Guttenplan (ed.), *Mind and Language* (Oxford: Clarendon Press, 1975), p. 148.
[22] Geach, 'Critical notice of *FPL*', pp. 444–5.

certain inaccuracy' he 'preferred to subsume them under the general category of first-level predicates' (*FPL*, p. 38). Moving on to characterize second-level predicates, he remarks, by contrast, that a quantifier is 'precisely such an expression' (*FPL*, p. 39); and later he confirms, with no suggestion that any inaccuracy is involved, that 'the universal quantifier (binding individual variables) is a simple sign of type [[i]]' (*FPL*, p. 48) – i.e. of type 1P(1PN). Why this asymmetry?

It will help clear the way to an answer to this question to consider first a naive complaint – hence, not Dummett's complaint – against the assimilation at the first level, and a rather bluff counter to it. The complaint takes off once again from the reasoning that earlier forced us to distinguish the name 'Herbert' from the second-level predicate 'φ(Herbert'), and which we can now set out in more explicitly hierarchical terms:

Herbert snores	P	0
Herbert	N	0
ξ snores	1P(N)	1
φ(Herbert)	1P(1PN)	2

The two must be distinguished, we said, because a characterization of the role that 'φ(Herbert)' plays in a proposition depends, via that of 'ξ snores', on a prior characterization of the role of 'Herbert'. The complaint construes this as a distinction between an abstracted component or pattern in the proposition, 'φ(Herbert)', and a genuine constituent of the proposition, 'Herbert', whose presence there makes possible that particular abstraction. And it now asks (speaking, as I said, naively): if Herbert must appear twice in this story, how is it that snoring appears only once? For 'ξ snores', when it appears here as a 1P(N), is likewise cast as an abstraction from the proposition, and there must surely be, in its case too, some constituent in the proposition whose presence there makes possible this particular abstraction. This constituent can only be the simple predicate 'snores', which thus deserves to be explicitly and separately acknowledged; but it is simply missing from Frege's scheme.

The bluff counter to this complaint charges it with railing against simple arithmetic. It is an immediate and obvious feature of the way the hierarchy is generated – by considering what remains constant when something is varied, and then what remains constant what that in turn is varied, and so on – that anything at level n will find a reflection at level $n+2$: what first appears as the bearer$_n$ of certain properties$_{n+1}$ is automatically reflected in the property$_{n+2}$ of those properties$_{n+1}$ that they hold of it$_n$. It is an even more obvious feature of the conception that there are no 'negative' levels, so no level that finds its reflection at level 1. Yet the complaint is asking that the first of these features be respected at the cost of the second, and it clearly cannot have it both ways. Moreover, the relation between a level n entity and its level $n + 2$ reflection is in any case not, as the complaint construes it, that between grounding constituent and derivative abstraction, because it will hold equally (for $n \geq 1$) when what we have at level n is only an abstracted component.

What we can gather from this naive exchange, at this stage, is only a warning against turning suggestions carried by one feature of the hierarchical conception into demands that are then pressed in contradiction to other features of this same conception. The warning will be relevant when we have considered Dummett's very different, and far from naive, complaint against Frege's assimilation.

He writes:

> Once we have acquired the notion of a complex predicate, we cannot refuse to allow, as a degenerate case, the 'complex' predicate 'ξ snores', considered as formed from such a sentence as 'Herbert snores' by omission of the name 'Herbert'; it would then seem quite redundant to insist on considering, as a separate linguistic entity, the simple predicate '... snores'. Nevertheless, it remains the case that, strictly speaking, if 'ξ snores' is treated as a complex predicate, on all fours with, say, 'If anyone snores, then ξ snores', we do need to recognize the separate existence of the simple predicate '... snores' as well: for, precisely because the complex predicate 'ξ snores' has to be regarded as formed from such a sentence as 'Herbert snores', it cannot itself be one of the ingredients from which 'Herbert snores' was formed, and thus cannot be that whose sense, on Frege's own account, contributes to composing the sense of 'Herbert snores'. (*FPL*, p. 31)

In speaking of how expressions are formed we describe relations of dependence holding between their sense. The 'complex' predicate 'ξ snores' is formed from a proposition such as 'Herbert snores' in that, as a general explanation of the significance of the quantifiers requires, the condition for this predicate to hold of a given object is explained as the condition for the truth of a proposition in which a name of that object occupies the argument-place of the predicate. That condition in turn must be explained by reference to the name and the predicate that constitute the proposition: such a proposition will be true if the named object satisfies the condition for the predicate to hold of it. But if we now try to explain *this* condition, as before, as consisting in the condition for the truth of a proposition in which a name of the object completes the predicate, then 'we shall go round in a circle'.

> We cannot explain what it is to grasp the condition for ['Herbert snores'] to be true in terms of grasping the condition for ['ξ snores'] to be true of an arbitrary [object], and then explain what it is to grasp *that* condition in terms of grasping the condition for ['George snores', 'John snores', 'Herbert snores'], etc., to be true. (*IFP*, p. 293, with example changed)

Circularity can be avoided, it seems, only if the satisfaction condition of the predicate quantified in 'Everyone snores' is distinguished from that of the predicate occurring in an instance of that generalization; and, if these conditions are distinguished, then our analysis must also distinguish the expressions whose sense is given by them.

There is, I think, only one way to counter this argument, and this way becomes apparent when we ask why the argument would not hold at a higher level. Just as first-order quantification requires us to recognize the general category of first-level predicates, so second-order quantifiers require the general category of second-level predicates. In each case the quantifier must be explained univocally, and so in a way that accounts for its application to complex predicates, although some of those to which it is applied will be simple (or perhaps only 'degenerately' complex). Let us, assuming it to be primitive, take '$\exists x\ \varphi x$' as a second-level example, corresponding to Dummett's first-level 'ξ snores'. '$\forall F\ \exists x\ Fx$' is explained, as is '$\forall x$ Snores(x)', as requiring the truth of all its instances. The truth-condition of an instance will then be explained by reference to the condition for the quantified predicate, '$\exists x\ \varphi x$' or 'ξ snores', to be true of what is referred to by the expression figuring as its argument in that instance. It is in the attempt to explain *this* condition, in Dummett's case, that circularity threatens; whereas, of course, in our case, of the simple second-level predicate '$\exists x\ \varphi x$', explanation proceeds smoothly down through the analysis, towards the condition for whatever replaces φ to hold of its arguments. What this shows is that the threatened circularity derives, not from the fact that the predicate whose application condition we are trying to explain is simple, since that is true in both cases, but from the more obviously relevant fact that the instance in which it is applied is an atomic proposition.

This explains the asymmetry in Dummett's remarks that we noted a few pages ago, but it also undermines the ground for it. For now that the source of the threatened circularity is clear the correct response to it is to accept it as a mark of an atomic proposition that this circularity cannot be avoided. This response does not question any of the premises of Dummett's argument. It remains true, and an essential part of Frege's conception, that

> (A) grasp of the condition for a predicate to hold of a given entity

is *in general* to be distinguished from, and explained by,

> (B) grasp of the condition for the truth of a proposition in which a designation of that entity occurs as argument to the predicate.

It is also true, as Dummett says, that explaining

> (A1) what it is for a simple, atomic predicate to hold of a given object

by reference to

> (B1) the condition for the truth of an atomic proposition in which a name of the object completes the predicate

gets us precisely 'nowhere'. But this is no ground for complaint. The reason it gets us nowhere is that there is nowhere further to go: analysis, having reached its atomic basis, terminates here. We cannot have, and should not expect, any account of what it is to grasp the application-condition of an atomic predicate that represents it as prior to, or as offering a non-circular explanation of, the ability to use that predicate in framing judgements. Rather the two are, as Frege said, simultaneous (*PW*, p. 17).

To insist on distinguishing the simple predicate '… snores' from that instance of the general category of first-level predicates that we have represented 'ξ snores', and thus to resist Frege's assimilation, is, I have claimed, to demand something impossible. In *this* respect (though only in this respect) Dummett's argument has something in common with the naive complaint countered above, and this emerges in his description of 'ξ snores' as a 'degenerate case' of a 'complex' predicate (*FPL*, p. 30). There should be no general objection to this notion. Earlier I adopted it from him, and agreed that it applies to the second-level 'φ(Herbert)'. This predicate is complex, by the explanation given, because analysis of its simplest occurrences, e.g. 'Herbert snores', will not represent them as involving it as a constituent; it is only a degenerate case because it is not, in any ordinary sense, compound. But note that this explanation presupposes the hierarchical structure we have described, as providing the more basic analysis of 'Herbert snores' in which 'φ(Herbert)' does not figure. This structure provides no more basic analysis of the simplest occurrences of 'ξ snores' than that they are completions of it. So to describe *this* as a degenerate complex predicate is to exploit a notion explained by the hierarchical structure in hankering after a still more basic analysis than any this structure provides; whereas, I have suggested, the only correct response for someone who respects this structure as providing the framework for analysis is to accept, as Frege did, that there can be no such account. To repeat, there is nothing wrong with the general notion of a degenerate complex predicate. But there are no first-level examples of it. There is, then, no need to distinguish the simple predicate that occurs in 'Herbert snores' from any such thing.

What this shows, I think, is that there can be no such *slight* inaccuracy as Dummett complains of in Frege's assimilation of simple predicates to the general category of first-level predicates. So far as concerns the interpretation of Frege this is a point of disagreement with Dummett, but one that is entirely peripheral when compared with the agreement against which it is set. His core distinction of simple and complex predicates is unquestionably sound, and is, as he says, 'not merely consonant with Frege's views, but important for the avoidance of a misunderstanding of them' (*IFP*, p. 292). So far as it concerns instead the evaluation of Frege's views, I am less sure that it marks any disagreement. For it implies that, if there is any sound thought underlying either of the complaints against Frege's assimilation that we have considered, then this thought cannot be accommodated by a *slight* revision to Frege's conception: no modest tweak of the conception will supply a place for what this thought hankers after. Satisfying the thought will call instead for a much more radical separation between the hierarchically structured categories required in the explanation of quantificational inference and any system of categories purporting to represent the articulation of atomic propositions and the facts they portray, a separation, that is, between the categories yielded by logic and those needed by ontology. This is the drift of the deeper challenge to Frege mentioned at the close of section 4 as deriving first from Wittgenstein but then more clearly and more powerfully from Ramsey. It is one that Dummett has very much in mind, and one whose force he clearly recognizes (*FPL*, pp. 61–7; *IFP*, pp. 319–22).

Assessing this challenge lies, as I said, far beyond the scope of this chapter; but I hope enough has been said to indicate that no reasonable assessment could proceed without the deep understanding of the issues involved that Dummett's discussion provides. For this reason I believe that the case we have looked at provides an example – just one example among many – of how thinking through Dummett's exposition is simply the same thing as thinking through what is most important in Frege.

Perhaps, though, that is too obvious to need saying. It certainly should be.

ALEX OLIVER

5 What is a predicate?

According to Frege, the sentence 'Socrates is mortal' can be analysed into the proper name 'Socrates' combined with the one-place predicate 'ξ is mortal'. It is uncontroversial that the two items are governed by different grammatical rules of combination. But Frege also introduced semantic distinctions between them that have been contended ever since. He made them refer to different types of thing. The proper name refers to an *object*, while the predicate refers to a *concept*, which he proceeded to identify with a function from objects to truth-values, e.g. from Socrates to truth. His distinction between objects and functions is exclusive: nothing can be both. He marked the difference by saying that an object is complete or saturated, whereas a function is incomplete or unsaturated. Proper names and predicates also differ in the way in which they refer to their referents. Frege uses 'refers to' as an umbrella term covering different relations, since his principle for individuating them dictates that the reference relation holding between a proper name and an object is of a different type from the relation holding between a predicate and a concept. He made further semantic distinctions between proper names and predicates at the intermediate level of *sense*, but discussion of them will not be necessary, except in §2.4.

These syntactic and semantic contrasts between proper names and predicates are relational in character. My question is whether proper names and predicates also differ in their intrinsic nature. Are they different sorts of thing, as different perhaps as the corresponding objects and concepts?

Frege answered that proper names and predicates 'differ essentially'.[1] Just as nothing can be both an object and a function, so

[1] Frege, letter to Russell, 13 November 1904, in *PMC*, pp. 160–6, at p. 161.

nothing can be both a proper name and a predicate. He characterized the difference between them by redeploying the adjectives that he used to contrast their worldly referents. Proper names, like objects, are said to be complete and saturated; predicates, like concepts, are not. There is little agreement about what Frege meant by this or whether he was right. The dispute focuses on predicates, since it is generally accepted that proper names can straightforwardly be construed as expressions. In part 2 I shall describe and evaluate Frege's account of predicates, taking issue with other exegetes. But first I need to map out various candidates for predicates and investigate the apparent competition between them.

1 PREDICATES

1.1 Four candidates

Plain expressions. I start with the simplest case, an atomic sentence '*Fa*' of the predicate calculus. Generations of logic students have learned to identify the expression '*F*' as the predicate occurring in '*Fa*'. Like the sentence, the expression is an abstract type. It may have any number of concrete tokens, of different forms and in different substances. And the type itself may reoccur in other sentences, or within a single compound sentence. But its grammar dictates that it occurs only once in an *atomic* sentence – at the front – and is followed by just one term. The grammatical rule governing the predicate '*F*' invokes a general method of constructing sentences, namely *predication* of a one-place predicate (we shall come to predicates of higher degrees when the need arises). This method of construction may be identified with the *linguistic* function that maps two arguments – a one-place predicate Φ and a term α – to a single value – the sentence $\Phi\alpha$.

Some authors characterize predicates in just these constructive terms: 'Expressions that yield sentences when thus attached to singular terms are called *predicates*'.[2] One might equally well come at predicates by decomposition or subtraction: '*F*' is what is left when '*a*' is subtracted from '*Fa*'. A third account would emphasize the

[2] Gerald J. Massey, *Understanding Symbolic Logic* (New York: Harper & Row, 1970), p. 226.

sharing of a single predicate by related sentences: '*F*' is the common element among '*Fa*' and its substitutional variants '*Fb*', '*Fc*', etc. Nothing hinges on the differences between these descriptions. They describe the same thing: the expression '*F*'.

I now introduce three further candidates, each intimately associated with the expression '*F*', and each leading naturally to the next. Each has some right to be called a *pattern*, but to avoid conflation and confusion, I give them different names.

Schemata. The sentences '*Fa*', '*Fb*', '*Fc*', etc. share a pattern '*Fx*'. In one sense of pattern, the common pattern is the very expression '*Fx*'.[3] The '*x*' in '*Fx*' functions as a schematic letter, and '*Fx*' is itself sometimes called a schema, which is the label I shall use; I sometimes call '*F*' a *plain* expression to distinguish it from the schema '*Fx*'. Accordingly, when one says that '*Fa*', etc. share the pattern '*Fx*', one means that they are each the result of substituting a term for '*x*' in '*Fx*'.

Linguistic functions. The substitution procedure applied to the schema '*Fx*' yields a corresponding linguistic function: *the result of substituting the term … for 'x' in 'Fx'*. Call this function *f* for short, then the sentence '*Fa*' is the value of *f* for the term '*a*' as argument, while the variants of '*Fa*' are values of the same function for different arguments.

According to a second sense of 'pattern', the pattern common to '*Fa*' and its variants is not the schema '*Fx*' but the corresponding linguistic function *f*. So when we say that '*Fa*' and its variants all share or illustrate a common pattern, we now mean that they are all values of the same function for different arguments.[4]

Instead of using an '*x*' in a schema, others might use a '*y*', or dots or lines or circled numerals or Greek consonants ('*F*…', '*F*___', '*F*①', '*Fξ*'). The choice is plainly arbitrary, but it prompts a search for a sense of pattern according to which these trivially different schemata all depict the same pattern. Our second sense of pattern gives us what we want, since the same linguistic function may be described in terms of substitution of a term for e.g. the letter '*x*' in '*Fx*', or the dots '…' in '*F*…'. Of course, it may also be described

[3] See e.g. Christopher Kirwan, *Logic and Argument* (London: Duckworth, 1978), p. 3.

[4] See e.g. Peter Geach, 'Names and identity', in Samuel Guttenplan (ed.), *Mind and Language* (Oxford: Clarendon Press, 1975), pp. 139–58, at pp. 142–3.

without making any reference to a schema, e.g. *the result of attaching the term … to the expression 'F'*.

The linguistic function *f* is related to, but distinct from, the linguistic function that I earlier identified with the predication construction. The latter takes *two* arguments, e.g. it maps '*F*' and '*a*' to '*Fa*', whereas *f* takes just *one*, e.g. it maps '*a*' to '*Fa*'. In effect, the expression '*F*' has been absorbed into the function *f*.

Properties of sentences. The sentence '*Fa*' is the value of the function *f* for the argument '*a*'. Hence '*Fa*' has the property *being the value of f for some term as argument*, a property which it shares with its variants '*Fb*', '*Fc*', etc. This common property may be called a pattern in yet another sense of the word.[5]

I have derived the common property from the corresponding linguistic function. But it too can be described more directly, e.g. as the property *consisting of 'F' followed by a term*. As well as depicting the function *f*, the schema '*Fx*' may be understood in a different way as depicting this property of sentences.

As I have explained, the pattern common to '*Fa*' and its variants may be regarded as either a schema, or a linguistic function, or a property of sentences. Indeed, the relevant *predicate* may be regarded as any of these things. Including the plain expression '*F*' with which we started, we now have four candidates. So: what is a predicate?

Different authors have given different answers. One has even given different answers at different times. Geach started by defining predicates as expressions: 'A *predicate* is an expression that gives us an assertion about something if we attach it to another expression that stands for what we are making the assertion about'.[6] Within three years, however, he had come to deny that a predicate is an actual expression: 'I should rather regard it as a common property of sentences'.[7] Finally he championed linguistic functions:

in 'Raleigh smokes', let us say, the two terms have totally different modes of significance. 'Raleigh' signifies just by being a man's name. We cannot

[5] See e.g. Michael Dummett, *Frege: Philosophy of Language* (London: Duckworth, 1973), pp. 31, 246, 250, who also speaks of 'features' of sentences.

[6] Peter Geach, 'Subject and predicate', *Mind*, 59 (1950), pp. 461–82, at p. 461.

[7] Peter Geach, 'Quine on classes and properties' (1953), reprinted in his *Logic Matters* (Oxford: Blackwell 1972), pp. 222–6, at p. 224.

sensibly ask what 'smokes' names; what is significant is not the bare word 'smokes' but a certain pattern – name followed by 'smokes'. And speaking of a common pattern in 'Raleigh smokes', 'Churchill smokes', etc., is just another way of saying what Frege would have expressed by saying we had values of a common function for a series of different arguments – the names 'Raleigh', 'Churchill', etc.[8]

1.2 Anything goes

Whereas other authors argue for different candidates, I propose that there is nothing to choose between them. Anything goes: each is equally serviceable. The rich multiplicity of candidates (there will be more) is not an embarrassment. The choice between them can be made arbitrarily, or, when the context allows, it can be left unmade.

I therefore need to rebut arguments that seek to show that some candidates are not fit for purpose. In the literature, the objections are targeted against plain expressions and against schemata. (Geach himself subsequently admitted that the difference between linguistic functions and properties of sentences 'seems to matter little'.)[9] Before rebutting them, it is important to emphasize that for any choice of candidate, we can give an appropriate sense in which a predicate *occurs* in a sentence, and so an appropriate sense in which a predicate can be *shared* by different sentences. When a predicate is construed as a plain expression like the '*F*' in '*Fa*' and '*Fb*', it occurs in them by simply being a part of them, albeit in a sense of 'part' that is appropriate for abstract types rather than concrete tokens. If another candidate for predicates is selected, an obvious compensating change must be made to the sense of 'occurs'. The schema '*Fx*' occurs in '*Fa*' means that '*Fa*' is the result of substituting a term for '*x*' in '*Fx*'. The associated linguistic function *f* occurs in '*Fa*' means that '*Fa*' is the value of *f* for some term as argument. And the corresponding property of sentences occurs in '*Fa*' simply means that it has the property. These accounts of occurrence are tailor-made for atomic sentences. More complex contexts demand more complex

[8] Geach, 'Names and identity', pp. 141–2.

[9] Peter Geach, 'Saying and showing in Frege and Wittgenstein', *Acta Philosophica Fennica*, 28 (1976), pp 54–70, at p. 61.

accounts of occurrence. But since all four candidates are so closely linked, what works for one can easily be adapted for the others.

1.2a Objecting to schemata

In the last paragraphs there is a certain linguistic awkwardness, of some philosophical interest. It is natural to speak of … the two-place predicate 'ξ killed ζ'; but here, as Frege would put it, by a kind of linguistic necessity we cannot quite say what we are trying to say … if we speak of the predicate 'ξ killed ζ' as figuring in 'John killed Mary' or 'Mary killed John', then again what we quote does *not* figure in these sentences. The actual expression 'ξ killed ζ' is neither a function nor a predicate: it serves however to identify a two-place predicate shared by many sentences, and this is the same thing as identifying a function yielding such sentences as 'John killed Mary' and 'Mary killed John' as its values when proper names are supplied as its arguments.[10]

Geach here thinks that the phrase 'the two-place predicate 'ξ killed ζ'' can only be understood as picking out a schema. Since he regards predicates as linguistic functions, 'we cannot quite say what we are trying to say'. But he himself allows that schemata depict (or, as he puts it, 'identify') linguistic functions. So in fact we could use another mode of expression to say what we are trying to say, namely 'the two-place predicate depicted by 'ξ killed ζ''. There is no need to do so, however. His account of quotation is too restrictive. The phrase 'the two-place predicate 'ξ killed ζ'' does not pick out a schema as a matter of 'linguistic necessity', since we can understand the quotation marks to be directing us towards any of the candidates for predicates, without the detour via depiction by an intermediary schema. Some famous philosophers' pronouncements notwithstanding, material enclosed in quotation marks may stand for a wide variety of things. Common examples are: expression tokens, whether written or spoken, meaningless or meaningful; expression types for all the different ways of typing them; meanings. As the Kneales quipped: 'Quotation marks were made for man, not man for quotation marks'.[11] Logicians understandably baulk at

[10] Geach, 'Names and identity', pp. 148–9.
[11] William and Martha Kneale, *The Development of Logic* (Oxford: Clarendon Press, 1962), p. 514.

such ambiguity, and invent different styles of quotation marks in order to resolve it. But there is a more relaxed way to disambiguate in ordinary use, namely the addition of an explanatory prefix, as in 'the printed token "loves"', 'the phonological type "loves"', 'the lexeme "loves"'.

To take our current topic, then, suppose that predicates are construed as linguistic functions. By using the prefix in 'the linguistic function "ξ killed ζ"' we can pick out the relevant function directly. Once it is settled that linguistic functions are the chosen candidates, the same goes for 'the two-place predicate "ξ killed ζ"'.

This point about quotation aside, why does Geach suppose that we cannot really mean to mention the schema when we use the phrase 'the two-place predicate "ξ killed ζ"', i.e. why cannot predicates be schemata? He claims that the schema 'ξ killed ζ' does not occur (or, as he says, 'figure') in 'John killed Mary', evidently on the ground that the schema contains Greek letters while the sentence does not. He has failed to understood 'occur' in the sense appropriate to the candidate. The schema 'ξ killed ζ' does indeed occur in 'John killed Mary' in the relevant sense, since the sentence results from substituting terms for the schematic letters. Geach cannot object to this reinterpretation of 'occur' to suit the candidate, for he himself must say that a linguistic function – his favoured candidate – occurs in the sentence in a sense of 'occur' quite different from that in which the plain expression 'killed' occurs.

1.3 Objecting to plain expressions

I now turn to a pair of arguments that aim to show that some or all predicates cannot be plain expressions. The first is Dummett's.[12] He draws a distinction between two kinds of predicate according as they play different explanatory roles. *Simple* predicates are needed to explain the syntactic and semantic structure of atomic sentences. *Complex* predicates, on the other hand, are introduced to explain the structure of sentences that feature variable-binding devices such as quantifier phrases, and to explain and to represent schematically the validity of arguments featuring such sentences. To use Dummett's own examples, 'Brutus killed Caesar' contains

[12] Dummett, *Frege: Philosophy of Language*, pp. 27–33.

the simple predicate 'ξ killed ζ', which is used in accounting for the construction and understanding of the sentence. But we might also need to represent the sentence as containing the complex predicate 'ξ killed Caesar' when explaining and representing the validity of an argument in which the original sentence figures alongside the quantified 'anyone who killed Caesar is an honourable man'.

Dummett claims that this difference in role is matched by a difference in nature: simple predicates are not the same sort of thing as complex ones. He asserts that simple predicates are plain expressions. The Greek letters in 'ξ killed ζ' are not elements of the simple predicate represented, but are merely used to indicate the location, nature and number of its arguments (this is another example of how material enclosed in quotation marks may be used to pick out different candidates for predicates in different contexts). Dummett quite reasonably counts even a discontinuous string of words as a plain expression, e.g. the simple predicate 'ξ took ζ to task' consists solely of the two discontinuous parts 'took' and 'to task'.

In contrast to simple ones, complex predicates are not expressions at all, but 'features' or 'patterns' of sentences, by which Dummett means properties of them. In order to establish this difference in nature, he focuses on just one, quite special kind of complex predicate, namely those that have argument-places that are, as Frege says, 'related',[13] i.e. places that must be occupied by occurrences of the same term. Related argument-places are indicated by repeating Greek letters, as in 'ξ killed ξ'. According to Dummett,

> There is no part in common to the sentences 'Brutus killed Brutus' and 'Cassius killed Cassius' which is not also part of the sentence 'Brutus killed Caesar': yet the predicate 'ξ killed ξ' is said to occur in the first two and not in the third. Such a complex predicate is, rather, to be regarded as a *feature* in common to the two sentences … it does not consist merely of some sequence of words or symbols … the complex predicate is thus not really an expression – a bit of language – in its own right.[14]

It is plain that Dummett's argument is limited in scope, since it only applies to predicates with related places. For all that he has said, a complex predicate without related places, e.g. his 'ξ killed Caesar',

[13] *Gg*, vol. I, §4.

[14] Dummett, *Frege: Philosophy of Language*, p. 31.

can be construed as a plain expression. Moreover, even if the argument works for predicates with related places, it can only show what they are *not*. It cannot establish Dummett's positive conclusion that complex predicates are properties of sentences. Why not schemata or linguistic functions? The same may be said against Geach's use of an identical argument, first to argue that predicates are properties of sentences,[15] and later to argue that they are linguistic functions.[16]

In any case, the argument does not even succeed on these limited and negative terms. Dummett tells us that an expression is 'a sequence of phonemes or of printed letters' or a word or string of words 'which can quite straightforwardly be written down'.[17] Yet even simple predicates cannot be expressions in this attenuated sense, since a predicate needs to be distinguished from its homophones and homographs, which may not obey the same rules of combination. In other words, if we use Dummett's notion of an expression, a simple predicate cannot be an expression *simpliciter*, but only an expression coupled with a grammatical rule. He himself notes that his simple predicate 'killed' is governed by the grammatical rule that it goes with two terms, *the same or different*, one on each side. But then he must allow that the same expression 'killed' may be governed by the different grammatical rule that it goes with the *same* term on either side. Although the expression remains constant, the predicate is different, since the expression is coupled with a different grammatical rule. Of course, inspecting the plain 'killed' does not by itself reveal the intended predicate's grammar. But it can be easily described, or else indicated by letters within quotation marks: 'ξ killed ζ' vs 'ξ killed ξ'.

It follows that simple and complex predicates, even those with related places, swing together ontologically. In particular, both can be regarded as expressions, in Dummett's sense, coupled with grammatical rules. Nothing changes if expressions are individuated more finely, by building grammatical rules into their identity conditions. It would still be true that both kinds of predicate may be construed as expressions. This not to deny that simple and complex predicates play different explanatory roles, but it is to deny Dummett's thesis that the distinct roles are reflected in distinct natures.

[15] Geach, 'Quine on classes and properties', p. 224.
[16] See e.g. Geach, 'Names and identity', pp. 139–40.
[17] Dummett, *Frege: Philosophy of Language*, p. 32.

Noonan tackles both predicates and functors in one go, and gives an altogether different argument against regarding them – all of them – as plain expressions. He favours regarding them as linguistic functions (his 'patterns'), and claims that we are inclined to regard them differently as expressions because

> in writing down the patterns exhibited by complex designations of numbers (like '2 + 3') or sentences (like 'Socrates is wise') we typically employ auxiliary expressions ('+', 'is wise') to construct the patterns thus exhibited. But we do not *always* do so (in mathematical symbolism the sign for the two-argument function *x raised to the power y* is 'x^y' and here there is no separable auxiliary expression which anyone could regard as the sign for the function), and we need *never* do so.[18]

It is true that '2^4' contains no expression standing for the function involved. It features only an *arrangement* of the terms '2' and '4'. But from the facts about this particular case it hardly follows that *no* functor, actual or possible, can be construed as an expression. When an 'auxiliary' expression does happen to be present, as in '2 raised to the power 4', it is a serviceable candidate for the relevant functor. Again, that there is a *possible* language in which no functor is an expression (Noonan's 'we need *never* do so') is compatible with the fact that in *our* language expressions comprise one kind of candidate for many functors. As for functors, so for predicates.

Noonan works with a crude dichotomy between two candidates: plain expressions vs linguistic functions. But contrary to his intentions, his example actually serves to introduce yet another candidate – a fifth – for predicates and functors, which I omitted from the initial list in §1.1. In '2^4' the relevant functor may be identified with the arrangement of the two terms, i.e. a particular relation holding between them. Turning to predicates, one candidate for a predicate in 'Brutus killed Caesar' is the relation that holds between two terms, the same or different, when one is to the left of the expression 'killed' and the other to the right. This idea will be familiar to readers of the *Tractatus*.[19] The corresponding candidate for a one-place predicate is not a relation between terms but a

[18] Harold Noonan, *Frege* (Oxford: Polity Press, 2001), p. 147.

[19] Ludwig Wittgenstein, *Tractatus Logico-Philosophicus* (London: Routledge and Kegan Paul, 1922), 3.1432.

property of a single term, which is distinct from but related to the property of *sentences* discussed in §1.1.

2 FREGE

2.1 Frege's own candidate: expressions with empty places

I have now assembled almost all of the materials required to understand and evaluate Frege's own account of predicates. I need to add two points about his logic which I have so far glossed over. First, he uses 'proper name' in an idiosyncratic, extended sense as including not only names such as 'Socrates' but also complex singular terms such as definite descriptions, and functional value terms obtained by applying a functor to its arguments, like '2 + 3'.[20] The second point concerns the kinds of linguistic item that take arguments and produce values. As with the case of 'occur' in §1.2, 'take' and 'produce' can be understood neutrally, or they can be given a specific sense appropriate to a particular candidate. It is now usual to distinguish such items according as their arguments are singular terms or sentences and their values are singular terms or sentences. Predicates take singular terms and produce sentences, sentential connectives take sentences and produce sentences, while 'functor' is often reserved (as in §1.3) for items that take singular terms and produce singular terms. Frege, however, lumps them all together as 'function-names'. The reason is a second idiosyncrasy, namely his misconceived assimilation of sentences to singular terms. They too count as 'proper names'.[21] He does indeed speak more specifically of concept-words and relation-signs, but these do not correspond to the modern notions of one- and two-place predicates, since they can take sentences as arguments, e.g. '(2 + 3 = 5) = (2 = 2)' is well-formed. (He distinguished function-names according to the 'level' of the functions for which they stand, e.g. higher-level functions that take other functions as arguments. But I shall have nothing to say about them, since they raise no new points of principle.)

In the passages I shall be quoting, therefore, Frege is out to characterize the nature of function-names in general, rather than

[20] See e.g. Frege, 'On sense and reference', in *CP*, pp. 157–77, at p. 158.
[21] See e.g. *ibid.*, p. 163.

predicates in particular. But it should already be clear from the discussion of Noonan in §1.3 that the multiplicity of candidates for predicates is easily replicated for connectives and functors.

Frege devotes §1 of *Grundgesetze* to explaining function-names and the functions they name. He first warns against confusing functions with expressions:

If we are asked to state the original meaning of the word 'function' as used in mathematics, it is easy to fall into calling a function of *x* an expression, formed from '*x*' and particular numbers by use of the notation for sum, product, power, difference, and so on. This is incorrect, because a function is here represented as an *expression*, as a concatenation of signs, not as what is designated thereby.[22]

He proceeds to argue that even a *function-name* cannot be regarded as an expression featuring the letter '*x*' such as '$2 + 3x^2$', where we would call '*x*' a free variable. For this kind of expression 'indeterminately indicates' a value of the relevant function, and therefore does not stand for the function itself. He does not conclude that function-names are not expressions at all, only that they are expressions of another, special kind: 'the expression for a *function* is *in need of completion, unsaturated*' (§1). What does this mean?

Frege repeatedly says that function-names are expressions with empty places. He means what he says; he never offers anything that would defeat a literal reading. I give four choice quotes; there are plenty more.

the expression for a function must always show one or more places that are intended to be filled up with the sign of the argument.[23]

The name of a function is accompanied by empty places (at least one) where the argument is to go; this is usually indicated by the letter '*x*' which fills the empty places in question. But the argument is not to be counted as belonging to the function, and so the letter '*x*' is not to be counted as belonging to the name of the function either. Consequently one can always speak of the name of a function as having empty places, since what fills them does not, strictly speaking, belong to them.[24]

every function sign must always carry with it one or more places which are to be taken by argument signs; and these argument places – not the

[22] *Gg*, vol. I, §1.
[23] Frege, 'Function and concept', in *CP*, pp. 137–56, p. 141.
[24] Frege, 'Comments on sense and reference', in *PW*, pp. 118–25, at p. 119.

argument signs themselves – are a necessary component part of the function sign.[25]

function names must differ essentially from proper names, the difference being that they carry with them at least one empty place – an argument place. And these argument places must always be preserved in a function name and be recognizable as such; otherwise the function name becomes a meaningless proper name.[26]

Empty places are hard to see, especially when they come at either end. Frege uses different devices to make empty places visible. One is a pair of brackets serving as an empty container: '( ) = ( )'. Another is the use of Greek letters: '$\xi = \zeta$', or the extreme case 'ξ', where the letter indicates the bare empty place that is his function-name for the identity function mapping any object to itself.[27]

He employs the Greek letters not just in the expressions he uses to mention functions, e.g. 'the function $\xi = \zeta$', but also in those he places between quotation marks to mention function-names, e.g. 'the function-name '$\xi = \zeta$''. This fact might lead one to think that he regards function-names as schemata. Thus Potts: 'A Fregean incomplete expression or function name is the same as a schema; it always contains at least one schematic symbol.'[28] But not so: Frege is clear that the Greek letters are not parts of his function-names.

where use is made of an expression like 'the function $\Phi(\xi)$', it is always to be observed that 'ξ' contributes to the designation of the function only so far as it renders recognisable the argument-places.[29]

when we say 'the function $1 + \xi - \xi$', the letter 'ξ' is not part of the function-sign … the role of the letter 'ξ' is to enable us to recognize the places where the supplementing proper name is to be put … This 'ξ' gives us a pointer for how to use the function-name.[30]

To repeat, the Greek letters do not here function schematically i.e. they are not parts of the function-name that are *replaced* by argument-terms when the function-name occurs in a larger context.

[25] Frege, letter to Peano, 29 September 1896, in *PMC*, pp. 112–18, at p. 116.

[26] Frege, letter to Russell, 13 November 1904, p. 161.

[27] Frege, *Gg*, vol. I, §26.

[28] Timothy Potts, *Structures and Categories for the Representation of Meaning* (Cambridge: Cambridge University Press, 1994), p. 71.

[29] Frege, *Gg*, vol. I, §1.

[30] Frege, 'Logic in Mathematics', in *PW*, pp. 203–50, at pp. 239–40.

They merely indicate empty places within the function-name. The empty places are *filled* when the function-name occurs in a larger context. (Note, too, that Frege's use of '$\xi = \zeta$' within quotation marks to mention an expression with empty places is another case of the flexibility of quotation with respect to reference.) So when he says that function-names are incomplete or unsaturated he cannot be talking in a roundabout, metaphorical way about the role of schematic letters within them, since they contain no such thing. Instead, his function-names are incomplete in the *literal* sense that they carry with them argument-places conceived of as 'empty places' or 'gaps'.

To sum up Frege's view, the schema '$\xi = \zeta$' is not a function-name occurring in 'Hesperus = Phosphorus'. It contains too much. The plain expression '=', on the other hand, contains too little: it omits empty places. A new kind of expression is needed: viz. '   =   ' with its two empty places. Frege thus supplies us with a *sixth* candidate for predicates.

2.2 *Frege according to other commentators*

Dummett's simple predicates are plain expressions. Although he acknowledges that they may be assigned 'gaps' or 'slots', he takes this to be a metaphor for a grammatical rule: 'the slot consists merely in the predicate's being subject to a certain rule about how it can be put together with a term to form a sentence'.[31] (As noted in §1.3, a discontinuous predicate may be regarded as a plain expression that consists solely of its discontinuous parts, i.e. it has no gaps built into it.) Since Frege's talk of empty places is not metaphorical, his predicates cannot be Dummett's simple ones.

As for complex predicates, Dummett acknowledges that they 'form the prototype for Frege's general notion of an "incomplete" expression'.[32] He notes that Frege says that they contain gaps, but again he wrongly makes such talk metaphorical, and now even more so. For he supposes that the incompleteness of an incomplete expression means that it is 'not really an expression – a bit of language – in its own right',[33] but is instead to be regarded as a property of expressions.

[31] Dummett, *Frege: Philosophy of Language*, pp. 32–3.
[32] *Ibid.*, p. 31.
[33] *Ibid.*

The same objection applies to Brandom's account,[34] since he identifies Frege's predicates with Dummett's complex ones.

A clue that something is wrong is that Dummett places all his emphasis on predicates with related argument-places. He believes that considering them makes Frege's notion of incompleteness 'immediately clear'.[35] Yet Frege himself does not single them out for special attention. On the contrary, he explains the idea of an incomplete expression by using predicates without related places (i.e. with either a single argument-place or several unrelated ones). For him, the incompleteness of an expression consists in its having at least one empty place. Incompleteness makes it a special kind of expression; it does not prevent it from being an expression. This holds even of predicates with related places. Frege needs only to indicate that their empty places are related, e.g. by using repeated occurrences of the same Greek letter as in '$\xi = \xi$' (contra Russell, who claims that Frege cannot indicate related places).[36]

When Geach himself started to regard function-names as linguistic functions, he was hesitant about attributing the same idea to Frege:

> So far as I know, Frege never explicitly adopts the view that linguistic functions are what symbolize numerical (or other) functions; but it seems likely that he would have adopted it if it had been put to him.[37]

Fifteen years later, however, he threw off any pretence to exegetical caution, and applauded Frege's 'fundamental insight that a concept is represented not by an expression within a sentence but by a function from e.g. proper names to sentences'.[38] Others who have claimed that Frege's function-names are linguistic functions include Hugly, Noonan, Rumfitt and Sullivan.[39] Stenius swithers between linguistic

[34] Robert B. Brandom, *Articulating Reasons: An Introduction to Inferentialism* (Cambridge, Mass.: Harvard University Press, 2000), pp. 131–2.

[35] Dummett, *Frege: Philosophy of Language*, p. 31.

[36] Bertrand Russell, *The Principles of Mathematics* (Cambridge: Cambridge University Press, 1903), Appendix A, §482.

[37] Peter Geach, 'Frege', in Elizabeth Anscombe and Peter Geach, *Three Philosophers* (Oxford: Blackwell, 1961), pp. 127–62, at p. 144.

[38] Peter Geach, 'Critical notice of Michael Dummett, *Frege: Philosophy of Language*', *Mind*, 85 (1976), pp. 436–49, at p. 440.

[39] Philip Hugly, 'Ineffability in Frege's Logic', *Philosophical Studies*, 24 (1973), pp. 227–44; Noonan, *Frege*, chs. 2 and 4; Ian Rumfitt, 'Frege's theory of predication:

functions and properties of sentences.[40] The same is true of the later Dummett. In *The Interpretation of Frege's Philosophy* he has Frege positing 'a congruence in logical type between the referents of expressions and the expressions themselves'.[41] Proper names *stand for* objects and *are* objects. In contrast, 'what stands for something incomplete, a function, is itself incomplete', e.g. a predicate 'may be viewed as a common property of certain sentences, or as a function whose values are those sentences'.[42] Here Dummett introduces linguistic functions as an alternative candidate for predicates, though he signally fails to decide between them and his earlier favourites, properties of sentences.

Identifying Frege's function-names with linguistic functions again wrongly imputes a metaphorical understanding of 'incomplete expression' and 'empty place'. It also strains credulity to suppose that he regarded function-names as functions, but felt no need to make his view explicit or to warn against possible confusion. He is insistent about the need to distinguish a function-name from the function named. If he regarded function-names as themselves functions, he would surely have told the reader that they are *linguistic* functions, not to be confused with the functions they name. He did nothing of the sort.

Frege's division of things into objects and functions is exhaustive as well as exclusive. Since function-names are not themselves functions, they must be objects. On his view of a function-name as a kind of expression, this is just what one would expect. Indeed, we shall see in §2.3 that he explicitly relies upon the objecthood of function-names in order to circumvent his difficulty in talking about functions in the material mode (the paradox of the concept *horse*).

Pace Dummett, then, Frege did not posit a congruence in logical type between the referents of function-names and the function-names themselves. Function-names *stand for* functions but *are*

An elaboration and defense, with some new applications', *Philosophical Review*, 103 (1994), pp. 599–637; and Peter M. Sullivan, 'The functional model of sentential complexity', *Journal of Philosophical Logic*, 21 (1992), pp. 91–108.

[40] Compare footnotes 2 and 6 of Erik Stenius, 'The sentence as a function of its constituents in Frege and in the Tractatus', *Acta Philosophica Fennica*, 28 (1976), pp. 71–84.

[41] Michael Dummett, *The Interpretation of Frege's Philosophy* (London: Duckworth, 1981), p. 485.

[42] *Ibid.*

objects. One must not be misled by his characterization of function-names and functions as both incomplete. 'Incomplete' means different things when it qualifies different things. At the linguistic level, function-names are incomplete in the sense that they have empty places, as opposed to complete proper names. But when Frege moves from linguistic items to their worldly referents, he contrasts incomplete functions with complete objects. According to this second, worldly sense of 'incomplete', function-names are as complete as proper names, since they are both objects. Or, to come at the matter from the other direction, functions cannot be said to be incomplete in the linguistic sense, since they are not expressions.

It is true that Frege once contrived to see an analogy between the incompleteness of function-names and the incompleteness of functions, by modelling the second on the first (*not* vice versa). He called 'the reference of a word part of the reference of the sentence, if the word itself is a part of the sentence',[43] and thus characterized functions, like function-names, as *incomplete parts* of a complete whole. For example, the function-name 'ξ conquered Gaul' is an incomplete part of the sentence 'Caesar conquered Gaul'. It is completed when the proper name 'Caesar' fills its empty place. By analogy, the function ξ *conquered Gaul* is an incomplete part of each of its values; it is completed when it is applied to one of its arguments ('the argument ... goes together with the function to make up a complete whole').[44]

This attempt to endow functions themselves with something like empty places is misconceived. It is utterly obscure how the values of the function ξ *conquered Gaul*, namely truth and falsehood, can have any kind of part, let alone the function itself as a part. Frege himself knew that he was on shaky ground in transferring the relation between the parts and the whole of a sentence to their corresponding referents.[45] But it was only late in his life that he presented a decisive objection. If a function goes together with an argument to form a whole – its value for the given argument – then the argument as well as the function must be a part of the value. Yet 'we cannot say that Sweden is a part of the capital of Sweden'.[46]

[43] Frege, 'On sense and reference', p. 165.
[44] Frege, 'Function and concept', p. 140
[45] Frege, 'On sense and reference', p. 165.
[46] Frege, '[Notes for Ludwig Darmstaedter]', in *PW*, pp. 253–7, at p. 255.

At this point it is necessary to counter Geach's story about how Frege came to the idea that predicates stand for functions. He tells us that

> Frege's first notion of a function was one that fitted only linguistic functions; but he later came to think that this view was insufficient – that functions belong to the subject-matter, not just the notation, of mathematics; his mind passed from linguistic functions, whose values and arguments are numerical expressions, to numerical functions, whose values and arguments are numbers; so also it was natural that he should pass from the recognition of the linguistic functions that occur in predication to the view that there are functions in reality which these predicational functions represent.[47]

This is fiction from start to finish. In support of its opening assertion, Geach cites the following definition from *Begriffsschrift*:

> Suppose that a simple or complex symbol occurs in one or more places in an expression … . If we imagine this symbol as replaceable by another (the same one each time) at one or more of its occurrences, then the part of the expression that shows itself invariant under such replacement is called the function; and the replaceable part, the argument of the function.[48]

But this definition actually contradicts his assertion that 'Frege's first notion of a function was one that fitted only linguistic functions'. It clearly applies 'function' to expressions, and, as Geach himself insists, a linguistic function is not an expression, even when its arguments and values are. Thus the 'functions' of this definition are not linguistic functions.

The story ends by supposing that Frege actually regarded predicates as linguistic functions. This is hard to reconcile with Geach's admission a few pages earlier that Frege 'never explicitly adopts the view'.[49] Worse, Frege explicitly contradicts it when he talks literally of expressions with empty places.

In conceiving of functions as expressions, the youthful Frege was following the mathematical custom of his day. By 1891, however, he had clarified his ideas, carefully distinguishing an expression for a

[47] Geach, 'Frege', p. 151.
[48] *Bs*, §9; Geach, 'Frege', p. 143.
[49] Geach, 'Frege', p. 144.

function (later 'function-name') from the function itself, and convicting others of confusing the two: 'a mistake, admittedly, that is very often met with in mathematical works, even those of celebrated authors'.[50] He never again applies 'function' to function-names.

2.3 *'On concept and object', footnote 8*

One other text has been cited in favour of construing Frege's predicates (or, more generally, function-names) as linguistic functions, viz. footnote 8 to 'On concept and object'.[51] In this paper, Frege engages for the first time with the paradox of the concept *horse*. He takes it for granted that the expression 'the concept *horse*' is (i) a proper name in good standing, i.e. it (ii) expresses a full sense and (iii) refers. Since (iv) a proper name refers to a single object, if anything, and (v) nothing is both an object and a concept, it follows that 'the concept *horse*' refers to an object, not to a concept. He now faces 'a quite peculiar obstacle … a certain inappropriateness of linguistic expression'.[52] Despite his best intention to mention a *concept* when using 'the concept *horse*', the expression itself refers to a proxy *object* that represents the concept. (Kerry, the critic to whom he is responding, had taken it to refer to something that is simultaneously an object and a concept, contrary to (v).) He also grants that 'ξ is not a concept' is (vi) a predicate, which is again in good standing, i.e. it (vii) expresses a full sense and (viii) refers. In particular, (ix) it refers to a concept that maps any object to truth, and therefore the sentence 'the concept *horse* is not a concept' is true. This is paradoxical because, prior to meeting Frege's semantic machinery, one would have intended the sentence to be false, and expected it to be so, by analogy with e.g. 'the volcano Vesuvius is not a volcano'. At this initial stage, Frege claims that (x) the paradox cannot be avoided: 'the obstacle is essential, and founded on the nature of our language',[53] i.e. there is no other way of mentioning or saying what we intend.

[50] Frege, 'Function and concept', p. 138. He was not given to self-criticism, but see the first note to *Gg*, vol. I, §1, and also footnote 40 to Philip E. B. Jourdain, 'Gottob Frege', a chapter from his 'The development of the theories of mathematical logic and the principles of mathematics' (1912), reprinted as the Appendix to *PMC*, pp. 179–206, which Frege commented on in manuscript.

[51] Frege, 'On concept and object', in *CP*, pp. 182–94, at p. 186.

[52] *Ibid.*, pp. 193–4.

[53] *Ibid.*, p. 194.

Each of (i)–(x) has been rejected by some critic or other (not least Frege's later self) as a way out of the paradox. Solving it is not my present concern, however, since I only need to set enough of the scene in order to understand the relevant footnote. It reads:

> A similar thing happens when we say as regards the sentence 'this rose is red': The grammatical predicate 'is red' belongs to the subject 'this rose'. Here the words 'The grammatical predicate "is red"' are not a grammatical predicate but a subject. By the very act of explicitly calling it a predicate, we deprive it of this property.[54]

According to Hugly's reading of the footnote,[55] Frege derives an analogous paradox concerning predicates from his view of them as functions. The proper name 'the grammatical predicate "is red"' fails to refer to a predicate, since a proper name cannot refer to a function, and so the sentence 'the grammatical predicate "is red" is not a predicate' is true. Geach hints at the same reading. He too thinks that Frege foresaw that his difficulty with the concept *horse* arises 'on the linguistic level too ... we see the futility of trying to escape Frege's difficulties by semantic ascent, by talking about words instead of the objects and concepts signified'.[56]

But this reading of the footnote cannot be right. As noted in §2.2, all the evidence is against construing Frege's predicates as functions. And as we shall see, he explicitly argues that his difficulties *can* be circumvented by moving to the linguistic level. Worse, to read the footnote as presupposing that predicates are functions makes a nonsense of the main text of 'On concept and object', in which he explicitly construes them as expressions (the same goes for the contemporaneous 'Function and concept', to which he points the reader in his concluding paragraph). He speaks of 'concept-words' interchangeably with 'predicates', and says e.g. that the *words* 'no other than Venus' stand for a concept. Or again, consider his account of why 'the concept *horse*' stands for an object, not a concept. If he had believed that it is linguistic functions which stand for concepts, he would have ruled out 'the concept *horse*' as standing for a concept simply on the ground that it is an expression, not a function. But he doesn't reason in this way. He denies that 'the concept *horse*' stands

[54] *Ibid.*, p. 186.
[55] Hugly, 'Ineffability in Frege's logic', §IXA.
[56] Geach, 'Names and identity', p. 149.

for a concept on the different ground that it is an expression *of the wrong kind*: 'this is in full accord with the criterion I gave – that the singular definite article always indicates an object'.[57]

Frege was soon to offer a solution to his difficulty in talking about concepts. He points out that 'the reference of the concept-word *A*' is as problematic as 'the concept Φ', since 'the definite article before "reference" points to an object'.[58] But he now thinks that the difficulty is avoidable, even in natural language:

> It would be better to confine ourselves to saying 'what the concept-word *A* stands for', for this at any rate is to be used predicatively: 'Jesus is what the concept-word "man" stands for' in the sense of 'Jesus is a man'.[59]

What is important here is what Frege does not say. He does *not* object to the definite article before 'concept-word', either in the illegitimate 'the reference of the concept-word *A*' or in its bona fide replacement 'what the concept-word *A* stands for'. But the definite article is only in order if concept-words are themselves objects, and not functions as Hugly and Geach contend.

Dummett's reading of the footnote is completely different from Hugly's and Geach's. For him, it contains an error, not an insight: 'Frege was quite wrong in pretending that the same ills affect the formal mode of speech'.[60] He supposes that the footnote is ambiguous. Frege may have intended assimilating his paradoxical 'the concept *horse* is not a concept' either to 'the predicate "is red" is not a predicate' (the more appropriate analogue) or to '"the predicate 'is red'" is not a predicate' (the sentence suggested by his actual words). But Dummett finds neither of them paradoxical: the first is straightforwardly false, the second straightforwardly true. However the ambiguity is resolved, then, the comparison between material and formal modes is void.

Dummett is right about the truth-values of the two sentences in the formal mode, but wrong to ascribe error and the offensive 'pretending'. Frege was surely not deluded. The unparadoxical truth-values follow immediately from his own view of the matter. And it goes against all we know of his writing – its strident and compelling honesty – to suppose that he was out to delude his readers.

57 Frege, 'On concept and object', p. 184.

58 Frege, 'Comments on sense and reference', p. 122.

59 *Ibid.*, p. 122.

60 Michael Dummett, 'Frege on functions' (1955), reprinted in his *Truth and Other Enigmas* (London: Duckworth, 1978), pp. 74–86, at p. 75.

These commentators have all missed the point of comparison intended in the footnote. They read Frege as meaning to assimilate *sentences* in material and formal modes, whereas he is actually assimilating the *proper names* 'the concept *horse*' and 'the grammatical predicate "is red"', while at the same time contrasting them both with 'the city of Berlin' and 'the volcano Vesuvius', a contrast that he mentions in the relevant passage of the main text. The assimilation and contrast do not turn on the failure or success of intended and expected reference. For while each of the second pair hits the target, it is not the case that each of the first pair misses. Unlike 'the concept *horse*', 'the grammatical predicate "is red"' does refer to what one intends and expects, namely a predicate. Hence Frege must be pointing to a different contrast, as follows. Whereas there is nothing peculiar about the make-up of 'the city of Berlin' and 'the volcano Vesuvius', it is quite different with 'the concept *horse*':

> The peculiarity of our case is indicated by Kerry himself, by means of the quotation-marks around 'horse'; I use italics to the same end. There was no reason to mark out the words 'Berlin' and 'Vesuvius' in a similar way.[61]

In fact, Kerry takes over Frege's own use of quotation marks in *Grundlagen* ('the concept "horse that draws the King's carriage"'),[62] while in the German original of 'On concept and object' Frege used expanded spacing between characters, as in 'the concept h o r s e'. Some such device is useful, since in the plain 'the concept horse', the doubling up of common nouns is hard to fathom, while 'the concept is a horse' is worse, since it is naturally read as a sentential clause. According to Frege's account, however, such a device is not just an aid to intelligibility; it also helps to create a context that shifts syntactic and semantic categories. He says that 'horse' is a predicate standing for a concept (in his informal writings, Frege commonly omits the copula and article, and sometimes fails to indicate empty places). But when the expression is italicized and given an appositional prefix as in 'the concept *horse*', the result is a proper name standing for an object. No such shift occurs between 'Berlin' and 'the city of Berlin', or between 'Vesuvius' and 'the volcano Vesuvius'. This is the contrast that Frege intends.

[61] Frege, 'On concept and object', p. 186.
[62] *Gl*, §46.

The function of Frege's footnote is now clear. Although he was concerned to emphasize the peculiarity of 'the concept *horse*', naturally he searched for more commonplace phrases of similar construction in order to rebut the charge of special pleading. The reference-shifting contexts of 'On sense and reference' were at the front of this mind. One of them – direct quotation of expressions – gave him the comparison that he desired. Just as italics help turn the predicate 'horse' into the proper name 'the concept *horse*', so quotation marks help turn the predicate 'is red' into the proper name 'the grammatical predicate "is red"'. The final sentence of the footnote – 'by the very act of calling it a predicate, we deprive it of this property' – is not intended to make the paradoxical claim that the predicate 'is red' is not a predicate. That would be to place the 'it' at the level of things mentioned, whereas Frege means to be describing the expression used. In other words, the 'it' is the expression 'is red', which normally functions as a predicate, but does not do so when it occurs within the context 'the grammatical predicate "is red"'.

In the footnote Frege is not at his lucid best. He does not take care to signal explicitly the comparison that he has in mind. Nor does he help by concluding it with an ambiguous sentence. But to read it as untypically sloppy is better than wrongly imputing error (Dummett), and better than ascribing a view of predicates as linguistic functions (Hugly and Geach) which goes against everything else Frege said, even in the very same paper.

This reading of the footnote is also consonant with Frege's later remarks in a letter to Russell. He once more faces the paradox of the concept *horse*, now in the more general form of the difficulty in saying of functions that they are functions: 'the nature of language forces us to make use of imprecise expressions … "*A* is a function" is such an expression: it is always imprecise; for "*A*" stands for a proper name'.[63] He points out that there is no such difficulty at the linguistic level:

Instead of using the imprecise expression 'ξ is a function', we can say: '"( ).3 + 4" is a function-name'.[64]

[63] Frege, letter to Russell, 29 June 1902, in *PMC*, pp. 135–7, at p. 136.
[64] *Ibid.*

But if '*A* is a function-name' is precise, function-names can be named by proper names that take the place of '*A*'. And this is exactly what he says of his example:

'"( ).3 + 4"' is a proper name, and its reference is the function name '( ).3 + 4'.[65]

If follows that Frege regards function-names as *objects*, since only objects can be named by proper names.

2.4 *Last thoughts*

In his last published work, 'Compound thoughts', Frege appears on a cursory inspection to deny that function-names are really unsaturated, and to deprive them of empty places. Was this an abrupt about-turn? It needs investigation. First, the context. His attention had turned to the intermediate realm of sense, in particular to *thoughts*, the senses of sentences. Although he had given up thinking of arguments and functions as parts of the corresponding values at the level of reference, he continued to transfer the relation of part to whole from sentences to the thoughts they express, and applied the idea that 'combination into a whole always comes about by the saturation of something unsaturated'[66] to these whole thoughts. In the paper in question, he investigates 'connectives', by which he means the kind of sense that joins several thoughts into one compound thought. Since the thoughts that are compounded are already saturated wholes, connectives must themselves be unsaturated in order to produce a saturated compound thought. He begins with the function-name 'and', which

seems doubly unsaturated: to saturate it we require both a sentence preceding and another following. And what corresponds to 'and' in the realm of sense must also be doubly unsaturated: inasmuch as it is saturated by thoughts, it combines them together. As a mere thing, of course, the group of letters 'and' is not more unsaturated than any other thing. It may be called unsaturated in respect of its employment as a symbol meant to express a sense, for here it can have the intended sense only when situated between

[65] *Ibid*.
[66] Frege 'Compound thoughts', in *CP*, pp. 390–406, at p. 390.

two sentences: its purpose as a symbol requires completion by a preceding and a succeeding sentence. It is really in the realm of sense that unsaturatedness is found, and it is transferred from there to the symbol.[67]

Appearances to the contrary, Frege is not in fact denying that the expression 'and' is unsaturated. He is only insisting that when it is so described it must be coupled with a sense. For if it is considered 'as a mere thing' it is divorced from any sense, and consequently there is nothing that could determine that it is subject to a grammatical rule of combination. But when it is understood as expressing a sense, the sense dictates that it 'requires completion by a preceding and a succeeding sentence'. (The distinction between thing and symbol also features in his celebrated discussion of identity;[68] the notion of an expression 'as a mere thing' is Dummett's attenuated notion of expression discussed in §1.3).

But what of empty places? Frege does not mention them in this passage. It might therefore be thought that unsaturatedness as it pertains to function-names no longer has to do with empty places, but is now merely a way of describing the grammatical rules that govern expressions. Turn the page, however, and empty places or 'gaps' appear repeatedly. For example, in the discussion of his fourth kind of compound thought, Frege says:

The connective is the sense of whatever occurs in 'A or B' apart from 'A' and B', that is, the sense of

'(or)'

where the gaps on both sides of 'or' indicates the twofold unsaturatedness in the connective.[69]

Frege continued equipping function-names with empty places to the end of his working life.[70]

2.5 *Must predicates have empty places?*

Frege's expressions with empty places are legitimate candidates for predicates. They are clearly distinct from the other five that

[67] *Ibid.*, p. 393.

[68] Frege, 'On sense and reference', pp. 157–8.

[69] Frege, 'Compound thoughts', p. 396.

[70] See Frege's unpublished 'Sources of knowledge of mathematics and the mathematical natural sciences', in *PW*, pp. 267–74, at p. 272, and the related letter to Hönigswald, 26 April 1925, in *PMC*, pp. 54–6, at p. 55.

I have discussed, viz. plain expressions, schemata, linguistic functions, properties of sentences, and properties of (or relations between) terms. But he does not claim that his candidate is merely one among many. On the contrary, he repeatedly implies that his way of regarding them is mandatory. Since he simply takes it for granted that function-names are expressions of some kind or other, he fails to argue against candidates that fail to be any kind of expression. He does argue, however, against regarding function-names as expressions of a kind different from his own. Against schemata, he says:

> when we say 'the function 1 + ξ – ξ', the letter 'ξ' is not part of the function-sign; for the proper name '1 + 3 – 3' is composed of the function-name and the proper name '3', and the letter 'ξ' does not occur in it at all.[71]

He is assuming that a function-name occurs in a more complex expression through being a part of it. Thus the function-name cannot contain a Greek letter, since the letter itself does not figure in the more complex expression in which the function-name occurs. Here he makes the same mistake as Geach (see §1.2a). He fails to allow for the different sense of 'occur', according to which the schema '1 + ξ – ξ' can quite properly be said to occur in '1 + 3 – 3' even if it is not a part of it.

In a single note Frege takes his style of argument one step further, distinguishing a function-name as it occurs on its own from a function-name as it occurs in combination. He says that 'sin ( )' (with empty brackets merely indicating the empty place) is 'meant only for the exceptional case where we want to symbolize a function is isolation. In 'sin 2', 'sin' by itself already symbolizes the function'.[72] In other words, a function-name in isolation is an expression with an empty place, but in combination it reduces to a plain expression. Evidently, he reasons that the function-name 'sin ( )' is not a part of 'sin 2' on the ground that the latter does not feature an *empty* place. He duplicates function-names, then, by using a narrow notion of part as it relates to complex expressions. Yet the notion is in fact quite elastic. There is a perfectly

[71] Frege, 'Logic in mathematics', pp. 203–50, at p. 239.
[72] Frege, 'What is a function?', in *CP*, pp. 285–92, at p. 291, fn. 3.

good sense of 'part' in which an expression with an empty place is a part of a more complex expression even though its place is then filled (think of the parts of a completed jigsaw puzzle). Indeed, Frege himself relied on this notion of part in his other explanations of the make-up of complex expressions. As to construction, he says 'If we call the parts of the sentence that show gaps unsaturated and the other parts complete, then we can think of a sentence as arising from saturating an unsaturated part with a complete part'.[73] As to decomposition, he says that a sentence 'can be imagined to be split up into two parts; one complete in itself, and the other in need of supplementation, or "unsaturated" … it contains an empty place'.[74]

In fact, for the purposes of describing the construction and decomposition of sentences, it is quite unnecessary to characterize function-names as parts of those sentences at all. Rather than describing '*Fa*' as formed by *filling* the empty place in '*F*', we can say instead that it is formed by *replacing* 'ξ' in the schema '*F*ξ', even though the schema is not part of the result. Likewise for decomposition, but in reverse. On this score, then, there is nothing to choose between schemata and Frege's own expressions with empty places as candidates for function-names. The same goes for plain expressions. The sentence '*Fa*' can equally well be described as formed by *attaching* the plain '*F*' to the term '*a*', while '*F*' itself is the result of the converse operation. In each of these three accounts, the function-name itself is a raw material for the construction, and a product of the decomposition. If function-names are regarded as things other than expressions, any of these accounts may still be given, but one will now use 'auxiliary expressions' (Noonan's phrase from §1.3) as the raw materials and products, rather than the function-names themselves.

Did Frege have any other reason for equipping function-names with empty places, rather than regarding them as plain expressions? He claims that without an empty place 'the function-name becomes a meaningless proper name'.[75] But building in an empty place is *not necessary* to make an expression a function-name and

[73] Frege, 'A brief survey of my logical doctrines', in *PW*, pp. 197–202, at p. 201.
[74] Frege, 'Function and concept', p. 146.
[75] Frege, letter to Russell, 13 November 1904, p. 161.

to prevent it from becoming a proper name. He is wrong to think that a proper name and function-name must 'differ essentially',[76] if by this he means that they are intrinsically different types of thing. Both can be regarded in the same way as plain expressions, since it is enough that they differ in their relational properties. Their grammatical behaviour is different, they express different types of sense, and they stand for different types of referent. In passing, it is worth noting that Frege's empty places are *not sufficient* by themselves to mark out an expression as a function-name. At the end of §2.3, we saw that the function-name '( ).3 + 4' is named by the proper name '"( ).3 + 4"'. His brief remarks on quotation[77] suggest that enclosure by quotation marks creates a context in which the enclosed expression shifts reference. It follows that the expression with an empty place

( ).3 + 4

does double duty. Outside of quotation marks, it serves as a function-name. But when placed between them, it serves as a proper name, in which case its empty place alone cannot mark it out as a function-name. His thesis that nothing can be both a proper name and a function-name must therefore be qualified by excluding such shifts in context.

Frege's talk of completion or saturation of incomplete expressions by complete ones indicates that empty places are supposed to play another role, namely in explaining how a sentence differs from a mere list. For example, a string of proper names is not a sentence. The proper names 'hold aloof from one another … however we put them together, we get no sentence',[78] whereas

> Concept words and proper names are exactly fitted for one another. Because of its need for completion (unsaturatedness, predicative nature), a concept word is unsaturated, i.e. it contains a gap which is intended to receive a proper name. Through such saturation or completion there arises a sentence.[79]

[76] *Ibid.*
[77] Frege, 'On sense and reference', pp. 159, 165.
[78] Frege, 'On concept and object', p. 193.
[79] Frege, letter to Hönigswald, 26 April 1925, p. 55, with 'sentence' replacing the translator's 'proposition'.

There are *two* contrasts in this area that may be thought to demand explanation. The first is the contrast between the one sentence and the many items in a string of proper names. The sentence is made of many items, but what makes them into one thing, while the string remains merely many? One answer is saturation.

It is disputable whether a string of proper names *is* its many items, for it may well be regarded instead as a single expression *made* from its many items. After all, it is *a* string. If so, saturation is not necessary to explain how one thing is made from many, since the string is such a thing, but does not feature saturation. Concatenation is enough.

If one resists treating such a string as a single expression, consider instead a word. It is one thing made from many letters, but even Frege did not suppose that some letters are saturated by others (which ones?). Or again, think of a single, semantically simple proper name consisting of several words, such as 'New York' or 'The Big Apple'.

The second contrast is independent of the first. Even if we regard a string of proper names as a single expression, it still does not count as a sentence. So what makes one expression a sentence and another not? Frege invokes saturation to explain the difference (it has to be saturation of a particular kind, since saturation of other kinds features in expressions other than sentences, e.g. '2 + 3'). But saturation is again not necessary for the job. If an account is demanded why a particular expression is or is not a sentence, general rules of grammatical combination may be given from which an answer can be deduced (at least in the case of a formal language). Contra Frege, there is no need to think that the different rules governing function-names and proper names have to be reflected in different intrinsic properties, the one kind having empty places, the other not. Nor can this kind of intrinsic difference be made to take over the work of grammatical rules, since ungrammatical combinations can easily be written down, e.g. '$\Phi(\ ) = X(\ )$' violates the rule that only proper names may flank '=' (*pace* Stenius, who thinks that Frege's account of function-names shows the "superfluousness" of a theory of types', i.e. he supposes that grammatically impermissible combinations are impossible).[80]

[80] Stenius, 'The sentence …', p. 79.

Frege's apparatus of empty places does indeed make the failure of grammaticality visually apparent: the empty places are plainly not filled. This is hardly an advantage of his notation, however, since in e.g. standard presentations of the predicate calculus one can effectively decide whether a particular predicate occurs in a well-formed combination on the basis of its font, case or other typographical features, together with similar features of the other expressions with which it is combined, and general rules specifying permissible combinations. It is not just an effectively decidable matter, it is easy to determine. No empty places are needed; we have managed well without them.

A final argument moves from the distinctive nature of functions to the need for empty places in the corresponding function-names:

> what is distinctive of a function, as compared with an object, is precisely its 'unsaturatedness', its needing to be completed by an argument; and this feature must also come out in the symbolism.[81]

I can accept that functions need to be completed by arguments if this means that it is in the nature of functions to be applied to arguments (though not in any sense having to do with parts and wholes). As for the corresponding function-names, one can say that they need to be completed by argument-terms, in the sense that their grammar dictates that function-names can only properly occur in larger contexts when so combined. They are governed by this kind of grammatical rule, since they are meant to stand for functions that need completion. In this sense, the nature of functions does come out in the symbolism. But it does not follow that function-names need to be equipped with empty places. It is enough that they have the grammar they do; they can look like anything you please. Again, the conclusion is the same: there is nothing to choose between plain expressions and expressions with empty places. Although I have concentrated in the previous few paragraphs on the alleged competition between these two rival candidates, what I have said can easily be adapted to show that any of the other candidates for function-names can do as well as Frege's own (as before, for some

[81] *Gg*, vol. II, selections in *Translations from the Philosophical Writings of Gottlob Frege*, 3rd edn, ed. Peter Geach and Max Black (Oxford: Blackwell, 1980), pp. 139–224, §147, note.

purposes one will need to invoke auxiliary expressions rather than the chosen candidates for function-names themselves).

Frege was the original arch-prescriptivist about logic. He claimed that natural language is replete with faults that are to be corrected in his ideal, symbolic language. For example, proper names *should* stand for one thing and one thing only (i.e. they should neither be empty nor plural); function-names *should* be neither vague nor incompletely defined; sentences *should* be neither devoid of truth-value nor be anything other than true or false. He voiced these prescriptions in the strongest terms, and held onto them with remarkable obstinacy. It went entirely against his grain to countenance alternatives. Although his arguments for his prescriptions are quite feeble, it has taken time for us to see his logical system for what it is: a brilliant development of one alternative among many. The paradigmatic status of the predicate calculus and its second-order extension shows that some of his prescriptions still retain their grip.

His view of the nature of function-names has some of the same characteristics: the strength of expression, the obstinate persistence, the criticism of natural language (now as 'covering up' the distinction between proper names and function-names), feeble supporting arguments. There is a striking difference, however. Unlike his prescriptions for logic, his candidate for function-names was never widely adopted. Indeed, as I have shown, his conception of them has been widely misunderstood. Different exegetes pin different, alien conceptions on him. Like him, they often argue that their chosen candidate for function-names is the right one, but in reality any will do. There is no competition and no uniquely right answer.

THOMAS RICKETTS

6 Concepts, objects and the Context Principle

In 1906, after giving up hope of vindicating logicism, Frege lists the results of his life's work that survive Russell's paradox. He begins:

> Almost everything is connected with the Begriffsschrift. Concepts conceived as functions. Relations as functions with two arguments. Concept-extensions or classes are not primary for me. Unsaturatedness of both concepts and functions. The essence of concepts and functions recognized.[1]

Twenty-two years earlier, in happier times, Frege took for his third guiding principle in *Die Grundlagen der Arithmetik* the admonition, 'never to lose sight of the distinction between concept and object'.[2] We cannot understand Frege's view of his epochal achievements in logic – let alone mark the distance that separates his understanding of quantificational logic from contemporary views – without an appreciation of the concept–object distinction, the centrality Frege assigns to it and its connections to his other views. The distinction embodies the quantificational understanding of generality that Frege sets against older conceptions of logic. This quantificational understanding of generality gives Frege the principle for determining the logical segmentation of sentences and the contents or thoughts expressed by sentences. I hold that Frege's Context Principle sets

[1] 'Was kann ich als Ergebnis meiner Arbeit ansehen?' in *NS*, p. 200 (*PW*, p. 184). (Parenthetical references following page references to Frege's *Nachgelassene Schriften* are to the English translation in *Posthumous Writings*. The translations of passages from Frege's writings are my own, made consulting the common English editions.)

[2] *Gl*, Introduction, p. x.

forth this connection between logical segmentation and quantificational generality.

Sections 1–3 lay out Frege's concept–object distinction and its place in his philosophy of logic. Section 4 describes Frege's introduction of higher-level concepts and his assimilation of sentences to proper names. Sections 5 and 6 explore the charge that the concept–object distinction incoherently undermines itself in the so-called Kerry paradox. Sections 7 and 8 discuss the bearing of the Context Principle and the concept–object distinction on Frege's logicism.

1

Frege takes as a given our capacity for objective knowledge, our capacity to recognize objective truths. There are two intertwined aspects to Frege's conception of objectivity. First, the truth or falsity of the contents we judge is independent of anyone's cognition, independent of anyone's grasping or judging those contents. Second, several individuals may judge the same contents true or false.[3] This conception of objectivity is built into Frege's conception of judgement: to make a judgement is to recognize the objective truth of an intersubjectively available content.[4] This capacity for knowledge includes a capacity for logical inference whose exercise enables us to recognize one truth on the basis of others. Frege aims to codify principles for logical inference in such a way that their application in any stated proof will force the explicit statement in the proof of any premise on which any conclusion or subconclusion depends. This enterprise is premised on the conviction that the inferential

[3] However, at the end of his career, in 'Thoughts', Frege qualifies intersubjectivity, positing in connection with the first-person pronoun thoughts that only a single individual can grasp.

[4] Frege thus urges in 'Logik' (1897), *NS*, p. 144 (132), that it is, as we might put it, pragmatically incoherent to deny the objectivity of what we know, on the grounds that this very denial would itself put forward as objectively true an intersubjectively available content. I see Frege voicing a similar attitude toward objectivity in 'Thoughts', p. 74. (Page references to Frege's published papers are to the pages of the original publication. These are marginally indicated in the leading German and English editions of Frege's papers.)

leaps in colloquial science and mathematics can be analyzed into a series of simple inferential steps.

These simple modes of inference must be applicable in proofs across the sciences. They must then abstract from the content that distinguishes the various sciences. Frege conceives this abstraction substantively. Logical laws are maximally general truths – generalizations whose statement requires only that topic-universal vocabulary required to express the results of any science, e.g. an expression for negation. Logical laws are then on the same level as the laws of the various special sciences.[5] The relation of logic to other sciences is that of a more abstract, less detailed science to a more detailed one. As Euclidean geometry is a body of knowledge that classical physics assumes and draws on, so logic is a maximally general science that every science implicitly assumes and draws on. On Frege's approach, then, fundamental laws of the maximally general science of logic capture topic-universal modes of inference. In this way, the capacity for inference proves to be the capacity to recognize one truth on the basis of another in accordance with logical laws, as Frege puts it.[6]

On Frege's approach to logic, the strategically central mode of inference is the inference from general to specific, from a generalization

[5] Frege's logical laws are not about a logical consequence relation. Nor is there any indication that he is guided by an intuitive semantic conception of consequence in framing them. I disagree, then, with Dummett's claim that Frege's formulation of logic rests on a semantic characterization of a logical consequence relation extracted from a semantic analysis of language. See Michael Dummett, *Frege: Philosophy of Language* (London: Duckworth, 1973), pp. 81–2. For further discussion, see Warren Goldfarb, chapter 3, this volume. I discuss how Frege conceives the task of codifying principles for logical inference in 'Frege's 1906 foray into metalogic', *Philosophical Topics*, 25 (1997), pp. 169–88, at §§1–2.

[6] For Frege's characterizations of inference, see 'On the foundations of geometry' (1906), part 2, p. 387, and his letter to Dingler, 31 January 1917, in *PMC*, pp. 16–17. Compare Frege's talk of inference modes (*Schlußweisen*) in *Gl*, §90. To be more accurate here, logical inference modes are captured in Frege's formulation of logic jointly by his logical axioms and inference rules. Although in his 1893 *Grundgesetze* formulation of logic Frege erects redundant inference rules to shorten derivations, he prefers to capture inference modes by logical axioms, restricting inference rules to the quantifier rules like Relettering that cannot be formulated by any single higher-order generalization. See *Gg*, p. vi, and 'Booles rechnende Logik', *NS*, pp. 43–4 (38–9). He does recognize the need in his formalism for, as we would put it, a truth-functional inference rule; he uses versions of *modus ponens* to this end in his formulations of logic.

to its instances.[7] Frege for the most part formulates his principles of inference as generalizations. These then get applied in proofs via a Substitution rule. The codification of laws of logical inference encounters immediate problems, however. Colloquial language is variously ambiguous, irregular, redundant and limited in its expression of logically relevant relationships.[8] Frege thus devises his Begriffsschrift to give unambiguous, perspicuous and more uniform expression to logically relevant relationships. His first task in devising a Begriffsschrift is to construct a notation that gives unambiguous, perspicuous expression to generality in order to make the relation of generalization to instance notationally recognizable.

By reflecting on arithmetical notation, Frege hits on the basic pattern for his Begriffsschrift. The central feature of mathematical notation that Frege latches on to is its use of letters to express generality, its use of letters as variables.[9] This use of letters as variables presupposes within the equations of arithmetic the isolation of numerical terms – numerals and the complex arithmetical expressions formed from them by iterated use of signs for the arithmetical operations – that are replaceable by variables to form generalizations of which the original equations are instances. Thus, we generalize

$$[2 \times (4 + 1)] + (4 + 1) = 3 \times (4 + 1)$$

to obtain

> The sum of a number with its double is its triple,

or

> Two times a number plus that number = three times that number,

[7] Frege emphasizes the centrality of the inference from generalization to instance in 'Logical Generality,' *NS*, p. 278 (258).

[8] Frege underscores the unsuitability of colloquial language as a medium for the codification of logical principles for the purposes of non-enthymematic proof in *Bs*, foreword, pp. iv–v, and *Gl*, §91.

[9] Indeed, the 1879 monograph is subtitled 'A formula language [*Formelsprache*] for pure thought modelled on the formula language of arithmetic'. On p. iv of the foreword to *Bs*, Frege identifies the use of letters as variables as the most direct way in which his notation is modelled on arithmetical notation; the first section of *Bs* presents the distinction between names and variables as a fundamental feature (*Grundgedanke*) of his approach.

or

$$2y + y = 3y.$$

We may then substitute any numerical term, simple or complex, for '*y*' to obtain further instances of this generalization. Via his notion of a proper name, Frege discerns the structure so conspicuous in arithmetical notation throughout language.

Frege's entire approach here makes a sharp break with traditional logic that is of the highest importance.[10] Traditional logic (syllogistic logic, the logic of categorical judgements) does not recognize as a distinctive mode of inference the inference from generalization to instance. Indeed, traditional logic lacks a *quantificational* conception of generality. The sentences

(A) Every human is mortal

and

(B) Socrates is mortal

are grammatically parallel. Both consist of the predicate 'mortal' joined by the copula to a subject. Traditional logic classifies the judgements expressed by these sentences together as universal affirmative categoricals. The inference to (B) from (A) and

Socrates is human

is thus just the syllogistic figure BARBARA.

Frege blames the limitations of traditional logic on the prominence in colloquial language of the grammatical distinction between subject and predicate. Traditional logic, blinded by the subject–predicate distinction, assigns a privileged position to categorical judgements, and consequently misrepresents the inference from general to specific as a matter of concept-inclusion. It is this misrepresentation that Frege laments, when he repeatedly chides his contemporaries for running together the distinction between an

[10] When I speak in this paper of 'traditional logic', I mean traditional logic as Frege viewed it. I believe that Frege pretty much identified traditional logic with syllogistic logic viewed through a Boolean lens. My concern is to understand Frege's viewpoint, and the rhetoric emanating from that viewpoint. Michael Potter also links the break Frege makes with traditional logic with his use of variables. See M. Potter, *Reason's Nearest Kin: Philosophies of Arithmetic from Kant to Carnap* (Oxford: Oxford University Press, 2000), p. 64 and pp. 33–4.

object's falling under a concept (expressed by B) and one concept's being subordinate to another (expressed by A).[11]

Moreover, in Frege's eyes, the salience of the subject–predicate distinction and the traditional logic erected on it also foster an inadequate view of judgement and inference. Frege sketches this view in 'Booles rechnende Logik':

> I have distanced myself further from Aristotelian logic than Boole has. For Aristotle namely, as for Boole, the logically basic activity [*logische Urtätigkeit*] is the formation of concepts by abstraction, and judgement and inference proceed by means of the unmediated or mediated comparison of concepts according to their extensions.[12]

On this view, the terms joined by copulae into expressions of categorical judgements are associated with mental representations, concepts in the subjective sense as Frege would say.[13] The formation of these mental representations is grounded in abstraction from experience so that the concepts represent the objects of experience. The difference between the mental representation associated with 'Socrates' and that associated with 'human' is that the former represents a single thing, but the latter represents a large number of things. Proper names and predicates are then both names of things: proper names designate just one thing; predicates typically designate several things. The items represented by a concept comprise that concept's extension. Judgement and inference centrally concern inclusions and overlaps among the extensions of concepts.[14]

Frege's approach to logic is very different from the approach he finds in traditional logic. In summarizing his life's work at the end of his career, Frege says:

[11] See 'Booles rechnende Logik, *NS*, p. 20 (18), where Frege notes that the subject–predicate structure of colloquial language obscures this distinction. See also Frege to Marty, 29 August 1882, *WB*, p. 165; Frege to Peano (undated), *WB*, p. 177; 'On concept and object', p. 201ff., 'A critical elucidation of some points in E. Schröder, *Lectures on the Algebra of Logic*', pp. 441–2; 'Logik in der Mathematik', *NS*, pp. 230–1 (213). I think that Frege views Boole and Schröder (but not Peano) as attempting to generalize the theory of concept-extensions contained in syllogistic logic. As they both lack a quantificational conception of generality, their work inherits what for Frege is the fundamental limitation of traditional logic.

[12] 'Booles rechnende Logik', *NS*, p. 16 (15).

[13] See *Gl*, §27, p. 37.

[14] Frege recognizes that some of his contemporaries, like Boole, add to categorical logic a propositional logic to treat the combination of categoricals into compound statements.

What is distinctive about my conception of logic comes out first in that I give top priority to the content of the word 'true' and then that I immediately introduce thoughts as that concerning which the question of truth arises. I therefore do not begin with concepts that I put together into thoughts or judgements. Rather, I obtain thought-components [*Gedankenteile*] by analysing [*Zerfällung*] thoughts.[15]

It is generality that compels the analysis of a thought into parts none of which is a thought, and sentences expressing thoughts into parts none of which is a sentence expressing a thought.[16] The simplest example of such analysis is the division of the thought expressed by sentences like

Socrates is mortal

into a proper name, 'Socrates', and the part that remains when this proper name is removed,

__is mortal.

The proper name in the sentence corresponds to a saturated, whole, complete part of the thought the sentence expresses; the leftover part to an unsaturated part that requires completion.

Thus analysed, our sentence says (expresses the thought) that a particular individual, Socrates, is mortal. By replacing 'Socrates' with other proper names, we can get sentences that say that various other things are mortal. In contrast to the contents expressed by these sentences, the content expressed by

Everything is mortal

or equivalently

x is mortal

[15] 'Aufzeichnungen für Ludwig Darmstaedter', *NS*, p. 273 (253). Compare 'Booles rechnende Logik', *NS*, p. 17 (16), where Frege says, 'In opposition to Boole, I begin with judgements and their contents instead of concepts … For me the formation of concepts arises only from judgements.' See also 'On the purpose of the *Begriffsschrift*', p. 4ff.

[16] See 'Aufzeichnungen für Ludwig Darmstaedter', *NS*, p. 274 (254); 'Einleitung in die Logik', *NS*, p. 203 (187); 'Kurze Übersicht meiner logischen Lehren', *NS*, p. 217 (201); and '17 Kernsätze zur Logik', *NS*, p. 189 (174).

does not say that any individual is mortal. This sentence is not about any particular thing: the content it expresses is general, and general *where* the contents expressed by our other sentences are specific. By replacing a proper name in a sentence with a variable, we confer generality of content on the original sentence.[17] While each of the proper names in our original family of sentences means a determinate individual thing, the variable indefinitely indicates individual things.[18]

Frege's understanding of the role of proper names to signify objects in this way is part and parcel of his quantificational understanding of generality. He does not take the notion of an object and of a name's signifying an object to be an independently available basis for introducing quantificational generality. In particular, the notion of a name's signifying an object in a sentence is not prior to that of a variable's indefinitely indicating an object. Frege takes as basic the inference from generalization to instance, the inferential relationship between a generalization and its instances. This is evident from the procedure Frege describes: he begins with thoughts, and it is generality – the inference from generalization to instance[19] – that prompts the recognition of parts of thoughts that are not themselves thoughts. In speaking of proper names as designations of objects and of sentences containing the names, in contrast to the corresponding generalizations, being about the named objects, Frege seeks to awaken an explicit awareness of this distinctive inference-mode: to analyse 'Socrates is mortal' as the completion of '__ is mortal' by the proper name 'Socrates' is to grasp the content expressed by this sentence as an instance of the corresponding generalization expressed by '*x* is mortal'.[20]

[17] 'Einleitung in die Logik', *NS*, p. 204 (188) and p. 206ff. (190); 'Kurze Übersicht meiner logischen Lehren', *NS*, p. 215 (199); and 'Foundations of Geometry' (1906), p. 307. See also 'Begründung meiner strengeren Grundsätze des Definierens', *NS*, p. 166ff. (154).

[18] In addition to the references in note 14, see *Bs*, §1; *Gg*, §17, p. 31ff.; 'What is a function?', p. 659ff.

[19] In 'Logical generality', *NS*, p. 278 (258), Frege says, 'The person who knows how this inference goes [the inference from general to specific] has also grasped what generality is (in the sense of the word here intended).'

[20] See Frege's remarks on the advantages of using variables to express generality in 'Logical Generality', *NS*, p. 280 (260).

Quantificational generality and identity are intertwined for Frege. Quantificational generality is generality over a multiplicity of discrete, determinate objects, determinate quanta so to speak. Discreteness is thus built into Frege's conception of an object: no entity without identity.[21] Frege, in his post-1891 elucidations of identity, says that the identity of objects *a* and *b* is the complete coincidence (*zusammenfallen*) of *a* and *b*.[22] If *a* and *b* completely coincide, then there is no difference between them so that whatever holds of *a* also holds of *b*, and vice versa. Here is the justification for the linguistic substitutions that encode the Leibniz inference: in a true sentence, replacement of a proper name by a proper name signifying the same object yields a truth. A grasp of the inference from generalization to instance in connection with proper names thus includes a grasp of objects as discrete, and so a grasp of identity and the Leibniz inference. These two inference-modes come together. It is not the inference from generalization to instance alone, but the interlock of this inference and the Leibniz inference that isolates proper names in sentences.[23]

I noted how Frege holds that traditional approaches to logic assimilate proper names and predicates. As a part of this assimilation, traditional logic also does not distinguish the copula from the expression of identity, symbolized in arithmetic by the identity-sign '=', and so misrepresents the Leibniz inference that equations

[21] W. V. Quine states this precept in 'Speaking of objects', in *Ontological Relativity and Other Essays* (New York: Columbia University Press, 1966), p. 23. See also the discussion on p. 19 and in 'On what there is', in *From a Logical Point of View* (Cambridge, Mass.: Harvard University Press, 1953), p. 4. In §3, in considering Frege's view of concepts, we shall see that Frege's understanding of this precept diverges from Quine's.

[22] See 'On sense and reference', p. 26, fn. 1, 'On concept and object', p. 194, fn. 2, and especially, 'Husserl review', p. 320.

[23] Given the irregularity of everyday language, the multiple expressions it offers for the same thoughts and especially the possibilities of forming grammatical subjects for grammatically singular sentences by nominalizations of adjectives, verbs and sentences, there may be uncertainty in some cases concerning the recognition of proper names. Nevertheless, as Frege views matters, our capacity for logical inference enables us confidently to distinguish proper names over a wide variety of sentences on many topics in colloquial language. This is all that is necessary in order to arrive at a pattern for the uniform, perspicuous expression of generality in a Begriffsschrift. I return to this position in connection with Frege's views on the application of logic in §8.

support.[24] Frege insists on a distinction here. The disambiguation of uses of 'is' in colloquial language as a copula and to express identity is the fulcrum that leverages Frege's rejection in logic of the subject–predicate distinction. In making this distinction, Frege calls attention both to object-signifying proper names – these are the terms of equations[25] – and with them to the relationship of generalization to instance. He then can extend mathematicians' use of variables to express generalizations across the board, and in this way notationally capture the inference from generalization to instance.

2

Frege's first published exposition of the concept–object distinction comes in *Grundlagen*, §51, and is explicitly directed against the view that predicates are common names, names common to many things.[26] Frege maintains that predicates signify concepts, not the things falling under concepts. These things are signified by proper names. Frege elaborates this contrast in the opening pages of 'On Concept and object', where he introduces the concept–object distinction by arguing that proper names cannot be used as predicates. To think they can, Frege maintains, is to confuse the use of 'is' to express identity in

(A) The Morning Star is Venus

[24] Insofar as the Leibniz inference is representable in syllogistic logic, it too is assimilated to BARBARA.

[25] By 'equation', I always mean singular equation. This mark of proper names fits with, indeed unifies, the other typical grammatical characteristics of proper names that Frege mentions: occurrence as the grammatical subject of expressions of singular judgements (*Gl*, §66, p. 77, fn **, and 'On concept and object', p. 198); and the use of the definite article and demonstratives with singular general terms in definite descriptions (*Gl*, §51, p. 63 and 'On concept and object', pp. 195–6). I do not suppose that presenting proper names as the terms of equations is a syntactic explanation of them, as opposed to presenting them as designations of objects. On Frege's approach, the identification of equations depends on a grasp of their sense. Neither elucidation of proper names is prior; each illuminates the other.

[26] For a slightly earlier, parallel discussion of the distinction, see Frege to Marty, 29 August 1882, *WB*, p. 164. Frege touches on the distinction in 'On the scientific justification for a *Begriffsschrift*', p. 50. For further expression of Frege's opposition to the view of predicates as common names, see Frege to Husserl, 24 May 1891, *WB*, p. 96, and 'Schröder elucidations', p. 454.

with its use as a copula in

(B) The Morning Star is a planet.

Frege works to elicit from his audience a recognition of the difference between these two uses. The two proper names in (A) may be interchanged to yield 'Venus is the Morning Star', an equivalent equation whose grammatical subject is 'Venus'. In this way, then, the proper names that are the terms of an equation behave symmetrically. In contrast, there simply is no such thing as interchanging 'The Morning Star' and 'a planet' in 'The Morning Star is a planet', to obtain a sentence parallel to it, as our two equations are parallel – a sentence in which 'a planet' is the grammatical subject.

Recognizing the difference in the use of 'is' in (A) and (B), we are to see that nothing counts as using a proper name as a predicate. There is no sentence consisting of two proper names joined by the copula to express a thought parallel to the one that (B) expresses.[27] Frege has been speaking in the formal mode, about language. He casually shifts into the material mode to state his final conclusion that extends the dichotomy of proper names and predicates to what they signify, objects and concepts:

> We have here a word 'Venus' that can never be a genuine predicate, although it can form a part of a predicate. The reference of this word can therefore never appear as concept but only as object.[28]

Proper names signify objects. Similarly, predicates signify concepts. By observing that there is no such thing as a sentence in which a proper name is used as a predicate, Frege attempts to instill in his audience an appreciation that these roles are distinct and mutually exclusive.

We shall return to this point shortly. First, however, we need to consider Frege's view of predicates as designations or names of concepts.

For Frege, a predicate is the result of removing one or more occurrences of a proper name from a sentence. The predicate itself contains a blank, an argument position, where the proper name is removed. When the proper name removed to obtain a predicate

[27] I discuss to this point further below.
[28] 'On concept and object', pp. 194ff.

is the grammatical subject of a simple sentence, the predicate approximates a grammatical predicate, which Frege without special comment assimilates to his predicates.[29] However, the proper name removed to obtain a predicate from a simple sentence need not be the grammatical subject; and several occurrences of a proper name may be removed from a compound sentence to form a predicate. Finally, a predicate isolated by the removal of a proper name from a sentence may appear in other, longer sentences that are not analysable as completions of it. For example, the predicate '__ is human' occurs in the conditional 'If Plato is human, then Plato is mortal'.

To understand Frege's view of predicates as names, let us consider his analysis of universal affirmative categoricals. The content expressed by

Every human is mortal

is false, if there is a counterexample to it, if there is something that is human while failing to be mortal. If, for example,

Plato is human and Plato is not mortal

were true, our categorical would be false. The categorical can then be analysed as the generalization of the negation of this sentence. So Frege introduces his material conditional

If *p* then *q*

as the denial that *p* and not *q*.[30] Our universal categorical is thus the generalization of

If Plato is human, then Plato is mortal

[29] So although Frege frequently presents general terms and adjectives as examples of predicates, his own view is that the copula is a part of the predicate, provided that we don't take the finite conjugation of the copula to express asserting force. See 'Über Schoenflies', *NS*, p. 192 (178). See also 'Foundations of geometry' (1903), p. 371; and 'Logik in der Mathematik', *NS*, p. 247 (229).

[30] Throughout his career, Frege points to the use of his material conditional to express universal categoricals to justify his reading it as 'if … then'. See *Bs*, §12, p. 23; 'On the purpose of the *Begriffsschrift*', p. 6; 'Booles rechnende Logik', *NS*, p. 12 (11); *Gg*, vol. I, §13, p. 23ff.; 'Einleitung in die Logik,' *NS*, pp. 203–7 (187–91); 'Kurze Übersicht meiner logischen Lehren', *NS*, pp. 214–18 (199–202).

i.e.

> If something is human, then it is mortal

or

> If *x* is human, then *x* is mortal.

This sentence, we have seen, is not about any individual thing. It expresses rather a generalization over things. As Frege views matters, when we recognize in this generalization the predicates '__ is human' and '__ is mortal', we analyse the sentence to say something about the concepts signified by these two predicates: it says that the concept *human* is subordinate to the concept *mortal*, that everything that falls under the first concept falls under the second.[31]

For predicates to be names of concepts is for them to be replaceable by variables to express generalizations over concepts. Indeed, in the one place where Frege asks whether predicates are names, he answers the question in the affirmative by observing that sentences containing incomplete expressions are thereby instances of generalizations.[32] He accordingly introduces variables in the positions determined by incomplete expressions to express these generalizations.

Consider a colloquial expression of such a generalization:

> Plato is everything that Socrates is.

As Frege would analyse it, this generalization has such instances as

> If Socrates is mortal then Plato is mortal
> If Alcibiades loves Socrates then Alcibiades loves Plato.

[31] See 'On the purpose of the *Begriffsschrift*', p. 8ff.; *Foundations*, §47; 'Booles rechnende Logik', *NS*, p. 19ff. (18); 'Einleitung in die Logik', *NS*, p. 210 (193); 'Logik in der Mathematik', *NS*, p. 230ff. (213); 'Aufzeichnungen für Ludwig Darmstaedter', *NS*, p. 274 (254).

[32] 'Introduction to Logic,' *NS*, pp. 209–10 (192–3). Here Frege raises the question as to whether incomplete expressions have a reference in addition to a sense. He says: 'When we say "Jupiter is larger than Mars", what do we speak of? Of the heavenly bodies themselves, of the references of "Jupiter" and "Mars". We say that they stand in a certain relation to each other, and we do this with the words "is larger than".' He thus presents the thought the sentence expresses as an instance of an existential generalization over relations, and concludes that as 'Mars' and 'Jupiter' mean planets, 'is larger than' means a relation.

These sentences respectively contain two occurrences of the predicates '__ is mortal' and 'Alcibiades loves __'. Our generalization can be expressed by replacing the two occurrences with a variable:

If *F*(Socrates) then *F*(Plato).

The variable in a predicate position is accompanied by parentheses, and in the parentheses will be whatever fills the argument position in the occurrences of the predicate that replace the variable in instances of the generalization.

We are now ready to return to Frege's understanding of the difference between the role of proper names to signify objects and the role of predicates to signify concepts.

Proper names and predicates are alike names, each with their associated variables. To grasp the use of letters to confer generality of content on a sentence is at its core to know what names may be substituted for it to form instances of the generalization. As a result of the incompleteness of predicates, variables in proper-name positions may not be replaced by predicates to obtain expressions of instances of the generalization. If we attempt to place a predicate into a proper-name position, we do not get a sentence, let alone an instance of the generalization; for nothing fills the empty place in the predicate. Similarly, proper names do not fit into predicate positions. As noted, the variable in a predicate position is accompanied by parentheses that contain the expression that fills the empty place in predicates in instances of the generalization. If we attempt to substitute a proper name for a predicate-variable, there will be no place for the expression in the parentheses accompanying the variable. In this way, proper names and predicates signify differently: whereas proper names refer to objects, predicates refer to concepts. Frege thus recognizes two types, two levels of generality – the generality expressed by variables in proper-name positions over what proper names refer to (objects), and the generality expressed by variables in predicate positions over what predicates refer to (concepts).

Frege's first fundamental principle in *Grundlagen* counsels his readers sharply to separate the psychological from the logical, the subjective from the objective. It encapsulates Frege's view of objective knowledge. His third fundamental principle sets forth the concept–object distinction. These two are linked by the enigmatic

Context Principle, which says that we 'must inquire after the meaning of a word in the context of a sentence, not in isolation'.[33]

I urged in §1 that Frege takes as basic the inference from generalization to instance so that his view of logical segmentation is founded on his quantificational conception of generality. The Context Principle encapsulates the connection between names and quantificational generality: an expression is a meaningful (designating) name by occurring in true or false sentences that express instances of generalizations expressible by replacing the name with a variable. The inference from generalization to instance together with the Leibniz inference isolates proper names in sentences, segmenting sentences into a proper name and the leftover part, the predicate with its blank. These predicates are also names that may recur with the same significance in other sentences and may be replaced by variables to express corresponding generalizations. Thus, the recognition of proper names and generality over objects brings with it the recognition of predicates and generality over concepts. In this way, the view of names and quantificational generality epitomized in the Context Principle leads to the recognition of two levels of generality, and to the distinction between objects and concepts.

3

My description of Frege's multi-levelled conception of quantificational generality assigns a certain priority to proper names and objects vis-à-vis predicates and concepts. This priority stands at the heart of Frege's understanding of the concept–object distinction. We need to consider it more closely.

To begin, we should bear in mind that Frege's concepts are not mental representations based ultimately on abstraction from experience. Frege's concepts are no more mind-dependent than his objects; he explicitly places concepts and objects on a par as regards their objectivity.[34] Moreover, as Frege recognizes concepts under which

[33] *Gl*, p. x.

[34] See Frege to Husserl, 24 May 1891, *WB*, p. 96, where Frege says 'objects and concepts have the same objectivity'. See also *Gl*, §27, p. 37, fn. 1. Frege emphasizes the objectivity of concepts in *Gl*: see §26, especially the final paragraph, and §47.

nothing falls, he does not believe concepts to be dependent for their existence on the objects falling under them.[35]

However, in 'Foundations of geometry' (1903), Frege says:

> It is clear that we cannot put a concept forward [*hinstellen*] as self-subsistent, like an object. Rather a concept can only occur in a combination. It can be said that the concept can be distinguished in the combination, but cannot be separated out of the combination.[36]

This remark and others like it are obscure. What does Frege mean, when he says that concepts can only occur in a combination? After the sense–reference distinction, there are in Frege's ontology no complexes that contain concepts analogous to the way that sentences contain predicates. Moreover, the Context Principle applies to proper names and predicates alike; and Frege is explicit in *Grundlagen* that the self-subsistence he assigns to objects should not be taken to contradict it.[37]

Frege's 1903 remark alludes to the incompleteness of predicates. Fregean predicates are characterized in terms of proper names and sentences, for a predicate is the result of removing occurrences of a proper name from a sentence. Consequently, while proper names and sentences are morphophonemic units, identifiable as series of marks including the mark SPACE,[38] predicates with their empty places are not. There is no way to write down a blank. In contrast to proper names, the only way to write down a predicate is to write down a sentence in which the predicate occurs with its empty places filled. In this sense, predicates are unsaturated, in need of completion.[39] Proper names, having no empty places, are

[35] See Frege to Husserl, 24 May 1891, *WB*, p. 96. See also *Gl*, §74, p. 87, 'The Law of inertia', p. 158, and 'Schröder elucidations', pp. 453–4.

[36] 'Foundations of geometry' (1903), p. 372, fn. 5. Frege is responding in this footnote to Russell's view in chapter 4 of *Principles of Mathematics* that concepts can occur predicatively or non-predicatively (as subjects) in Russell's (non-linguistic, non-mental) propositions. Frege makes a similar remark in his 1882 letter to Marty, *WB*, p. 164. See also 'Booles rechnende Logik,' *NS*, p. 19 (17), where a connection with the incompleteness of concept-names is explicit.

[37] *Gl*, §60, p. 72.

[38] In a language like German with case endings, different series of letters may count as the same proper name.

[39] This view of predicates creates a problem in interpreting Frege's discussions of them. Frege introduces quotation names of proper names in the usual way: by

saturated, are integral wholes (*abgeschlossen*). Frege extends these characterizations to concepts and objects themselves.

This extension highlights that the application of these logical syntactic categories to sentences is an analysis of the thought expressed by the sentence 'Two is a prime number' as the completion of the predicate '__is a prime number' by the proper name 'two'. This is to analyse the sentence to say that the number 2, the object signified by the proper name, falls under the concept prime number, the concept signified by the predicate. Frege explains thoughts as that concerning which the question of truth arises. In the present case, as analysed, this question of truth becomes the question whether the designated object falls under the designated concept. Frege's talk of objects as saturated and concepts as unsaturated highlights the asymmetry of subsumption, an asymmetry constitutive of Frege's conception of objects and concepts.[40] In this way, as Frege says in an unpublished manuscript, 'Concept and object are fundamentally [*ursprünglich*] dependent on each other, and in subsumption we have their fundamental [*ursprünglich*] connection.'[41] However, we must take care how we understand this talk of subsumption. Frege's analysis of 'Two is a prime number' into proper name and predicate finds no separate symbolization of the relation of subsumption of object to concept. What is intended by the words 'falls under' in Frege's paraphrase of the original sentence is already meant in the original predicate. Concepts are unsaturated, predicative: subsumption is, so to speak, built into them.

enclosing the word-forms (*Wortbilder*) in quotation marks, we form names of the words, signs for signs. See 'On sense and reference', p. 28. Frege thus appears to use quotation marks to form names of proper names and sentences in accord with the contemporary convention for quotation names. However, without special comment, Frege uses quotation-marks to form names of predicates like:

'Socrates taught__'

If my remarks about the incompleteness of predicates are correct, this is a distinct use of quotation-names, one that cannot be assimilated to the contemporary convention. There is, however, no obstacle to taking these expressions formed using quotation marks to be proper names of predicates.

[40] See *Gl*, §51, p. 64, where Frege says: 'With a concept the question is always whether anything falls under it, and if so what. With a proper name, such questions make no sense.'

[41] 'Über Schoenflies', *NS*, p. 193. Frege sounds similar themes much earlier in his 29 August 1882 letter to Marty, *WB*, p. 164.

So far, we have a complementary difference between objects and concepts, but no priority attaching to objects. Frege maintains some such priority, especially when he says that objects, but not concepts, are self-subsistent (*selbständig*). In *Grundlagen*, Frege links the self-subsistence of objects to identity:

> In the sentence 'the number 0 belongs to the concept *F*', 0 is only an element in the predicate (taking the concept *F* to be the real subject) ... Precisely because it forms only an element in what is asserted, the individual number shows itself for what it is, a self-subsistent object. I have already drawn attention above to the fact that we speak of 'the number 1', and by means of the definite article put 1 forward as an object. In arithmetic this self-subsistence comes out at every turn, as for example in the equation 1 + 1 = 2.[42]

The self-subsistence Frege assigns to objects is the applicability to them of the relation *identity*. In contrast, the relation *identity* is not applicable to concepts. For concepts are what predicates mean, and predicates, on account of their incompleteness, cannot be the terms of equations. This point, however, does not carry us beyond the previously noted complementary difference between objects and concepts.

To understand the priority of objects to concepts, we need to consider a difficulty with Frege's conception of predicates as names that I have postponed. As names of concepts, predicates carry with them the possibility for introducing variables to generalize over the concepts. However, quantificational generality over concepts, parallel to that over objects, assumes that concepts comprise a multiplicity of discrete, determinate items. In the case of objects, the identity-predicate gives expression to this determinate discreteness. What does the 'determinate discreteness' of concepts come to?

The Leibniz inference sets the standard for the univocal use of proper names across sentences to designate the same objects. Suppose the thoughts expressed by 'Ø(*a*)' and by the equation '*a* = *b*' have been recognized to be true, but the thought expressed by 'Ø(*b*)' has been rejected as false. This result would show that either '*a*' is being used in the two premises to designate different objects, or that '*b*' is being used in the second premise and conclusion to designate

[42] *Gl*, §57, p. 68. See also §56, p. 67 and §62, p. 73.

different objects.[43] Let us then ask what substitution principle for predicates sets a parallel standard for their univocal use as names.

Consider first a sentence analysable as the completion of a proper name by a predicate. In principle, every predicate is obtainable from such a sentence. Such a sentence says that the object signified by the proper name falls under the concept signified by the predicate. So, regardless of what object the proper name designates, the truth-value of the sentence will be preserved, if we replace the predicate by a predicate that signifies a concept under which fall the *same objects* as fall under the concept signified by the original predicate. In sentences analysable as proper name and predicate, we may then substitute *salve veritate* for the predicate any predicate signifying a coextensive concept

Frege generalizes this principle to all occurrences of predicates in sentences, taking coextensiveness of concepts to set the standard for a univocal use of predicates generally. In this way, subsumption is the logically fundamental relation. Frege makes this point in his manuscript 'Ausführungen über Sinn und Bedeutung', where he explains that predicates signifying coextensive concepts are intersubstitutable *salve veritate.*

> The logically basic relation [*logische Grundbeziehung*] is that of an object's falling under a concept. All relations between concepts are reducible to it. Insofar as an object falls under a concept, it falls under every concept with the same extension, from which what has been said follows. As proper names of the same object can substitute for each other [*einander vertreten*] and preserve truth, so the same holds also of concept-words, if the concept-extension is the same.[44]

The contents expressed by sentences are what are recognized to be true or false. When the sentence is analysable as the completion of

[43] Still another possibility is that either '*a*' in the first premise or '*b*' in the conclusion is not being used as a proper name. As Frege's discussions in 'On sense and reference' show, there are in this connection a number of possibilities.

[44] 'Ausführungen über Sinn und *Bedeutung*'. *NS*, p. 128 (118). See also pp. 131 (120) and 132 (121); see also 'Husserl review', p. 320. In saying that every relation between concepts is reducible to subsumption, Frege does not, of course mean 'definable just in terms of subsumption'. It takes more than subsumption to define, for example, the relation of equinumerosity of *Gl*, §72. Rather, Frege means that the holding of any relation over concepts is fixed by the objects falling under those concepts. I.e., if concepts *F* and *G* stand in some relation, then any concepts coextensive with *F* and *G* stand in that relation.

a predicate by a proper name, truth or falsity is a matter of whether the designated object falls under the designated concept. The substitution principle for an arbitrary predicate that Frege extracts by considering sentences that are completions of that predicate thus reveals that concepts are, so to speak, individuated by their extensions. Coextensiveness is for concepts the analogue of identity for objects.

Concepts *F* and *G* are coextensive if all and only those objects falling under *F* fall under *G*. This analogue for a principle of individuation for concepts, in generalizing over objects, presupposes the discreteness of objects, the applicability to them of the relation of identity. Identity – the discreteness of objects – is for Frege a fundamental, irreducible given. The analogue of individuation for concepts thus depends on the discreteness of objects, but not conversely. Here we have the ontological counterpart to the syntactic identification of predicates in terms of sentences and proper names. This is what the priority of objects vis-à-vis concepts comes to.

We can now better appreciate the significance of Frege's quantificational conception of generality. On a traditional view of logic, the generality of a judgement is a matter of the abstractness of the concepts that enter into the judgement. This generality or abstractness is a comparative matter: the concept *mammal* is more abstract than the concept *horse*, and the concept *animal* is more abstract than the concept *mammal*. Logic, in order to be applicable to any judgement, must prescind from the contents of judgement, from the concepts that figure in judgement. Logic achieves this universal applicability by treating just the forms of judgements, the ways in which concepts of whatever degree of generality/abstractness are combined into judgements. So we find Kant saying:

> For the advantage that has made it so successful logic has solely its own limitation to thank, since it is thereby justified in abstracting – is indeed obliged to abstract – from all objects of cognition and all the distinctions between them; and in logic, therefore, the understanding has to do with nothing further than itself and its own form.[45]

[45] Immanuel Kant, *Critique of Pure Reason*, trans. and ed. Paul Guyer and Allen W. Wood (Cambridge: Cambridge University Press, 1998), p. 107 (B ix). See also D.198 (A 60 = B 85).

Frege's quantificational conception of generality opens the way to a very different conception of logic. Logical laws are the laws of a general theory of objects and concepts statable using only the topic-universal vocabulary needed for any special science. Logical laws get applied in the construction of proofs in a special science when we infer instances from them, instantiating variables by names drawn from the proprietary vocabulary of that special science. Frege's conception of a predicate is vital here, for it is generalization into the positions of predicate as well as those of proper names that enables Frege's logical laws to prescind from the content that distinguishes the statements of the various special sciences.[46]

Furthermore, as observed in §1,[47] Frege associates traditional logic with a psychological view of concepts as mental representations formed by abstraction from experience. Talk of the form as opposed to the content of judgement then assumes a psychological cast. In Frege's eyes, as traditionally conceived, logic all too easily becomes entangled with general accounts of cognition, with psychology. In contrast, on Frege's view, logic, as the maximal general science, does not concern itself directly with cognition and cognizers. Logic assumes a place among the sciences at their centre: its topic-universality establishes a framework for all of science.[48] In turn, the study of cognition takes its parochial place as a part of the special science of psychology. In this way, Frege's universalist conception of logic brings with it a clean and sharp distinction between the psychological and the logical.[49]

Frege's universalist conception of logic also gives him a logical conception of objecthood. The notions of *proper name* and of *object* are correlative; so are the notions of *predicate* and *concept*. Objects are the sort of item proper names designate: they are what are indicated by the letters that supplant proper names in corresponding generalizations. Similarly for concepts and predicates. Frege's

[46] For further discussion of Frege's view of the applicability of logic, see Goldfarb, chapter 3, this volume, especially §1. While a quantificational understanding or generality makes possible Frege's universalist conception of logic, it does not require it. See Goldfarb's comparison of a schematic conception of logic with Frege's universalist conception.

[47] See p. 6, especially the quotation from 'Booles rechnende Logik'.

[48] See *Bs*, Foreword, p. vi. See also the opening paragraph of 'On formal theories of arithmetic'.

[49] I am grateful to Michael Friedman for comments here.

logical laws spell out the intertwined conceptions of what objects are – not objects of this or that sort, but simply objects – and what concepts are, much as the axioms of Euclidean geometry spell out what points, lines and planes are.[50] There is no further extralogical content to the conception of *object* – nothing metaphysical, epistemological, or transcendental. Many commentators dwell on Frege's posit of causally inert (*unwirklich*), non-spatial objects, including logical objects. Frege himself emphasizes this point. On the interpretation I am urging, what is central in Frege is not his recognition of logical objects, but the logical conception of objecthood that makes the recognition of *unwirklich* objects look innocent.[51]

4

I have been discussing Frege's contrast between proper names with the objects they signify and predicates with their concepts. Predicates are not the only incomplete expressions Frege recognizes, concepts not the only unsaturated items. By removing occurrences of two proper names from a sentence, marking their positions with differently styled blanks, we form a dyadic predicate. For example, removing 'Socrates' and 'Plato' from 'Socrates taught Plato' yields

__taught __.

Or, following Frege and using the Greek consonants 'ξ' and 'ζ' as proper-name blanks, we get

ξ taught ζ.

As monadic predicates signify concepts, so dyadic predicates signify relations. As dyadic predicates, like monadic ones, contain blanks, so relations, like concepts, are unsaturated. However, a dyadic

[50] There is a difference between the two cases. 'Point and line' are among the primitive vocabulary of Euclidean geometry. The predicates 'object' and 'concept' do not appear in Frege's automatization of logic. His notion of objecthood is, in a way, linguistically expressed in Begriffsschrift not so much by a predicate, but in the use of letters in proper-name positions to generalize over objects without restriction. The same holds for Frege's notion of concepthood. However, it is trivial to formulate in Begriffsschrift a one-place first-level predicate true of all objects, and a second-level predicate true of all concepts.

[51] See *Gg*, p. xix, where Frege says, '**Wirklich** is only one of many predicates and no more concerns logic than the predicate **algebraic**, predicated of a curve.'

predicate can no more be substituted for a predicate in a sentence than can a proper name: one of its blanks would remain unfilled. Similarly, a monadic predicate cannot be substituted for a dyadic predicate. There is then a difference of type between concepts and relations like the difference of type between concepts and objects.[52]

Besides relations, Frege recognizes higher-level concepts. I noted how the isolation of proper names in sentences segments sentences into proper names and predicates. These predicates can then be recognized as meaningful units, as names, in more complex sentences. We may remove occurrences of a predicate from a sentence, marking its position with a predicate-blank. Just as expressions formed from sentences by removing proper names designate concepts under which objects fall, so expressions formed from sentences by removing predicates designate second-level concepts under which first-level concepts fall. Removing 'ξ is mortal' from 'Socrates is mortal' yields

φ(Socrates).[53]

This second-level predicate signifies the second-level concept that subsumes every first-level concept under which Socrates falls. So it subsumes the concept *philosopher* but not the concept *horse*. Removing the predicate 'ξ is a martian moon' from the sentence 'Something is such that it is a martian moon' yields

Something is such that φ(it).

This second-level predicate designates the second-level concept *existence* under which falls every first-level concept that subsumes at least one object. The concept *martian moon* falls under it; the concepts *venusian moon* and *not self-identical* do not. Finally, if we remove 'ξ is human' and 'ξ is mortal', from 'Everything is such that if it is human, then it is mortal', we get

Everything is such that if φ(it) then ψ(it).

This dyadic second-level predicate signifies the second-level relation in which one first-level concept stands to another if the first is subordinate to the second. Although Frege does not explicitly discuss

[52] See 'Function and concept', p. 29, and *Gg*, vol. I, §23, pp. 39–41.

[53] As Frege uses 'ξ' and 'ζ' to mark argument places for proper names, so he uses 'φ' and 'ψ' to mark argument-places for first-level predicates. See *ibid.*, §22.

the matter, it is clear that he takes higher-level concepts, like first-level ones, to be extensionally individuated.

The recognition of higher-level concepts is important for Frege's development of a notation adequate for multiple generality, for it enables him to overcome limitations on the use of (free) variables to express generalizations. We have seen how Frege uses variables to supplant names in sentences in order to confer generality of content on those sentences. However, this notation is not adequate for the translation of compound colloquial sentences with a component that is a generalization in its own right. The simplest example of its inadequacy, the one Frege presents in *Grundgesetze* §8, is the difference between the generalization of a negation and the negation of a generalization.

Frege accordingly introduces a new notation to express generalizations:

$$(\forall a)\varphi(a).$$

Here we have a second-level predicate that yields a sentence when a first-level predicate is inserted into the blank for first-level predicates. The second occurrence of '*a*' fills the empty place in the predicates inserted into this position. This second-level name forms an indivisible unit: in particular, '$(\forall a)$' has no significance on its own. Our second-level predicate signifies the second-level concept that subsumes those first-level concepts under which every object falls. Frege thus understands the universal quantifier over objects to be a second-level concept. Every object falls under the concept signified by 'if ξ is human then ξ is mortal'. So

$$(\forall a)(\text{if } a \text{ is human then } a \text{ is mortal})$$

is true. With this notation, we can express both the generalization of a negation and the negation of a generalization. To express the (false) generalization of the negation of

$$2 + (3 \times 1) = 5 \times 1,$$

fill the blank in our second level predicate '$(\forall a)\varphi(a)$' with the first level predicate 'Not $[2 + (3 \times \xi) = 5 \times \xi]$' to get

$$(\forall a) \text{ Not } [2 + (3 \times a) = 5 \times a].$$

In contrast,

$$\text{Not } (\forall a)[2 + (3 \times a) = 5 \times a]$$

expresses the (true) negation of the generalization of our sample equation.

We saw above how Frege explains the use of letters in proper name positions to confer generality of content on the sentence from which the names were removed. After introduction of a second-level predicate to express generality, Frege retains this use of letters, stipulating in effect that when (free) object-variables occur in sentences that are used to make assertions, those sentences express the same content as the corresponding generalization formed by completing in the predicate position in '$(\forall a)\varphi(a)$'.[54] So the sentence 'If x is human then x is mortal', if used to make an assertion, expresses the same thought as '$(\forall a)$(if a is human then a is mortal)'; and 'Not $[2 + (3 \times x) = 5 \times x]$' expresses the generalization of a negation.

Frege's treatment of quantifiers over first-level concepts parallels his treatment of quantifiers over objects. As the universal quantifier over objects is symbolized by a second-level incomplete expression, so the universal quantifier over first-level concepts is symbolized by a third-level incomplete expression. Removal of the two occurrences of 'ξ is mortal' from

Socrates is mortal or not (Socrates is mortal)

yields the second-level incomplete expression

φ(Socrates) or not φ(Socrates).

Frege's universal quantifier over concepts is

$$(\forall F)\mu_\alpha F(\alpha).$$

[54] See *Gg*, vol. I, §17, p. 31, and *Bs*, §11, p. 21. What I have referred to as free object-variables, Frege calls object-letters; he uses assorted Latin letters for this purpose. Frege uses gothic vowels in his notation for the universal quantifier to fill the argument places of first-level predicates. Frege does not explicitly assert in *Gg* that free variable generalizations prefixed by the judgement stroke express the same thoughts as their universal closures, although his rhetoric both in §17 and in §32, p. 50 suggests it. In later writings, Frege does speak of free variable generalizations as expressing the same thoughts as their colloquial counterparts. See 'Einleitung in die Logik,' *NS*, p. 206 (189ff.); 'Kurze Übersicht meiner logischen Lehren', *NS*, p. 217 (201); and 'Logische Allgemeinheit', *NS*, p. 280 (260). Frege's Begriffsschrift then gives us two options for asserting generalizations. As Frege explains in *Gg*, §17, this notational redundancy promotes simplicity and perspicuity in the formulation of inference rules.

Here '$\mu\alpha \ldots \alpha$' is a blank for a second-level incomplete expression. In filling this blank with a second-level predicate to obtain a sentence, 'F' is to be placed in the empty position in the second-level predicate. The second-level predicate is obtained by removal of first-level predicates from a sentence. The empty position in occurrences of the removed first-level predicates must be filled by something in that sentence. This something remains in the second-level predicate and must fill the parentheses that accompanies the 'F'. The 'α' in the blank is a reminder of this feature of second-level incomplete expressions. Completion of our third-level predicate by the sample second-level predicate then yields the sentence:

$$(\forall F)(F(\text{Socrates}) \text{ or not } F(\text{Socrates})).$$

The universal quantifier over concepts signifies the third-level concept that subsumes any second-level concept under which every first-level concept falls. Our sample second-level incomplete expression signifies a second-level concept which subsumes every first-level concept under which Socrates either falls or fails to fall. Every first-level concept satisfies this condition. Hence the sentence is true.

Frege's treatment of generality thus leads him to the recognition of second-level and third-level concepts. Although second-level concepts, like first-level ones, are unsaturated, they are unsaturated in a different way. First- and second-level concepts are thus fundamentally different in a way comparable to the fundamental difference between objects and concepts.[55]

In all this, Frege appears to have taken the first steps in generating a hierarchy of higher-level concepts with the structure of the simple theory of types. We begin with the recognition of objects (level 0) and first-level concepts. Recognition of concepts at any level is a basis for recognition of quantifiers over those concepts; and quantifiers over level n entities are level $n+2$ concepts. However, Frege, having introduced a quantifier over first-level concepts, introduces free variables over third-level concepts and then stops.[56] The application of logic he envisions for the development of the arithmetic of

[55] See 'On concept and object', p. 201; 'Function and concept', p. 31.

[56] These variables are needed to fill the empty place in '$(\forall F)\mu\alpha\, F(\alpha)$' in the statement of Basic Law IIb.

the natural and real numbers requires only free-variable generalizations over third-level entities.[57] He never offers any general characterization of a hierarchy of concepts comparable to Russell's theory of types.

There is one conspicuous feature of Frege's view of logical segmentation that I have left out of account. In connection with his introduction of the sense–reference distinction in 1891, Frege identifies sentences themselves as proper names of the truth-values, the True and the False; the first-level concepts designated by Fregean predicates are accordingly reconceived as functions that map each object to a truth-value. Higher-level concepts become higher-level functions. For example, the universal quantifier over objects is that function that maps a first-level function to the True if the first-level function maps each object to the True, and otherwise maps the first-level function to the False.

This strange sounding treatment of sentences fits with my interpretation of Frege's view of logical segmentation. I noted in §1 how, for Frege, the inference from generalization to instance compels the recognition in logic of logically relevant parts of sentences that are not sentences. There are also a number of elementary inference-modes that require the recognition of sentences as logically relevant parts of compound sentences. Contraposition and hypothetical syllogism are examples here.[58] In his formulation of logic, Frege seeks to capture such inferences-modes by general laws that accordingly use variables in sentential positions in compound sentences. So, among Frege's logical axioms in *Begriffsschrift* we find

If *a* then (if *b* then *a*)

and

[57] In the attempted deductive development of arithmetic in *Gg*, Frege uses Basic Law V to avoid higher-level generalizations so far as possible. He alludes to this strategy at the end of 'Function and concept', p. 31: 'One might think that this [the introduction of ever higher-level concepts] could continue. Probably, this last step [the introduction of second-level concepts] is not so rich in consequences as the earlier, because in further developments, instead of second-level functions, we can consider first-level functions, as will be shown in another place.'

[58] Frege presents a number of these patterns in 'Compound thoughts'. Frege does appreciate the need in his codification of logic to capture some such inference by means of an inference-rule rather than a generalization.

If (if *b* then *a*) then (if not *a* then not *b*).[59]

Sentences may be substituted for these variables to form instances of these generalizations.[60] Frege thus takes the inferences represented by in the application of these laws to abstract from the content of the special sciences in the same way as other logical inferences: the segmentation of sentences within compound sentences that accompanies these inference-modes is the segmentation of names within sentences.[61]

Even granting that sentences are compound names, what motivation does Frege have for taking them to be proper names like 'The teacher of Plato' or '$3 \times (4+1)$'? On Frege's quantificational understanding of generality, there is no entity without identity or surrogate for identity. So to quantify sentence positions requires that either the identity-predicate or some surrogate be available to voice claims of identity and difference over what sentences mean.

[59] The first law is proposition 1, *Bs*, §14, p. 26; the second is proposition 28, §17, p. 43. Hypothetical syllogism is expressed by proposition 28, a theorem, in §15, p. 32.

[60] The letters that occur in these Begriffsschrift formulas should be sharply distinguished from the schematic letters that appear in contemporary presentations of truth-functional logic. The schematic letters '*P*' and '*Q*' in a formula like 'If *P* then (if *Q* then *P*)' are placeholders that mark the positions of the component sentences of a truth-functionally compound sentence. The formula with these placeholders is not a sentence – it does not say anything; it is not true or false. Instead, it represents a form of truth-functionally compound sentence, and thus gives us a convenient way to specify an infinite class of such sentences, namely the sentences that result from uniformly replacing '*P*' and '*Q*' in our formula by sentences. For further discussion of the significance here of the distinction between variables and schematic letters, see Goldfarb, chapter 3, this volume.

[61] Frege does not present the segmentation of sentences within compound sentences as resting on the relationship of generalization to instance. See especially Frege's 1906 manuscripts 'Introduction to Logic' and 'A brief survey of my logical doctrines' and his final published paper 'Compound thoughts'. I think that Frege's presentation of truth-functional inferences is guided by pedagogical considerations. He could scarcely initially present truth-functional inferences and the segmentation accompanying them via general laws, for in colloquial language there is no need for quantification of sentence positions in truth-functionally compound sentences. Indeed, it is only the codification of logical inferences that calls for such quantification. Only after his audience has come to understand quantificational generality by appreciating how it is expressed, to the extent it is, in colloquial language, and how it may be expressed by the use of indicating letters in the place of names, can Frege retrospectively assimilate the segmentation of sentences within compound sentences to the segmentation of names within sentences.

Sentences are complete expressions: together with subsentential proper names, they are the basis for the identification of varieties of incomplete expressions. The completeness of sentences has two consequences. First, Frege's surrogate for identity for first-level concepts, namely coextensiveness, is an equivalence relation over first-level concepts whose characterization requires generalization over objects, the arguments for first-level concepts. In this way, then, the surrogate for identity for concepts presupposes generality over the multiplicity of discrete objects. In contrast, because sentences are complete expressions, there can be no surrogate for identity for what they designate parallel to this surrogate for identity for concepts. Second, as sentences are complete expressions, the logical segmentation imposed by generality creates no barrier to sentences' filling the argument places in the identity-predicate. Indeed, for this reason, any alleged surrogate for identity for the designata of sentences can be, and so should be, assimilated to identity.[62] If nothing bars the component sentences in expressions of compound thoughts from playing the role of proper names, then the attempt to formulate logical laws, as Frege conceives of such laws, requires that we recognize them to play this role.[63]

It is natural enough to assume that if sentences are names, then they designate the judgeable contents or thoughts Frege associates

[62] Formulation of the basic laws of any science, including logic, is guided by a maxim of economy. See 'Booles rechnende Logik', *NS*, p. 40 (36).

[63] Richard Heck has insightfully proposed another motivation for Frege's identification of sentences as proper names: this identification greatly facilitates the introduction of double value-ranges that are the objects that in Frege's system go proxy for relations. By making sentences proper names, Frege makes '$1 < \xi$' and '$\grave{\alpha}(\xi < \alpha)$' into designations of functions of the same type so that both may fill the argument position in the second-level function name '$\grave{\varepsilon}\Phi(\varepsilon)$'. See Richard Heck, 'The Julius Caesar objection', in R. Heck (ed.), *Language, Thought, and Logic: Essays in Honour of Michael Dummett* (Oxford: Oxford University Press, 1997), pp. 281–5. This additional motivation is compatible with the one I have urged. Indeed, it may well have been this use of truth-values as objects that prompted Frege to take the audacious step of taking sentences to be proper names of truth-values. Heck's motivation is not, however, the entire story here. It does not explain why Frege, after abandoning efforts to rehabilitate value-ranges in the wake of Russell's paradox, continues to take sentences to be proper names of truth-values and in 1906 lists the assimilation of concepts to functions as among his chief logical achievements. See 'Aufzeichnungen für Ludwig Darmstaedter', *NS*, p. 276 (255) and 'Was kann ich als Ergebnis meiner Arbeit ansehen?', *NS*, p. 200 (184).

with them. This assumption, however, clashes with the assimilation of sentences to proper names. If sentences are proper names, then there must be equations whose terms are sentences. Let '$\Phi(a)$' be any compound proper name containing proper name 'a'; let '$\Phi(b)$' be the corresponding proper name, with proper name 'b' replacing 'a'. The logic of identity that Frege incorporates into his formulation of logic commits him to

If $a = b$, then $\Phi(a) = \Phi(b)$.

In this way, the designation or meaning of any compound proper name remains unchanged under replacement within it of component proper names by proper names that designate the same thing. In 'Function and concept,'[64] Frege observes that the two sentences

> The Morning Star is a planet whose orbital period is less than that of the Earth,

and

> The Evening Star is a planet whose orbital period is less than that of the Earth,

express different thoughts: a person who does not know that the Morning Star is the same as the Evening Star might, nonetheless, understand both sentences and hold the one true and the other false. However, since 'the Morning Star' and 'the Evening Star' both designate Venus, these two sentences must designate the same thing.[65] Observing nothing other than the truth or falsity of a sentence that is guaranteed to be preserved under substitution of co-referring proper names, Frege concludes that sentences mean truth-values.[66]

[64] 'Function and concept', p. 14. See 'On sense and reference', p. 32. See also Frege to Russell, 28 December 1902, *WB*, p. 235, where Frege explicitly links the point with the use of sentences as terms of equations.

[65] Matters are particularly striking when we reflect that equations themselves are compound names. Here we get such instances of Leibniz's law as

If $2^4 = 4^2$, then $(2^4 = 4^2) = (4^2 = 4^2)$.

In this way, any true equation designates the same as an instance of the principle of identity, 'Everything is self-identical' ('$x = x$').

[66] I discuss Frege's identification of sentences as names of truth-values and of truth-values as objects in more detail in 'Quantification, sentences, and truth-values', *Manuscrito: Revista International de Filosofia*, 26 (2003), pp. 389–424.

5

We saw how in *Grundlagen* Frege says that objects are what proper names signify and concepts what predicates signify. As the linguistic categories of proper name and predicate are disjoint, so are the corresponding ontological categories of object and concept. Frege's contemporary Benno Kerry objected to Frege's linguistic elucidation of the concept–object distinction, maintaining that concepts may be designated by either proper names or predicates.[67] Consider, for example, the sentence,

The concept *horse* is a concept easily attained.

The subject of this sentence, 'The concept *horse*', is a Fregean proper name, and it must surely designate a concept. However, Frege's explanation of the distinction commits him to the contrary. Indeed, his explanation commits him to affirming the paradoxical statement:

The concept *horse* is not a concept.

This statement – the so-called Kerry paradox – appears to be an outright contradiction.

Frege writes 'On concept and object' to give a fuller exposition of the concept–object distinction in the face of Kerry's objection. Early in the paper, Frege dismisses the charge that his way of drawing the distinction leads to contradiction. He urges that designations like 'the concept *horse*' resemble quotation-names of linguistic expressions. Although a quotation name typographically contains the expression it names, it is itself unstructured, containing no other names as proper parts. Similarly, 'the concept *horse*' is an unstructured proper name. As proper names mean objects and no object is a concept, the Kerry paradox – however odd it may sound – is true.[68]

[67] My interest lies in what a series of philosophers beginning with Bertrand Russell made of Kerry's objection, and I examine the objection from Frege's perspective. For a treatment that sets the objection in the context of mid-nineteenth-century German views of logic, see Eva Picardi, 'Kerry und Frege über Begriff und Gegenstand', *History and Philosophy of Logic*, 15 (1994), pp. 9–32.

[68] Here I follow Anthony Kenny, *Frege* (London: Penguin, 1995), pp. 122–5. On p. 125, Kenny insightfully suggests, 'The expression "the concept …" is really meant to serve the same purpose with regard to concepts which quotation marks

Frege recognizes that this manoeuvre to sidestep the Kerry paradox does not resolve all the issues the paradox raises. He says:

> It is indeed unmistakable that we encounter here a linguistic awkwardness, one that I admit to be unavoidable … Language finds itself here in a predicament [*Zwangslage*] that justifies the departure from what is ordinary … In logical investigations, the need not infrequently arises to say something about a concept, and to clothe this in the usual form for such ascriptions so that the ascription is the content of a grammatical predicate. Hence one would expect a concept for the meaning of a grammatical subject. But a concept, due to its predicative nature, cannot just appear as a subject. It must be transformed into an object, or, to be more precise, it must be represented [*vertreten*] by an object which we designate by prefixing the predicate with the words 'the concept', e.g.
>
> The concept *human* is not empty.[69]

To understand better this linguistic awkwardness, let us consider Frege's analysis of existence sentences, including particular categoricals ('Some *F* are *G*'). As we saw in section 4, he analyses such sentences as the completion of a second-level predicate by a first-level predicate, not the completion of a first-level predicate by a proper name. The Fregean logician brings this point out in the material mode by reproducing the content of existence sentences to emphasize that these sentences say something about concepts, not objects:

> The sentence 'Human beings exist' (or 'There are human beings') says that the concept *human being* is non-empty.

Everyday language forces the Fregean logician for this purpose to paraphrase the existence sentence by a sentence that is the completion of a first-level predicate by a proper name. But the attempt misfires, for 'the concept *human being*' is a proper name, not a concept-name. The Fregean logician thus tries to use a sentence parallel to 'Socrates is a human being' to paraphrase 'There are human beings' in order to highlight the difference between the thought

serve in relation to predicates.' That is, the phrase 'the concept …' is used with conventional predicates as quotation marks are used with phrases generally to form unstructured proper names. Kenny thinks that this observation completely resolves the difficulty, neglecting the more serious problems that arise with Frege's use of 'concept' and 'object' as contrasting predicates.

[69] 'On concept and object', pp. 196–7.

expressed by the completion of the second-level existence predicate and that expressed by the completion of a first-level predicate.[70]

Frege reflects on this awkwardness in the famous antepenultimate paragraph of 'On concept and object':

> I do not at all contest Kerry's right to use the words 'object' and 'concept' in his own way. I would only like to secure the same right for myself, and to maintain that with my version I have fastened on to a distinction of the highest importance. Of course, a singular obstacle stands in the way of reaching an understanding with the reader. As a result of a certain necessity of language, sometimes my expression, if taken entirely literally, misses my thought in that I name an object where I mean a concept. I am myself fully aware that in such cases I am dependent on the goodwill of the reader who will meet me midway, who will not begrudge a pinch of salt.[71]

Frege, ever ready in his polemics to take the words of his contemporaries uncharitably literally, asks his audience to meet him halfway. They are to look beyond the content of the flawed paraphrase of the original existence claim to glom on to the distinction that Frege intends to communicate.

What, though, is this distinction? How is it to be described? Frege's evocations of it involve the use of pairs of contrasting predicates – 'object' and 'concept', 'saturated' and 'unsaturated'. The predicates are contrasting in that Frege clearly thinks the classifications they mark are non-vacuously mutually exclusive: while there are objects and there are concepts, no object is a concept. Here

[70] The same sentence may be analysable in different ways. So 'Some philosopher taught Plato' is analysable either as a singular sentence about Plato or as an existence sentence saying that the concept *philosopher who taught Plato* is non-empty. Frege presents a parallel example of multiple analysability in 'Einleitung in die Logik', *NS*, p. 203 (187). The relevant point here is that these two analyses are independent. Despite some indications to the contrary. I do not take Frege to be committed to the claim that any thought expressible by a sentence analysable as the completion of a second-level predicate is also expressible by a sentence analysable as the completion of a first-level predicate. In particular, I do not take Frege to believe that every sentence about a first-level concept (function) expresses the same thought as a corresponding sentence about the extension of that concept (the value-range of that function), although he does embrace this equivalence for the special case of generalized equations and equations. See 'Function and concept', pp. 10–11. This topic raises murky issues about Frege's view of the individuation of thoughts.

[71] 'On concept and object', p. 204.

(A) No object is a concept

appears to be parallel to

(B) No whale is a fish.

However, as Frege construes (B), 'whale' and 'fish' are both first-level predicates that signify concepts. (B) thus says that no object that falls under the concept signified by the first of these predicates falls under the concept signified by the second. The generalization is non-trivial, for there are objects falling under the concepts meant by these two predicates. If we attempt to construe (A) similarly, we encounter problems. Every object falls under the first-level concept signified by 'object'; and no object falls under the first-level concept signified by 'concept'. This second first-level concept is then empty, like the concept *not self-identical.* So construed, (A) is then equivalent to a generalization with an evidently vacuous antecedent. It thus sets forth a trivial generalization over objects, and so fails to capture Frege's views. To express the mutual exclusivity of the concept–object distinction we need predicates that can be completed both by object-names and concept-names, and a variable that generalizes simultaneously over both objects and concepts. As we saw in §2, the concept–object distinction bars such predicates and variables.

Kerry's objection thus touches more than proper names of the form 'the concept …'; it concerns the use of 'object' and 'concept' as contrastive predicates. Frege uses these words as first-level predicates repeatedly in his logical discussions, including the exposition of his Begriffsschrift notation in *Grundgesetze.* As a first-level predicate, the only names that can complete 'ξ is a concept' are proper names. This predicate cannot then be used to affirm that anything is, so to speak, a concept. The Kerry paradox thus reveals that none of Frege's logical discussions that turn on this contrastive use of 'object' and 'concept' can be taken at face-value.

Frege acknowledges this linguistic predicament in his letter of 29 June 1902 to Russell.[72] In his next letter, he says that since the concept *concept* is a second-level concept, use of the first-level

[72] Frege to Russell 29 June 1902, *WB*, p. 218. In this letter, Frege is discussing the more general distinction between function and object. See also 'Über Schoenflies', *NS*, p. 192 (177ff.).

predicate 'concept' is 'in fact logically to be rejected'.[73] Frege observes in his 29 June letter that in Begriffsschrift we can introduce a second-level predicate to do what we cannot do with 'ξ is a concept'. Corresponding to 'ξ is an object', we might introduce a second-level predicate '$\mathrm{Cpt}_a\varphi(a)$'. This predicate has the same grammar as Frege's universal quantifier '$(\forall a)\varphi(a)$'. We can fill the argument-place in 'ξ is an object' with any meaningful proper name and obtain a sentence expressing a truth, for example

2 is an object.

Similarly, we can fill the argument-place in '$\mathrm{Cpt}_a\varphi(a)$' with any meaningful predicate and obtain a sentence expressing a truth, for example

Cpt_a(*a* is a prime number).

Here, it seems, is an accurate expression for the thought aimed at with the sentence

The concept *prime number* is a concept.

However, Frege immediately goes on to observe that this second-level predicate in combination with the first-level predicate 'ζ is an object' does not give us replacements either for

The concept *prime number* is not an object

or for

2 is not a concept.

The argument place of 'ξ is an object' takes only proper names; a first-level predicate does not fit. The reverse holds for the argument place of '$\mathrm{Cpt}_a\varphi(a)$'. Furthermore, this disparity between the argument places of these predicates blocks as well their use in a generalization to replace 'No object is a concept'.

Even though the awkwardness of language exhibited in Frege's logical discussions can then be partially overcome by use of higher-level predicates within the framework of Begriffsschrift, nevertheless, there remain colloquial sentences that figure in Frege's logical

[73] Frege to Russell, 28 June 1902, *WB*, p. 224.

discussions that do not have Begriffsschrift surrogates.[74] Confronted with the unavoidable awkwardness of language that infects the contrastive use of 'object' and 'concept' in logical discussions, in his 28 July letter to Russell Frege advises resort to the formal mode, to talk about language:

If we want to express ourselves precisely, there is no other option than to speak of words or signs. We can analyse the sentence '3 is a prime number', into '3' and 'is a prime number'. These parts are essentially different. The former is in itself complete; the latter is in need of completion. In the same way, we can analyse the sentence '4 is a square number' into '4' and 'is a square number'. Now we can sensibly [*sinnvoll*] put together the complete part of the first sentence with the incomplete part of the second. (That the resulting sentence is false is another matter.) But we cannot sensibly put together the two complete parts: they don't stick to each other. Just as little can we sensibly put 'is a square number' in the position of '3' in the first sentence.[75]

Predicates, proper names and sentences are all objects for Frege.[76] There is then no difficulty in setting forth the mutual exclusivity of Frege's classification of names into those that are complete and those that are incomplete. However, if we restrict ourselves to talk of language, it seems that we leave out the basis for and the point of the distinction between complete and incomplete expressions.

Frege's 28 July letter to Russell continues:

To this distinction in signs must correspond a distinction in the domain of references, although it is not possible to speak about it without transforming what requires completion into something complete, and thereby actually falsifying the matter. We do this when we say 'the reference of "is a square number"'. However, the words 'is a square number' are not referenceless [*bedeutungslos*]. The analysis of a sentence corresponds to an analysis of the thought, and this in turn to something in the domain of references; and this I should like to call a basic logical fact [*logische Urthatsache*].[77]

[74] Terence Parsons makes this point in 'Why Frege should not have said "The concept *horse* is not a concept"', *History of Philosophy Quarterly*, 3 (1986), pp. 449–65, at p. 462.

[75] Frege to Russell, 28 July 1902, *WB*, p. 224.

[76] Frege explicitly calls names of function-names proper names in Frege to Russell 29 June 1902, *WB*, p. 218.

[77] Frege to Russell, 28 July 1902, *WB*, p. 224. Frege makes a similar remark in 'Foundations of geometry' (1903), p. 370. There he speaks of 'the analysis into

Here once again, in moving beyond talk of signs to talk of what signs refer to, Frege runs up against the linguistic predicament. The problem is that there is no way to set forth the difference in the domain of reference corresponding to the difference between the complete and incomplete parts of sentences without the use of contrasting predicates that in this context will fail, if taken literally, to draw the intended distinction. Here the sentence

(A) No object is a concept

is emblematic. We have seen that as Frege analyses it, it can only express a trivially true generalization over objects. We have found no other sentence, either in colloquial language or in the framework of the Begriffsschrift, that we can embrace as a literal and precise expression of a thought that the Fregean logician aims at, but misses with (A).

This consequence is troubling. Frege's basic logical laws are evident maximally general truths, the principles of a general theory of objects and concepts that establish a framework for all of science. The Begriffsschrift is a notation for expressing logical laws perspicuously and unambiguously so as to make possible the notationally secured rigour of gap-free proofs. Frege envisions the expansion of this language by incorporation into it of the predicates necessary to formulate the truths of the various sciences. Now, however, we appear to have encountered in connection with Frege's concept–object distinction truths that are not expressible in the framework of Begriffsschrift.

This point can be sharpened into a charge of incoherence at the core of Frege's philosophy of logic. Frege's universalist conception of logic commits him to the expressibility of every truth within the framework of Begriffsschrift. There is, however, no expression in Begriffsschrift for those thoughts Frege aims at in sentences like (A) that use 'object' and 'concept' as contrasting predicates. There are then truths, truths Frege points toward in his elucidation of the concept–object distinction, that are not expressible in Begriffsschrift, contrary to his core commitment. In this way, then, in the context

a saturated and an unsaturated part' as 'a logically basic phenomenon [*logische Urerscheinung*], which must be simply recognized but cannot be reduced to anything simpler'.

of his universalist conception of logic, Frege's explanations of the concept–object distinction appear self-thwarting.[78]

6

Frege himself sees in the linguistic predicament exhibited in his discussions of the concept–object distinction no threat to his enterprise. He concedes the unavoidable inaccuracy of some of his remarks, but shrugs off the difficulties that arise here as merely linguistic.[79] To understand this cavalier attitude toward the linguistic predicament, I want to consider Frege's response to one salient strategy for avoiding it.

In day-to-day life, including day-to-day science, there is little occasion to voice generalizations into predicate positions.[80] Not surprisingly, everyday language is impoverished in its devices for expressing such generalizations. The most salient means everyday English offers here is the use of words like 'everything' and 'something'. These words (and pronouns linked to them) may occupy proper-name positions; so used, they express generalizations over objects. When we use these words to generalize into predicate positions, grammar requires the retention of a verb to join subject and predicate. For example:

> Socrates taught philosophy.
> Plato taught philosophy.
> Socrates and Plato *have* something in common.
> There is something that Socrates and Plato both *did*.

[78] The objection that Frege's concept–object distinction is self-thwarting is a venerable one. Russell makes it in *Principles of Mathematics*. See §§49, 481 and 483. Without reference to *Principles*, Max Black revives it in 'Frege on functions', reprinted in E. D. Klemke (ed.), *Essays on Frege* (Urbana: University of Illinois Press, 1968), pp. 223–48, especially p. 242. I am indebted to Michael Resnik's careful and trenchant formulation of the objection in 'Frege's theory of incomplete entities', *Philosophy of Science*, 32 (1965), pp. 329–41, especially p. 339. There Resnik calls Frege's theory of incomplete entities 'self-referentially inconsistent'.

[79] See especially 'Foundations of geometry' (1903), p. 372, fn. 5.

[80] In 'Function and concept', p. 2, Frege observes that it is with 'higher analysis' that mathematicians sought to state laws about functions. Only with this comparatively recent development did the need arise to generalize into positions for signs for functions.

Such use of 'something' suggests that the position following the verb in these sentences, the position occupied by the expression of generality, is an accusative position, and leads to the use of various nominalizations of the predicate to express instances of the generalization.

> Socrates and Plato have something in common.
> They both have the property of having taught philosophy.
> They both fall under the concept *taught philosophy.*

The possibility of such nominalizations suggests that we might get by with just one level of generality. We might uniformly nominalize grammatical predicates, say by use of Frege's 'the concept'+ italicized predicate locution. The sentence 'Socrates is mortal' becomes

> Socrates falls under the concept *mortal.*

'Every human is mortal', is transformed into

> Everything falling under the concept *human* falls under the concept *mortal.*

We can follow Frege in taking the subject and object positions of 'falls under' to be proper-name positions, and so argument positions of the same type. We can thus use a single kind of variable to generalize into both positions. The object–concept distinction becomes a straightforward distinction within the range of this variable.

In the penultimate paragraph of 'On concept and object', the paragraph directly following the request for a pinch of salt, Frege considers this suggestion. I quote the entire paragraph:

> This difficulty [the linguistic predicament] might be thought to be artificially created: we do not need to bring anything as unmanageable as what I call a concept into consideration, and can follow Kerry in viewing an object's falling under a concept as a relation in which what on one occasion appears as an object, on another appears as a concept. The words 'object' and 'concept' then serve merely to indicate the difference in position in the relation. We can do this, but it is a gross error to believe that in this manner the difficulty is avoided. It is only shifted. For not all the parts of a thought may be complete. Rather, at least one part must be unsaturated or predicative – otherwise they would not stick together. For example, the sense of the phrase 'the number two' does not join with that of the expression 'the

concept *prime number'* without an adhesive [*Bindemittel*]. We supply just such an adhesive in the sentence 'The number two falls under the concept *prime number*'. It is contained in the words 'falls under' that requires completion in two ways. Not until it is filled out in this two-fold way do we have a complete sense, do we have a thought. I say that such words or phrases stand for a relation. Now we have with relations the same difficulty that we wish to avoid for concepts. For with the words 'the relation of an object's falling under a concept' we do not designate a relation but an object. The three phrases 'the number two', 'the concept *prime number*', and 'the relation of an object's falling under a concept' are to each other as mutually antipathetic [*verhalten sich ebenso spröde zueinander*] as the first two by themselves. However we assemble them, we do not get a sentence. Thus, we easily recognize that the difficulty that lodges in the unsaturatedness of a part of a thought can be shifted, but not avoided. Of course, 'complete' and 'unsaturated' are only metaphorical expressions, but in any case here I intend to and can only give hints.[81]

There is a natural, but mistaken, interpretation of Frege's point, as Cora Diamond has argued.[82] In this passage and in the 28 July letter to Russell, Frege observes that two proper names cannot be *sensibly* put together. The same goes for two predicates. The analysis of sentences into complete and incomplete parts corresponds to an analysis of the thoughts they express into complete and incomplete parts; and this analysis in turn corresponds 'to something in the domain of meanings', as Frege says in the letter. The natural interpretation finds an explanation at the level of reference for the cited linguistic phenomenon. The combinatorial valencies of words – the fact that some series of the signs constitute the expression of a thought, whereas other series of signs are nonsensical – is explained in terms of the combinatorial valencies of the senses of these words, and these valencies in turn are grounded in features of the items the

[81] 'On concept and object', pp. 204–5. See also Frege to Russell, 27 July 1902, *WB*, p. 224, 'Foundations of geometry' (1903), p. 372, and 'On Schoenflies', *NS*, p. 192 (177).

[82] My critique of the natural reading of the passage is based on and echoes Diamond's critique of the natural view of nonsense, which she develops with reference to Frege's concept–object distinction. She presents her critique in essays 2–4 of her collection *The Realistic Spirit: Wittgenstein, Philosophy and the Mind* (Cambridge, Mass.: MIT Press, 1991). I am especially indebted to essay 4, 'What does a concept-script do?' I am also indebted to Warren Goldfarb's discussion of Diamond's view in §2 of 'Metaphysics and nonsense: On Cora Diamond's *The Realistic Spirit*', *Journal of Philosophical Research*, 22 (1997), pp. 57–73.

words, in virtue of their sense, signify. Objects are saturated: they do not 'stick together'.[83] Because objects are saturated, they can be signified only by complete expressions that express complete senses. As the senses of two object-designating names do not fit together to form a thought, so the words expressing these senses do not fit together to form a sentence. Although 'two' and 'three' are meaningful proper names, 'two three' is nonsense. Because features of the items signified by names fix the combinatorial possibilities of those names into sentences, these features cannot themselves be described.[84]

The natural interpretation begins with the linguistic phenomenon that some series of words are thought-expressing sentences, and the others fail to express a thought. It seeks to explain this linguistic phenomenon by invoking the content (sense and reference) of names. From the perspective of Frege's Context Principle, this attempt at explanation is thoroughly wrong-headed. Consider the following two series of words:

(A) Three four

[83] See 'Foundations of geometry' (1903), p. 372.

[84] The natural interpretation thus sets Frege up for the charge of self-thwarting incoherence. Peter Geach seeks to avoid this charge, arguing that Frege's distinction between concept and object is a precedent for Wittgenstein's *Tractatus* distinction between saying and showing. As Geach views the matter, the sentence 'No object is a concept' is a flawed attempt to say an ineffable truth about logical categories that can only be shown. See Peter Geach, 'Saying and showing in Frege and Wittgenstein', *Acta Philosophica Fennica*, 28 (1976), pp. 54–70. I think that Geach is correct in finding a convergence here between Frege and Wittgenstein. However, following Diamond, I differ with his interpretation both of Frege on concepts and objects and of Wittgenstein on saying and showing. As regards Frege, Frege's conception of a thought as that concerning which the question of truth arises together with Frege's universalist conception of logic rules out any resort to ineffable truths about logical category-distinctions.

James Conant, in elaborating Diamond's interpretation of the *Tractatus*, suggests that Frege's view of the concept–object distinction wavers between the one Diamond and I attribute to him on the basis of the Context Principle and the one Geach attributes to him. See James Conant, 'Elucidation and nonsense in Frege and early Wittgenstein', in A. Crary and R. Read (eds), *The New Wittgenstein* (London: Routledge, 2000), pp. 174–217, especially, p. 177. (Note, however, that in fn. 12 Conant refrains from endorsing his suggestion as Frege interpretation.) Passages in Frege like the one from the end of 'On concept and object' can all too easily be taken to confirm Conant's suggestion of a tension in Frege here. Below I present a reading of Frege's talk in these passages of items not fitting together that reconciles such rhetoric with Diamond's and my interpretation of the concept–object distinction.

and

(B) Three is a square number.

Frege begins with thoughts and sentences like (B) that express them. A grasp of generality, of the inference from generalization to instance, leads to the recognition in thought-expressing sentences like (B) of proper names and thence predicates. The recognition of an expression as a proper name thus depends on a prior discrimination of those series of marks or sounds that express thoughts from those that do not. Given the priority of thoughts and sentences to their logically segmented parts, there can be for Frege no explanation of why (A) fails to express a thought, or why (B) does, that invokes the notion of a name.

As regards (A), we recognize it to be a juxtaposition of the words T-H-R-E-E and F-O-U-R that occur as proper names in sentences like (B). We could make (A) into a sentence containing the proper name 'three'. We could do this by giving 'ξ four' a predicative use, say by stipulating

> Something four if and only if it is a positive integer that is equal to the sum of its factors smaller than itself.

Then 'Six four' would express a truth, and (A) a falsehood. (A) would not, however, contain the proper name 'four'. Similarly, we could give 'three ξ' a predicative use so that (A) would be a sentence containing the proper name 'four'. Finally, we might stipulate that 'ξ ζ' – SPACE preceded and followed by proper name argument-places – is to be an expression for a dyadic relation, say the relation of identity. But as things now stand, (A) is not a sentence: it neither contains the proper name 'three', as the sentence (B) does, nor does it contain the proper name 'four'.

The point of Frege's remark about the failure of proper names to stick together into sentences is then straightforward. For a proper name to occur in a sentence is for that sentence to be the completion of a predicate by that proper name. Any sentence analysable as containing a proper name must then contain a predicate as well. Similarly, any sentence analysed as containing two proper names is thereby analysed as the completion of a dyadic predicate by those proper names. There is then no sentence that contains just two proper names and no incomplete expression. Frege conceives of sentences and thoughts as structurally parallel. In particular, the analysis of the expression of a thought into complete and incomplete parts is an analysis of the thought expressed into corresponding complete and incomplete parts.[85] This last point about the analysability of sentences thus equally applies to thoughts.[86]

The argument in the penultimate paragraph of 'On concept and object' is an application of this point about the analysability of sentences and thoughts. Frege allows that we can replace

(A) 2 is a prime number

with

(B) 2 falls under the concept *prime number.*

(B) is a relational sentence, analysable as the completion of 'ξ falls under ζ' by the proper names '2' and 'the concept *prime number*'. The dyadic predicate 'ξ falls under ζ' thus appears in the sentence as the designation of a relation. The same awkwardness that surrounds Frege's talk about concepts, about what monadic predicates signify, reappears in his talk about relations, about what dyadic predicates signify. The linguistic awkwardness has not been avoided, but at best only shifted.

We can now appreciate why Frege repeatedly terms his explanations of the concept–object distinction hints.[87] In connection with his debate with Hilbert over Hilbert's consistency proofs for geometry, Frege distinguishes elucidations (*Erläuterung*) from genuine

[85] See above, pp. 164–5.

[86] I have explained Frege's remarks about the complete parts of a sentence and the thought it expresses not sticking together. In 'Foundations of geometry' (1903), p. 372, Frege talks of objects themselves in these terms. (Perhaps this rhetoric is a survival from his pre-1891 conception of judgeable contents.) In this paper, Frege presents his application of the terms 'saturated' and 'unsaturated' to objects and concepts as metaphorical (*bildlich*). I suggest that his talk of objects not sticking together is a part of the metaphor. It points toward the differences in logical type built into to Frege's conception of quantificational generality. For further discussion of this point, see the opening three pages of §3.

[87] In the discussion of elucidation in the following three paragraphs I am indebted to Joan Weiner, *Frege in Perspective* (Ithaca, N.Y.: Cornell University Press, 1990), especially ch. 6, pp. 249–56.

definitions.[88] Elucidations are the remarks investigators make in the effort to bring an audience to a shared understanding of the basic vocabulary that figures in the statement of the basic results of a branch of science. Frequently, there is no word in colloquial language that unequivocally and uniformly expresses the exact sense an investigator wishes to attach to a term. An investigator has to talk around the matter, relying on goodwill and good sense to lead her audience to glom on to the intended sense.[89] Then, once the basic, primitive vocabulary of a science is in place, new words may be introduced by definition as abbreviations for complex expressions constructed from the basic vocabulary.[90]

Frege faces a daunting elucidatory task in explaining the fundamental vocabulary of the science of logic.[91] Here it is not only a matter of conveying unfamiliar notions – the material conditional, value-ranges and even identity – that receive no unambiguous, uniform and salient expression in everyday language. Frege has to explain the use of a language with a different grammar from colloquial language to express thoughts. This is first of all a matter of explaining how familiar generalizations, first-level generalizations over objects, are expressed by Frege's quantifier-variable notation. In addition, he has to get his audience to understand as well the expression of higher-level generalizations whose expression in everyday language is frequently awkward, if available at all. Frege is thus led to use 'object' and 'concept' as contrasting predicates to point to the difference between the argument positions determined by proper names and predicates, to point to the difference in the use of variables in those positions to express generalizations, while

[88] See Frege to Hilbert, 27 December 1999, *WB*, p. 63; 'On the foundations of geometry' (1906), p. 301; 'Logik in die Mathematik', *NS*, p. 224 (207).

[89] See 'On the foundations of geometry' (1903), p. 301.

[90] In addition to inference rules, Frege equips his Begriffsschrift with rules for introducing new names by definition.

[91] Apart from 'On concept and object', p. 193, Frege does not apply the word 'elucidation' (*Erläuterung*) to his explanations of the concept–object distinction, preferring to call these explanations 'hints'. While a grasp of the concept–object distinction is necessary for understanding Begriffsschrift, neither 'object' nor 'concept' is a primitive Begriffsschrift name: this distinction is not one capturable by contrasting predicates. Perhaps this is why Frege speaks in this context of hints. I am indebted here to Göran Sundholm and Michael Kremer.

denying that this difference is like the distinctions marked by other pairs of contrasting predicates.[92]

Although Frege's elucidations of the concept–object distinction thus sound confused, their self-stultifying character points his audience in the right direction. In Begriffsschrift, the difference between concepts and objects is expressed by the use of different types of variables, and so cannot be expressed by contrasting predicates. For the audience that has mastered Begriffsschrift, the confusion latent in Frege's elucidations becomes manifest when they attempt to paraphrase these remarks into Begriffsschrift. They find here no thought towards which 'No object is a concept' aims. There is only the trivial generalization that comes to little more than the tautology that no object is a non-object. Recognizing this, the master of Begriffsschrift is free to discard the contrasting use in Frege's elucidations of the predicates 'object' and 'concept', of 'saturated' and 'unsaturated', as so much hand-waving. There is no residue that goes unexpressed in Begriffsschrift. So it is in hopeful anticipation of his reader's mastery of Begriffsschrift that Frege requests a pinch of salt.

Frege's argument at the end of 'On concept and object' assumes his view of generality and logical segmentation. It may then appear to beg the question against Kerry, for Kerry's objection can be read as questioning this view. Kerry maintains that what a predicate designates may be designated as well by a proper name. If everything is designatable by proper names, then we can use variables in proper-name positions to generalize simultaneously over everything, Frege's objects, concepts, relations and higher-level concepts alike. Variables in other grammatical positions will then presumably have a restricted domain in comparison to the unrestricted variables in proper-name positions. Nominalization of predicates along the lines discussed at the beginning of this section opens the way to replacement of such a restricted predicate variable by an unrestricted proper-name variable. The tables are turned as regards Frege's argument at the end of 'On concept and object'. The

92 Peter Geach, in 'Saying and showing in Frege and Wittgenstein', emphasizes the role of Frege's contrastive use of 'object' and 'concept' to inculcate in his audience an understanding of Begriffsschrift. See pp. 55 and 58.

relation designated by the dyadic predicate 'falls under' is one more object, designated as well by the proper name 'the relation of an object's falling under a concept'. This is essentially the approach to segmentation and generality Russell adopts in *Principles of Mathematics*.[93]

This formulation of the Kerry objection places great weight on the notion of designation (*bezeichnen*) or meaning (*bedeuten*). The identity Kerry maintains between concepts and some objects cannot be asserted by using a proper name and a predicate as terms of an equation. We have to go metalinguistic here and resort to semantic ascent:

> The predicate 'ξ is a horse' designates the concept *horse*

or

> The proper name 'the concept *horse*' and the predicate 'ξ is a horse' both designate the same thing.

In Frege's eyes, semantic ascent provides no way around the linguistic predicament. Although Frege himself speaks of predicates' designating concepts, he recognizes that this talk is problematic. In the sentence

> 'Socrates' designates Socrates

we find a dyadic first-level predicate, 'ξ designates ζ'. We can then form the predicate

> ξ designates something

i.e.

> $(\exists a)(\xi$ designates $a)$.

No completion of this blank by the name of a predicate will yield a truth. In his 28 July letter to Russell, Frege accordingly advises, 'We

[93] See especially *Principles of Mathematics*, ch. 4. As mentioned above, Russell urges the Kerry paradox as an objection to Frege's concept–object distinction. However, the unrestricted nominalization of predicates leads to Russell's paradox. In order to avoid the paradox, Russell tentatively advances the simple theory of types in appendix I of *Principles*.

cannot really say of a concept-name that it designates something, although we can say that it is not meaningless.'[94]

To appreciate the force of this point against Kerry, we need to consider more closely Frege's conception of designation. I urged in §1 that Frege's conception of designation is rooted in the contrast between the use of a sign in a sentence as a proper name and the use of a sign in that position as a variable. The variable generalizes over a multiplicity of determinate objects, one of which is designated by each of the ideally unambiguous proper names that may replace it to form instances of the generalization. We identify objects by the use of equations whose terms are proper names. So, for example, we can identify the object designated by the proper name '2' by use of an equation of the form:

The object designated by '2' =___.

In this way, the identification of the object designated by a proper name requires the use of a proper name in an equation. Furthermore, we will always be able to identify the object designated by a proper name by the use of that proper name:

The object designated by '2' = 2.

Finally, once '2' has been recognized in sentences as a designating proper name, we can inquire after the identity of the object designated by the name '2' by inquiring after the truth of thoughts expressed by equations of the form

2 = ___.

Talk of designation drops away.

I have been talking about proper names and designation. What about predicates? In speaking of predicates as designating concepts, Frege brings out the parallels in the use of variables in both proper name positions and predicate positions to confer generality of content on sentences. However, as we have seen, these elucidations land Frege in the linguistic predicament. Nevertheless, there is an

[94] Frege to Russell, 29 June 1902, *WB*, p. 219. On this suggestion, 'meaningless' cannot then be defined using the dyadic predicate 'ξ means ζ'. See also 'Ausführungen über Sinn und *Bedeutung*', *NS*, p. 133 (122).

innocent enough notion of designation applicable to predicates that avoids the predicament.

To understand the application to predicates of the analogue to the notion of designation for proper names, consider what corresponds to the relation of identity as regards the multiplicity generalized over by predicate variables, namely coextensiveness. As equations identify objects, the concept *prime number* is, so to speak, identified by the generalized biconditional:

> Something is a prime number just in case it is a positive integer > 1 which is divisible without remainder only by 1 and itself.

Frege observes that there is in colloquial language a predicative use of 'what the predicate ___ means', as in

> Jesus is what the predicate 'human' means.[95]

The proper colloquial analogue then for the predicate 'human' to 'The reference of "2" = 2' is:

> Something is what the predicate 'human' designates just in case it is human.

As with proper names, we have to use a predicate to identify the reference of a predicate, so to speak, and can use that very predicate to do so.

Furthermore, nothing formally blocks the expression in Begriffsschrift of this use of 'designates' in connection with predicates. To specify reference for first-level monadic predicates, we might introduce a dyadic mixed-level predicate whose first argument position is for proper names and whose second argument position is for monadic first-level predicates:

> ξ DESIGNATES$_a$ (φ(a)).

So we have sentences like

> 'ξ is a prime number' DESIGNATES$_a$ (*a* is a prime number),

[95] *Ibid.*, p. 133 (122). I am indebted in this paragraph and the next to Michael Dummett's discussion of this passage in *Frege: Philosophy of Language*, pp. 212–19.

a sentence that might be taken to give accurate expression to the thought aimed at with the sentence

> 'ξ is a prime number' designates the concept *prime number.*

However, we will not be able to use this higher-level predicate in combination with the first-level 'ξ means ζ' to formulate accurately Kerry's identification of the reference of a predicate with the reference of a proper name. Here we encounter again the same obstacles that block the use of 'ξ is an object' and '$\mathrm{Cpt}_a\varphi(a)$' to formulate accurately 'No object is a concept'.

We can now appreciate why Frege would remain unmoved by the latest version of the Kerry objection. Frege's Context Principle makes the notion of a name, the occurrence of a name in judgement-expressing sentences, prior to talk about designation, about what those names designate. The central, underlying purpose of Frege's talk of designation is to make salient within everyday language the segmentation of sentences into proper names and predicates. In this way, I have argued, Frege both exposes the notational pattern for Begriffsschrift and instructs his audience in the use of variables corresponding to different types of names to express generality.[96] In

[96] I believe that talk of designation is purely elucidatory for Frege. It belongs to the preliminaries that establish a notation for the full, unambiguous and perspicuous expression of scientific knowledge. Once this notation is in place, talk of designation is not required for the communication of any proper scientific knowledge, since 'identifications' of particular objects and concepts can be expressed without semantic ascent Some will think that, in relegating talk of designation to the elucidatory propaedeutic for the communication of proper scientific knowledge, I neglect the need for a notion of designation within science to state a theory that describes how sentences are determined to be true or false by the designata of the names into which they are analysed. Michael Dummett and commentators who follow him see such a semantic theory to underlie Frege's codification of logic. For a discussion of some of the issues here, see Richard Heck, chapter 9, this volume.

My concern here has been with the connection between Frege's talk of designation and the concept–object distinction. I have not then argued for my strong claim about designation. In particular, the type theory which Frege's concept–object distinction imposes on any theory does not by itself exclude semantic ascent. In my view, semantic theorizing of the sort Dummett, Heck and others find in Frege is irrelevant to logic, as Frege understands it on his universalist approach. My portrayal of Frege's elucidations of the concept–object distinction in the context of his view of the role of the Begriffsschrift shows how many of Frege's remarks about designation do not have to be understood as a proto-semantic theory. I do think that Frege's view of truth excludes a notion of designation this setting, we have no purchase on the notion of designation apart from the use of a proper name or predicate to 'identify' (in the appropriate fashion) what a given proper name or predicate designates. In contrast, the Kerry objection, by permitting the use of a proper name to specify the designation of a predicate, makes the notion of designation independent of the logical syntactic categories of proper name and predicate.

After the opening exposition of the concept–object distinction in 'On concept and object', Frege says:

> Of course, Kerry thinks that logical determinations cannot be based on linguistic distinctions. However, no one who makes such determinations can avoid basing them on linguistic distinctions in the way that I do. For without language we cannot come to an understanding, and so are, in the end, thrown back on the confidence that others understand words, forms, and sentence-construction essentially as we do. As I have already said, I do not intend to give definitions, but only to give hints by appealing to the general feeling for the German language.[97]

Communication of any determinations relies on a shared understanding of the language used to this end. This dependence is especially weighty in logical investigations, as Frege conceives of and conducts them. Frege assumes that his audience can, as he puts it, recognize the same thought in various linguistic guises.[98] In particular, Frege assumes his audience's tacit grasp of the various ways quantificational generality is expressed in everyday language, to the extent that it is there expressible. By assembling hints that highlight occurrences of proper names in sentences, Frege aims to bring his audience to an explicit awareness of the inference from generalization to instance. From his perspective, Kerry's objection rests on a failure first to distinguish this distinctive inference-mode, and second to appreciate how the difference it exposes between the complete and incomplete parts of sentences and the thoughts they express evinces the striation of quantificational generality into levels. For Frege, the debate stops here. The striation of quantificational

from within science. See my 'Logic and truth in Frege', *Proceedings of the Aristotelian Society*, supplementary volume 70 (1996), pp. 121–40.

97 'On concept and object', p. 195.

98 Thus, in a section of the 1897 manuscript 'Logik' entitled 'Separation of a thought from its wrapping', Frege commends the value of foreign-language study for aspiring logicians. *NS*, p. 154 (142).

generality into levels, thus conceived, is the 'logically fundamental fact' of which Frege speaks in his letter to Russell.[99]

7

Frege's Context Principle is the second of three fundamental principles that guide his logico-mathematical investigation of the concept of number in *Grundlagen*. It remains to examine how the Context Principle shapes Frege's analysis of the concept of number and his attempt to reduce arithmetic to logic via the introduction of extensions for concepts.[100] I shall maintain that the quantificational conception of generality encapsulated by the Context Principle gives shape and substance to Frege's task of analysing the concept of number. It constrains what will count as an analysis; it leads Frege to logical resources for his attempted analysis; finally, it prompts the introduction of extensions, for only in this way can Frege bring those logical resources to bear on the analytic task at hand.[101]

The foundations for arithmetic Frege aspires to lay involve proofs of fundamental laws for systems of numbers from non-arithmetical premises, proofs that are based on analyses, definitions, of the basic vocabulary of arithmetic in non-arithmetical terms. By integrating arithmetic into the body of science as a deductive development of a more abstract body of knowledge, these proofs will clarify the subject matter of arithmetic, the character of our knowledge of it, and its regulative status as a common presupposition of the various sciences. Frege points to the rigorization of the calculus as a kind of model for his inquiry. There, however, the arithmetic of the real numbers provides resources for analysis. When we come

[99] See the references in footnote 61.

[100] More generally, via the introduction of value-ranges for functions.

[101] Michael Dummett sees the Context Principle as having these roles as well. See his paper 'The Context Principle: Centre of Frege's philosophy', in Max Ingolf and Werner Stelzner (eds.), *Logik und Mathematik: Frege-Kolloquium, Jena 1993* (Berlin: De Gruyter, 1995). In addition, however, Dummett thinks the Context Principle serves another, more elusive purpose, namely to answer the Kantian question, 'How are numbers given to us?' I shall maintain that in the setting of Frege's views on judgement and logic, the Context Principle transforms the Kantian question precisely by fulfilling the roles that both Dummett and I acknowledge.

to the positive integers, we appear to have reached conceptual bedrock: here we find no ready-to-hand resources for proof-generating analysis and no unclarity to be removed. Frege emphasizes this point in the introduction to *Grundlagen:*

> So free from all difficulty is the concept of positive whole number held to be that an account of it fit for children can be both scientific and exhaustive; and that every school boy, without further reflection or acquaintance with what others have thought knows all there is to know about it … The truth is quite the other way: the concept of number, as we shall be forced to recognize, has a finer structure than most of the concepts of the other sciences, even though it is still one of the simplest in arithmetic.[102]

To begin, if we focus on arithmetical equations and inequalities – for example, '15 + 8 = 23' – we find no fine structure for the concept of number. Indeed, we will likely succumb to a crude formalism which Frege rejects on the grounds that it severs the logical relevance of pure arithmetical equations to statements of number, statements answering answer the question, 'How many F are there?' Instead, in chapter 2 of *Grundlagen*, Frege focuses on statements of number themselves in order, as he puts it, 'to assign to *number* its place among our concepts'.[103] He argues first against the naive view according to which, in statements of number, a number is predicated of a heap or collection of things. Perhaps the most telling objection is that the naive view is committed to affirming and denying the same number of the same heap of footwear in my closet (8, not 4, shoes; 4, not 8, pairs of shoes). After rejecting the naive view, Frege goes on to dismiss the view that these statements predicate numbers of ideas of collections of things.

With our hindsight, we know the direction in which Frege is moving: statements of numbers are not about groups of things; nor are they about ideas of groups; rather, they are about objective concepts. This way of thinking of matters is, I believe, misleading, for it places Frege's view of statement of numbers on a level with the ones rejected in chapter 2. In chapter 2, in suggesting that statements of number predicate number of something, and asking 'What are the subjects of this predication?', Frege follows the surface grammar of these statements with their seemingly adjectival use of numerals

[102] *Gl*, pp. iii and iv.
[103] *Ibid.*, §21, p. 27.

and does not question traditional subject–predicate analysis. The entire polemic is supported by the first guiding principle that bids us separate the psychological-subjective from the logical-objective. In chapter 3, the weight of the polemic shifts to the second two guiding principles, to Frege's distinctive view of logical segmentation. Let's consider this crucial transition in Frege's exposition of his views.

Despite the rejection of the naive view of *number* in chapter 2, Frege devotes the bulk of chapter 3 to consideration of Euclid's maxim, 'A number is a collection of units'. In Stanley Jevons's work, Frege encounters a development of Euclid's view. A person distinguishes the items in a group. Ignoring the distinguishing features of these items, she use repetitions of a mark, say a stroke '|', to represent each. So the telephone, the printer, and the pack of printer paper on the right-hand corner of my desk might be represented by '| | |'. I can then introduce the sign '3' as an abbreviation for '| | |': or, using a mark, '+', to replace sheer concatenation, '3' abbreviates '|+|+|', and '3+2' stands as an abbreviation for: '|+|+|+|+|'. In this way, a number is a sum of units, a sum of ones.

Frege's critique focuses on the notion of representation operative here. Jevons's account slides from the 1–1 correlation of the items to be counted with an array of strokes to the identification of the strokes with occurrences of the numeral '1'. Frege exploits this slide to embarrass Jevons's account. His polemics turn on the following point. The manipulations of arithmetic terms in equations turn on treating various occurrences of numerals, including the numeral '1', as signifying the same thing. Jevons's account, in positing wholesale equivocation here, renders these manipulations of terms logically unintelligible, but cloaks this equivocation by sliding from talk of units to talk of (the number) One.[104]

It is here that the basis for Frege's polemic shifts to the second two fundamental principles. Jevons's view takes the strokes used in an underlying notation of numerals to symbolize or designate the distinct units of arbitrary collections. Accordingly, the way in which in the stroke notation a stroke is used to signify a unit of a collection is distinct from the sense in which a personal proper name or a demonstrative phrase is used to signify a person.

[104] See *Gl*, §38, p. 48ff.

Furthermore, Jevons's notion of signification, based on a correlation of marks with the units to be counted, is signification both outside of and allegedly prior to the use of numerals in statements of number. Frege's polemic against Jevons in effect observes that Jevons's notion of signification cannot be captured in terms of the notion of meaning that the Context Principle makes available, a notion of meaning based exclusively on the logical segmentation of statements. We cannot understand the strokes to be unambiguous proper names of distinct items; if we take them to be ambiguous names we lose our purchase on the manipulation of numerical terms in equations. Finally, Frege observes that the word 'unit' is a predicate, while the numeral '1' is a proper name. The slide from talk of a collection of units to talk of a sum of one's then represents a failure to attend to the difference between concept and object.

At the beginning of §46, Frege returns to statements of number, now explicitly invoking the Context Principle:

> To throw light on the matter, it will be good to consider number in the context of a judgement where its basic application comes to the fore.

Frege does not, however, at this stage focus on the role of the numerical terms that occur in statements of number. He focuses on the content of these statements taken as wholes, returning to his earlier suggestion that statements of number predicate number of something. The context for the suggestion now, however, is the quantificational view of logical segmentation encapsulated in the Context Principle. Frege's rhetoric over the next nine sections is designed to reshape his audience's logical gestalt to initiate them into this viewpoint without instructing them in Begriffsschrift. In these sections, above all in §46 and §51, Frege explains the distinction between proper names and predicates, between objects and concepts, along the lines considered in my §§1 and 2 above. In §53 Frege notes that existence-statements are examples of statements that say something about a concept. They give Frege a model for the segmentation of statements of number. Just as the existence-statement 'There are martian moons' says that there is at least one object falling under the concept *martian moon*, so the statement of number 'There are (exactly) two martian moons' says that there are two and only two objects falling under the concept *martian moon*. In this way statements of number predicate something of concepts.

The comparison of statements of number with existence-statements does more than reveal the logical segmentation of the former. It makes the first, decisive step in uncovering logical resources for the analysis of *number.* Frege observes that statements of number that assign to a concept the number 0, statements like 'There are 0 venusian moons' or equivalently 'There are no venusian moons', simply deny what the corresponding existence-statement, 'There are venusian moons', asserts. Once we look at matters in this way, it is clear that the content of the second-level predication in any statement of number can be reproduced in logical terms by exploiting embedded quantifiers over objects and the identity sign. For example, 'There is one *F*' goes over into

$$(\exists x)[Fx \,\&\, (\forall y)(Fy \rightarrow x = y)].$$

'There are two *F*' is paraphrased by:

$$(\exists x)(\exists y)[Fx \,\&\, Fy \,\&\, {\sim}(x = y) \,\&\, (\forall z)(Fz \rightarrow (z = x \vee z = y))]$$

In the logical elements used to reproduce the content of the predicates in statements of number, we encounter resources for the analysis of the concept of number.

How are the logical resources for the analysis of number Frege has uncovered to be applied? Frege's views on logical segmentation shape his view of analysis. For Frege, analysis is definition, and definitions present notational equivalences by setting forth a simple expression as equivalent to a compound expression. Thus we can always adopt the pretence that a definition introduces the definiendum into the language for the first time as an abbreviation for the definiens.[105] Definitional equivalences must respect logical types, so that 'expression' in this context means 'Fregean name' – proper name, first-level predicate of one argument-place, first-level predicate of two argument-places, higher-level predicate. Frege's rhetoric in the opening sections of chapter 4 of *Grundlagen* is to

[105] As Joan Weiner has argued, this view of definitions is a stable, deeply entrenched element in Frege's thought. It fits the very brief discussion of definitions in *Bs*, §24, as well as the definitional practice of that monograph. I do not think that Frege ever entertains the possibility of 'contextual' definitions of expressions that are not names, not logically segmented units, as Russell does with his contextual definitions of incomplete symbols like definite descriptions and class abstracts.

initiate his audience into this viewpoint as regards the definitions that constitute the analysis of *number.* That will involve establishing the basic logical segmentation of pure and applied arithmetic discourse and then bringing the logical resources Frege's investigation has uncovered to bear on that discourse.

Frege's regimen for quantificational paraphrases of statements of number is in effect a routine for defining a series of second-level predicates: 'the number 0 belongs to (the concept) *F*', 'the number 1 belongs to *F*', 'the number 2 belongs to *F*', etc. These definitions treat each second-level numerical predicate as an unstructured whole. We have then no definitions for the apparent proper names 'the number 1', 'the number 2' – only definitions for predicates that contains these phrases as logically unsegmented parts. This means that these definitions for these second-level predicates give us no general conception of *number.*

Frege had observed in chapter 3 that numerals appear as proper names in the statements of pure arithmetic. Indeed, as he puts it in §57, 'The form of equations is the predominant form in arithmetic.' Furthermore, Frege notes that statements of number themselves can be paraphrased as equations: 'There are two martian moons' goes over into 'The number of martian moons = 2'. The logical connections between pure arithmetic and statements of number are then forged by the Leibniz inference. On this basis, Frege concludes that numbers are objects. Michael Dummett criticizes this argument, suggesting that §56 may be 'the weakest in the whole of *Grundlagen*'.[106] Dummett's point is that Frege assumes that '2' must be recognized as a proper name in both statements of number and statements of pure arithmetic in order to do justice to the content of these statements and to the logical connections between them. There is, however, another option available within Frege's logical framework, a framework that admits higher-order quantification. We can identify individual numbers, so to speak, with second-level concepts, for example, the number 2 with that second-level concept that subsumes just those first-level concepts under which exactly two objects fall. In this way we can retain the analysis of statements of number presented by their quantificational paraphrases,

[106] Michael Dummett, *Frege: Philosophy of Mathematics* (London: Duckworth, 1991), p. 105.

reconstruct the statements of pure arithmetic, and do justice to the logical links between pure arithmetic and statements of number.

To understand better Frege's argument in §§56–7, we need to consider more closely his conception of the application of logic in and to the sciences. In my general exposition of the Context Principle, I said that Frege takes the inference from generalization to instance as basic. This approach does not commit Frege to taking every apparent proper name in German and English – both languages with multifarious grammatical options for nominalization – to be a genuine proper name, every apparent identity statement to be a genuine one. Critical perspective here is, however, internal to logic and its application.[107]

Frege is concerned with language as a means for the public expression of knowledge, of systematic knowledge, of science. He states:

> It is a fundamental principle [*Grundsatz*] of science to minimize the number of axioms [*Urgesetze*]. The essence of explanation consists in the domination of a large, perhaps an unsurveyable, manifold by one or a few sentences.[108]

Frege understands 'domination' in terms of his quantificational conception of generality:

> The value of a law for our knowledge rests on its containing as special cases many, indeed infinitely many, individual facts as particular cases. We profit from knowledge of a law by extracting from it an abundance of pieces of particular knowledge.[109]

Frege's picture is this. Investigators gather knowledge of a domain, including knowledge of general laws. This knowledge is systematized by its axiomatic organization in the framework of Begriffsschrift together with Frege's codification of logic. The regulative principle of axiomatization is economy: the basic proprietary laws of the science should be minimized so as to maximize explanation. To this end, the basic vocabulary of the science should be minimized via definitions. This ideal of system-revealing, explanation-creating

[107] In contrast, Russell appeals to an epistemic notion of acquaintance in applying logic to the sciences. For further discussion of the deep differences between Frege and Russell here, see Peter Hylton, chapter 13, this volume.

[108] 'Booles rechnende Logik', *NS*, p. 40 (36).

[109] 'Logische Allgemeinheit', *NS*, p. 278 (258).

economy extends from particular sciences to the whole of science. Across the sciences, investigators should seek to maximize explanation by means of definition.

Frege takes it to be a part of our logical capacity defeasibly to recognize on the basis of our grasp of thoughts both generalizations and equations. Exercising this capacity, when we begin systematization of a science, we find in its linguistic expression at least a provisional segmentation into proper names and predicates. On Frege's type-theoretic view of logical generality, recognition of higher-level names depends on recognition of lower-level ones. Higher-level primitive predicates proprietary to the science will be recognized only as required on the basis of the antecedent discovery of lower-level predicates.[110] In the interests of economy, we may eliminate as redundant stylistic variants those nominalizations (apparent proper names) whose use is not necessary to express the singular claims of the science in question, whose use does not add to the explanatory power of the evolving systematization. After stylistic redundancies have been pruned back, and we have a body of general laws dominating the particular facts of a science, we have only the surface to go on. If, after such logical reflection, an expression appears to be a proper name, then it is. This reflection is what Frege takes to be fundamental as regards the application of logic in the systematization of a body of knowledge. He has no perspective from which to gainsay its deliverances.[111]

In the case of elementary arithmetic, the deliverances of this logical reflection are unambiguous. The use of numerals as the terms of equations in pure arithmetic marks them as proper names. Applications of arithmetic afford no reason to question this

[110] Of course, predicates that are unstructured with respect to an axiomatization of a branch of science may be defined when foundations for that science are given in terms of some more abstract science. This is what happens to the predicate 'number' in Frege's logicist foundations for elementary arithmetic.

[111] In chapter 4 of *Frege: Philosophy of Language*, 2nd edn (London: Duckworth, 1981), Michael Dummett considers how, among the expressions satisfying semi-formal criteria, spurious proper names might be distinguished from genuine proper names. The factors he mentions – presence of individuation standards for the putative objects, a suitably rich vocabulary of predicates for describing the putative objects, the eliminability of putative proper names without loss of content – will all figure in the application of Begriffsschrift to arrive at a revealingly economical formalization of a branch of science in which the putative proper names figure.

appearance, once we appreciate that they can themselves be paraphrased by equations. This outlook is deeply embedded in Frege's thought. Only long after the paradox, toward the end of his career, does he even tentatively question whether numbers are objects.[112]

It is in this recognition of numerals to be proper names and numbers to be objects that the Context Principle shapes Frege's logicist project. The definition of 'the number one' must take the form of a definition of a proper name; the definition of 'number' will be the definition of a first-level predicate.[113]

8

We still face the task of bringing the logical resources Frege has discovered to bear on an analysis of *number* within the constraints of the logical segmentation of pure and applied arithmetic. *Foundations* §§62–63 introduce Frege's definitional strategy here. An invocation of the Context Principle opens §62:

> How is a number to be given to us, if we can have no mental representation [*Vorstellung*] or intuition [*Anschauung*] of it? Only in the context of a sentence do words mean something. It is then a matter of explaining the sense of sentences in which a number-word occurs. At first this makes things too arbitrary. But we have already determined that what number-words stand for are to be understood as self-subsistent objects, so that we are given one variety of sentence that must have a sense, namely sentences that express an identification [*Wiedererkennen*].[114]

Here, as elsewhere in *Grundlagen*, Frege's formulation of the Context Principle is completely general, applying to all of language. He is not then specifying our cognitive access to numbers as opposed to either spatio-temporal actual (*wirklich*) bodies or geometrical

[112] See especially the 1919 'Aufzeichnungen für Ludwig Darmstaedter', *NS*, p. 277 (257). In the last year and a half of his life, Frege goes back and forth on the issue of whether numerals are genuine proper names. See the diary entry for 23 March 1924, *NS*, p. 282 (261) and 'Zahl', *NS*, p. 284 (265).

[113] The series of definitions of second-level numerical concepts yielded by the quantificational analysis of statements of number are not then the core of the desired analysis of number. I suggest, however, that the logical derivability of the equivalence of the quantificational version of statements of number with the equational version Frege's definitions gives Frege an adequacy test for his definitions.

[114] *Gl*, §62, p. 73.

forms. Rather, Frege is inquiring after constraints on definitions of number, an inquiry which is conceptual, not narrowly epistemic. The logical segmentation of pure and applied arithmetic shows numerals to be proper names, which accordingly appear as terms of equations. Definitions of numerals must then fix a sense for these numerical equations, and do so without use of arithmetical vocabulary. In fixing a sense for numerical equations, the definitions for numerals and for the concept *number* must build in standards of identity, standards of individuation, for numbers. Standards of individuation for numbers will look especially problematic to a Kantian. Empirical objects, actual (*wirklich*) objects – dogs for examples – are individuated in terms of spatio-temporal continuity; their individuation presupposes geometry. No such individuation standards are relevant to numbers.[115] What other individuation standards are there?

The earlier treatment of statements of number showed that the numbers of arithmetic are the numbers that belong to concepts: to be a number is to be the number of a concept. We can then take proper names of the form 'the number of *F*' to provide the fundamental designations of numbers; sentences that express identifications of number will be equations whose terms are proper names of this form. Frege's analytic problem is to reproduce the content of equations involving these proper names. He now calls attention to the standard of equality for numbers (*allgemeines Kennzeichen für die Gleichheit von Zahlen*) given by the Cantor-Hume principle:

> The number of *F* = the number of *G* just in case there are exactly as many *F* as *G*.

Despite verbal appearance, the right-hand side of this standard does not invoke arithmetical notions, for it can be expressed as an existential generalization over first-level relations:

> There is a relation that pairs 1–1 all *F* and all *G*.

[115] This then is the point of Frege's Kantian rhetoric here. In earlier sections of *Gl* Frege has argued that nothing in the conception of number present in pure and applied arithmetic motivates the identification of numbers with empirical or geometrical objects – motivates the identification of the number one with any objects that can be identified in empirical terms (*Vorstellung*) or geometrical terms (*Anschauung*).

Indeed, this existential generalization can be stated in what for Frege are logical terms. The Cantor-Hume principle then fixes in logical terms the content of equations whose terms are of the form 'the number of__'. Furthermore, as the second-level relation *as many as* over concepts specified by the right-hand side of the Cantor-Hume principle is an equivalence relation, the sense the principle gives to these numerical equations will not run afoul of the laws of identity. Might we have here the core of the analysis of *number*?

Frege argues that we do not. On the proposal under consideration, we would be using the right-hand side of the Cantor-Hume principle to stipulate what identity comes to as regards numbers. But the relation of identity is not restricted to numbers: the identity predicate, the notion of identity, is not proprietary to arithmetic. For Frege, it is a part of logic, and so figures in every science. We cannot then separately in logical terms stipulate a sense for numerical equations. As Frege puts the point,

> The relation of identity does not occur only with numbers. From this it appears to follow that it should not get a special explanation for this case. We should have thought that the concept of identity is already firmly in place and together with the concept of number must yield when numbers are identical to each other, without need of a special additional definition.[116]

The problem is that we do not yet have in logical terms a definition of *number.* Frege obscurely proposes

> to form the content of a judgement that can be conceived as an equation, each side of which is a number. We don't want to explain identity just for this case. Rather, by means of the already familiar concept of identity, we want to obtain that which is to be considered as identical. Admittedly, this appears to be a very unusual sort of definition, one that logicians have not yet sufficiently considered.[117]

He goes on in the next three sections to discuss of the parallel attempt to define 'the direction of straight line *a*' via the equivalence:

[116] *Gl*, §63, p. 74.

[117] *Ibid.* As I read the entire passage, in the first quoted remark, Frege criticizes the proposal to use the Cantor-Hume principle to define numbers 'contextually'. The second quoted passage points towards the definition by abstraction that invokes extensions.

The direction of line a = the direction of line b just in case line a is parallel to line b.

Let's switch to Frege's simpler example. We might propose to take this equivalence to give an analysis of *direction* on the grounds that it stipulates a content for sentences of the form 'the direction of line a = the direction of line b'. This content appears to be the content of an equation, because, from it and the fact that the relation *parallel* is an equivalence relation over lines, we can establish such instances of the theory of identity for directions as:

> If (the direction of a = the direction of b), then if (the direction of b = the direction of c), then (the direction of a = the direction of c)
>
> If (the direction of a = the direction of b), then (the direction of b = the direction of a).

However, the reasoning that establishes these statements manipulates the 'directions'-sentences as unstructured wholes on the basis of their stipulated equivalence to sentences about parallel lines. This is as it should be, for from the perspective of Frege's view of logical segmentation our putative stipulation treats sentences of the form 'the direction of line a = the direction of line b' as unstructured wholes. The content our stipulation gives to these directions-sentences has, on its face, nothing in common with the content of sentences like 'The Morning Star is Venus,' or 'The number of martian moons = 2'. In this way, to use the rhetoric of §63, we would have explained identity only for directions, and that is to explain neither *identity* nor *direction*.

Frege brings this point out in §66 by his second use of the Julius Caesar objection. The identity predicate, the first-level relation of identity, is a fixed logical point for Frege's inquiry. To have defined 'the direction of line a' so as to give content to genuine *equations* of the form

> The direction of line a = the direction of line b

is to recognize

> the direction of line $a = \xi$

as a genuine concept-designating predicate, which in turn must yield a true or false sentence when its argument position is filled by any designating proper name, 'Julius Caesar' for example. Our definition, however, gives no content to this predicate. I noted that on Frege's view of logical segmentation definitions must respect logical categories by in effect introducing 'new' names as abbreviations for compound names. The definiendum is then a logically unstructured unit. At best, the proposed definition of direction defines only a simple, unstructured two-place predicate, 'the direction of ξ = the direction of ζ', and so does not yield an analysis of the concept *direction*. Frege's Julius Caesar objection is thus intended only to establish that the Cantor-Hume principle does not give us a definition of *number*, does not give us a definition of genuine equations whose terms are numerals.[118]

Frege's strategy for defining *direction* (and *number*) is to form a judgement that can be conceived as an equation whose terms are 'direction of' names – that is, to find a statement that, by reproducing in other terms the content of 'the direction of *a* = the direction of *b*', sets forth a standard of equality, of individuation, for directions. Then, 'using the already familiar relation of identity, we want to obtain that which is to be regarded as identical'. We have the geometrical equivalence:

> The direction of line *a* = the direction of line *b* just in case line *a* is parallel to line *b*.

[118] This explains why Frege does not raise the Julius Caesar objection in connection with the posit of value-ranges in *Gg*: there is no question but that Basic Law V is an axiom, not a definition of Frege's value-range names.

A second point deserves mention here. Some commentators suggest that the Julius Caesar objection is Frege's reason for rejecting an attractive strategy of contextual definition of numerical contexts in favour of explicit definitions treating numerals as proper names. In pre- and post-*Gl* writings Frege, by example and precept, restricts definitions to explicit definitions of names in the setting of his type-stratified conception of generality. I find no evidence in *Gl* that Frege waivers from this view and seriously considers contextual definitions as an option. Appearances to the contrary represent Frege's efforts informally to inculcate in his audience an appreciation of the constraints his logical segmentation of pure and applied arithmetic discourse imposes on this definitional enterprise.

The Julius Caesar point reappears in Frege's two fundamental principles of definition in *Gg*, vol. II, §§55–67. This discussion is directed at Peano's partial definitions that, like the attempted definition Frege considers in *Gl*, §64, explain an expression only for restricted linguistic contexts. In *Gg*, vol. II, §67, Frege advises that partial definitions that treat the identity-sign as both known and unknown be replaced by definitions by abstraction that exploit Basic Law V.

The right-hand side of this equivalence states in available geometrical terms a standard of equality for directions. When we logically segment the equivalence, we see 'the direction of ξ' as the designation of a function that maps lines to their directions. Directions are the range of this function. Viewing the function as the unknown, our equivalence states a second-level condition a first-level function must satisfy in order to be the *direction of* function. Frege now invokes extensions to define the *direction of* function: the direction of x = the extension of the concept *parallel to x*. As the relation *parallel to* is an equivalence relation, we can prove that the function Frege defines to be the *direction of* function satisfies the second-level condition we extracted from the equivalence. This result certifies the adequacy of the definition. Thus the unusual sort of definition that logicians[119] should pay more attention to is so-called definition by abstraction that define a first-level predicate by identifying a range of objects with equivalence classes obtained from an antecedently recognized equivalence relation over an antecedently recognized range of objects.

In *Grundlagen*, §68, extensions appear abruptly. Frege provides no explanation of them, famously saying in a footnote that he 'presupposes that it is known what the extension of a concept is'.[120] Given

[119] For a discussion of the antecedents of definitions by abstraction in the geometry of Frege's day, see the illuminating discussion in Mark Wilson, 'Frege: the Royal Road from geometry', in William Demopoulos (ed.), *Frege's Philosophy of Mathematics* (Cambridge, Mass.: Harvard University Press, 1995).

[120] *Gl*, §68, p. 80. Frege returns to this point at the end of *Gl*, §107, p. 117, adding that he 'places no decisive weight on bringing in extensions' to remove the 'Julius Caesar' problem. This is an obscure remark, for nothing in Frege's earlier or later surviving writings indicates any alternative to extensions for the analysis of *number.* The only other indication that Frege thought there might be an alternative comes from a telegraphic description Heinrich Scholz gives of a lost manuscript in his catalogue of the Frege Nachlass. See Albert Veraart, 'Geschichte des wissenschaftlichen Nachlasses Gottlob Freges und seiner Edition. Mit einem Katalog des ursprünglichen Bestands der nachgelassenen Schriften Freges', in Matthias Schirn (ed.), *Studien zu Frege*, vol I: *Logik und Philosophie der Mathematik* (Stuttgart: Fromman-Holzboog, 1976), p. 95, the description of item 47. Perhaps Frege in §107 is voicing a worry about his use of the term 'extension'. Frege's conception of extensions, linked as it is to his objective view of concepts, cannot be straightforwardly identified with extant views of extensions. Frege may also be alluding to his realization of the point I make below, that familiarity with extensions is not something that may be presupposed, given the concept–object distinction.

the concept–object distinction, Frege is not entitled to this assumption. Frege tells us that statements of number predicate something of concepts, the objective designata of predicates. Frege's concepts then take over the role that other accounts of *number* assign to collections or groups of things. Thus Frege's exposition of his viewpoint in *Grundlagen* makes the introduction of extensions in addition to concepts appear utterly otiose. The explanations Frege might offer would likely only sow confusion. He might have told us that the extension of concept *F* is the same as the extension of concept *G* just in case concepts *F* and *G* are coextensive. But this sharp, informal statement of Basic Law V only invites the question, 'What is the difference between a concept and its extension?' Any answer here will become obviously and embarrassingly entangled in the awkwardness of language that permeates Frege's explanations of the concept–object distinction – above all, his use in this setting of 'concept' and 'object' as contrasting first-level predicates. Indeed, I think that it is fairly clear that, in *Grundlagen*, Frege is cognizant of this awkwardness, and endeavours there to minimize the expository difficulties it creates.[121]

We have here a genuine problem. On the one hand, I have argued that Frege's view of logical segmentation commits him to recognizing numerals as proper names, numbers as objects. On the other hand, the logical resources for the analysis of *number* Frege has uncovered are at higher levels. These resources cannot be applied to the analytic task at hand within the constraints imposed by the logical segmentation of the statements of pure and applied arithmetic. Now it indeed looks as if extensions do enter on stage as a *deus ex machina* to save the day. The posit of extensions via Basic Law V promises to yield a definition of the concept *number*, definitions of the series of proper names of numbers, '0', '1', '1 + 1', etc., and a proof of the infinity of numbers.

Frege's most extensive pre-paradox discussion of Basic Law V comes in *Grundgesetze*, vol. II, §147. Frege observes there that

[121] I think that this is indicated by Frege's remark in the footnote in *Gl*, §68 that he might have simply used the word 'concept' in place of the phrase 'extension of the concept'. Frege's tone in 'On concept and object' is the tone of a man irked because someone is uncharitably taking his words overly literally and so missing his point.

mathematicians use letters in isolation to express generalizations, as they say, over 'functions':

$$\text{If } f = g, \text{ then } \ldots$$

As function-names cannot occur in sentences in isolation, without a completion for their argument-places, there is no substitution of function-names for these letters to obtain an instance of this generalization. To treat

$$\text{If } (x^2 - 1) = (x + 1)\cdot(x - 1), \text{ then } \ldots$$

as though it were an instance of this generalization exhibits a tacit reliance on the inference codified by Basic Law V. Similarly, the use of the first-level predicate 'function' to express generalizations over functions exhibits the same tacit reliance.[122] Frege accordingly presents Basic Law V as the explicit codification of a basic logical inference implicit in the mathematics that springs from the calculus as well as in logicians' talk of extensions:

> So with this transformation [Basic Law V] we do not really do anything new; but we do it with full consciousness and with appeal to a basic logical law.[123]

How are we to view these points? At the beginning of 'Function and concept,' Frege says:

> A scientific expression first appears with a clear-cut meaning where it is needed for the expression of a lawful regularity. For functions, this case arises with the discovery of higher analysis. There for the first time it is a matter of setting forth laws that hold for functions in general.[124]

I suggest that, as Frege views matters, colloquial language has well-developed resources for expressing singular contents – for saying

[122] In 'Function and concept', p. 10, Frege remarks, 'In many common mathematical turns of phrase, the word 'function' corresponds to what I here have called the value-range of a function. But function, in the sense of the word used here, is the logically prior notion.

[123] *Gg*, vol. II, §147, p. 148. See also the opening paragraph of *Gg*, vol. I, §9, p. 14, where Frege says, 'This possibility [of transforming the generalization of an equation into an equation of value-ranges, and vice versa] must be seen as a logical law, of which incidentally use is always, if tacitly, made when concept extensions are spoken of.'

[124] 'Function and concept', p. 1ff.

that an object falls under a concept. And colloquial language has more or less serviceable resources for expressing generalizations over objects. However, the exigencies of everyday life and science only rarely require generalizations over concepts. Colloquial language is accordingly impoverished in resources to express higher-level generalizations. Higher-level generalizations become salient in mathematics only with the development of analysis, which requires generalization of positions occupied by incomplete function-signs. It is with this development that we need to see

$$x + y = y + x$$

as an instance of $f(x,y) = f(y,x)$.[125] Frege's Begriffsschrift remedies the notational limitations of everyday language as regards higher-order generalizations.

Of course, the vigorous development of analysis did not wait for Frege's Begriffsschrift. Mathematicians used the notational devices mentioned in the previous paragraph to generalize over, as they said, 'functions', thereby appearing, from Frege's vantage point, to treat functions as objects. However, this linguistic usage by itself scarcely warrants the claim that reliance on Basic Law V inference is implicit in mathematics. In explaining his views, Frege treats sentences like

Socrates falls under the concept *mortal*

as a stylistic variant on

Socrates is mortal.[126]

Having accepted Basic Law V as a logical axiom, he treats 'the concept *mortal*' as the designation of an extension. However, apart from the posit of extensions, there is no reason from a Fregean viewpoint to take this step. We might maintain that the predicates 'ξ is mortal' and 'ξ falls under the concept *mortal*' are not only synonymous, but that the second predicate, despite grammatical

[125] A better example from analysis would run into the problems with fitting Leibnizian notation into Begriffsschrift that Frege mentions at the end of 'What is a function?'

[126] For example, see 'Booles rechnende Logik,' *NS*, p. 18 (16); 'Über Schoenflies', *NS*, p. 193 (178); and 'Logik in der Mathematik', p. 231 (214).

appearances, contains no proper name. This redundancy can then be pruned away when we go from everyday language to Begriffsschrift. Both sentences exhibit just the segmentation of a simple singular sentence. Things are less straightforward, when we consider uses of the phrase of the form 'the concept …' outside of the context 'falls under the concept …'. Again, in line with Frege's use of the word 'subordinate' in his exposition, we might argue that the sentence

The concept *whale* is subordinate to the concept *mammal*

is synonymous with

Everything which is a whale is a mammal.

As before, we might further maintain that the first sentence, despite grammatical appearances, contains no genuine proper name. That is, we might maintain that the sentence is, on Frege's principles, logically unsegmented, and so is a misleading expression of the content that is, for this reason, more perspicuously expressed by the second sentence. On this approach, phrases of the form 'the concept …' are to be eliminated in favour of the first-level predicates we have to recognize anyway. Similarly, phrases of the form 'the function that …' can be replaced by the corresponding first-level incomplete expressions. For example, the sentence

The function of addition is commutative

goes over into

$x + y = y + x.$

Frege himself leaves the door open to this option as regards the logical segmentation of colloquial language.

Our colloquial languages [*Volkssprachen*] are not created for the purpose of conducting proofs. The deficiencies that arise from this fact are for me the chief reason to erect a Begriffsschrift. The task of colloquial language is essentially fulfilled if people who interact with each other attach the same thought, or approximately the same thought, to the same sentence. For this, it is not necessary that the individual words by themselves [*für sich*] have a sense and a reference so long as the entire sentence has a sense. Matters are different if inferences are to be drawn. There it is essential that the same expression occurs in two sentences and has in both the same reference. It must then in itself have a reference that is independent of the other parts of the sentence.[127]

Following Mark Wilson, I suggest that Frege in §147 has another raft of considerations in mind to motivate Basic Law V, considerations that he does not sharply separate from the linguistic usages I have just noted. Wilson observes that nineteenth-century mathematicians, as we might put it, expand the ontology of mathematics by characterizing new objects in terms of the behaviour of relations and functions over some domain. The new objects then figure in generalizations alongside the old objects. A paradigm of this procedure is the posit of non-Euclidean points in projective geometry, Frege's branch of mathematics. Indeed, the directions of lines Frege discusses in *Grundlagen*, §§64–68, are the projective geometer's points at infinity. The laws of projective geometry generalize over both Euclidean points and the added ones; the concepts that figure in these generalizations must take both the Euclidean and the non-Euclidean points as arguments. We will not then be able, in projective geometry, to paraphrase away talk of non-Euclidean points in favour of talk of equivalence relations over Euclidean objects.

Throughout his career, Frege is concerned with the introduction of new domains in mathematics, with the 'creation' of new mathematical objects. He vigorously polemicizes against formalist accounts of this practice and aims to develop an alternative to it.

[127] Frege to Peano, 29 September 1896, *WB*, p. 183. In this passage, Frege is discussing both predicates like 'heap' that, lacking sharp boundaries, are logically speaking meaningless and Peano's incomplete definitions of predicates. Frege maintains that these definitions do not confer a sense and reference on the definiens, but only on a range of larger expressions in which the definiens occurs, but not as a logically segmented part.

I should observe that is not clear how to extend this approach to generalizations of claims make using nominalized predicates and function signs. What is the relationship of the sentences

> If one concept is subordinate to a second, and the second subordinate to a third, then the first is subordinate to the third,

and

$$(Fx \to Gx) \to [(Gx \to Hx) \to (Fx \to Hx)]?$$

One might take the first sentence to be a logically unstructured expression of the second-level generalization that the second sentence more perspicuously expresses. Alternatively, one might reject the first sentence as a pseudo-sentence

Frege's own approach here shines forth in a comment on Dedekind's account of the real numbers.

> The most important thing for an arithmetician who recognizes in general the possibility of creation [of mathematical objects] will be to develop in an illuminating way [*in einleuchtender Weise*] the laws governing this in order to prove in advance of each individual creative act that the laws allow it. Otherwise, everything will be imprecise, and proofs will degenerate to a mere appearance, to a good-willed self-delusion.[128]

The desired foundation will be provided by formulating a logical law that, in the context of other logical laws, will yield as a theorem the existence of the desired new objects. Basic Law V is Frege's proposed law that affords a foundation, a codification, for mathematical practice here. It is in this way that Frege claims with Basic Law V not to have introduced anything new.

I began with a depiction of Frege's stratified view of quantificational generality rooted in the Context Principle. The quantificational patterns that pervade colloquial language presentations of mathematics and the mathematical sciences appear to be more or less capturable within a second-order setting.[129] This appearance is abetted by the use of nominalizations of incomplete expressions. This manoeuvre yields more than stylistic variants of essentially higher-level statements. This, I suggest, is the lesson Frege extracts from the mathematical practice exhibited in projective geometry. Basic Law V in the context of impredicative second-order logic is Frege's attempt to disentangle matters so as to provide a unified foundation for all of arithmetic. He is well aware of the iterative possibilities of Basic Law V within his formulation of logic. He takes these possibilities to offer the prospect of establishing within logic the ontology of mathematics.

that does not express a thought at all precisely because it arises from treating nominalizations of predicates as though they were genuine proper names.

[128] *Gg*, vol. II, §140, p. 142.

[129] Frege's view of the quantifiers as themselves higher-level concepts prompts the introduction into logic of second- and third-level predicates. Moreover, the formulation of Basic Law Iib in connection with quantification over first-level functions requires the use of a free variable over third-level functions. However, these upward forays are auxiliaries to Frege's formulation of second-order logic within the setting of his universalist conception of logic. In contrast to Russell after the adoption of the theory of types, Frege shows no inclination to identify logic with a type-stratified theory of entities organized into a countable hierarchy.

Thus we hope to be able to develop the wealth of objects and functions that mathematics treats from the eight functions enumerated in vol. I, §31, as it were from a seed.[130]

It was not to be. With his conception of foundations as notationally exact axiomatic systematization, and the depth and coherence of his view of logic and logical segmentation, Frege paints himself into a corner as regards logicist foundations for arithmetic.[131]

[130] *Gg*, vol. II, §147, p. 149.

[131] I am indebted to Michael Kremer for comments on an earlier version of this chapter and to Michael Friedman, Juliet Floyd, Warren Goldfarb, Peter Hylton, Peter Sullivan, Judson Webb and Mark Wilson for extensive discussions on the topics of the chapter. An earlier version of material in the last two sections was presented at an Arché Abstraction Weekend at the University of St Andrews.

MICHAEL KREMER

7 Sense and reference: the origins and development of the distinction

Frege's distinction between sense (*Sinn*) and reference (*Bedeutung*) has been his most influential contribution to philosophy, however central it was to his own projects, and however he may have conceived its importance. Philosophers of language influenced by, or reacting against, the distinction and historians of philosophy commenting on it, have all contributed to the voluminous literature surrounding it.[1] Nonetheless in this essay I hope to shed new light on the distinction by considering it in the context of the development of Frege's thought, and connecting it more intimately than is usually done with Frege's interests in logic, especially his views on judgement, truth and inference, which were central to his own projects as he conceived them.

Frege does not employ the terminology of sense and reference in his first great logical-philosophical work, the *Begriffsschrift* of 1879 (*Bs*). However, *Bs* already contains the seeds of the distinction in its notion of 'content' (*Inhalt*). Tracing out the difficulties inherent in

[1] I avoid detailed discussions of the secondary literature. A few references will be given in footnotes, but it will be obvious to many that my debts are far more wide-ranging than can be acknowledged here. The most important overall influences on my interpretation are Brandom, Dummett, McDowell, Ricketts, Sluga and Weiner. Burge on truth, Mendelsohn on identity and Taschek on sense are each of central importance at specific points. A detailed and rich recent reading of Frege, with many points of contact with my interpretation, is that of Michael Beaney. Since the composition of the main argument of this essay, several relevant works have appeared of which it has not been possible to take account.

I first studied Frege under Bob Brandom at the University of Pittsburgh. I sometimes think that everything I say about Frege is a dim recollection of something Bob said in a lecture. I am also in the debt of the many students who have tolerated my spinning these tales in my own lectures. Finally, thanks are due to Jim Conant, Mike Beaney, Gottfried Gabriel, Ed Zalta, Marian David and especially Tom Ricketts for helpful comments and discussion.

Frege's early talk of 'content' illuminates the need for this distinction, as well as his further difficulties in formulating it. *Bs* contains two distinct, yet interrelated, ancestors of the sense–reference distinction. Section 1 discusses the first root of the distinction, which lies in Frege's notion of *judgeable content*, expressed by sentences. The second root lies in his account of identity sentences, and the associated idea of 'modes of determination' of a content. Section 2 explores this account in detail, and reveals some of the difficulties inherent in it. Sections 3 and 4 show how the needs of Frege's project in the philosophy of mathematics brought these difficulties to the fore and led to the development of the mature sense–reference distinction. Sections 5 to 8 expound Frege's mature vision; section 9 then examines some of the remaining difficulties in the light of the development of the distinction.

1 BEGRIFFSSCHRIFT: JUDGEABLE CONTENT

In the preface to *Bs*, Frege set as his goal to determine the epistemological status of arithmetical truths. This requires investigating whether they can be proved on the basis of logical laws alone, or need some other source of support, such as Kantian pure intuition. To this end, he constructed a new logical system in which proofs could be carried out without 'gaps', so as to display explicitly all presuppositions and assumptions employed. His logical notation was to express only that which is relevant to inference, which he called 'conceptual content' (*begrifflicher Inhalt*).[2] In §3 of *Bs*, Frege explains this notion through an example: 'At Plataea the Greeks defeated the Persians' and 'At Plataea the Persians were defeated by the Greeks' have the same conceptual content, since the same *consequences* follow from each, in conjunction with any set of additional premises one might consider (*Bs*, §3, pp. 112–3). Frege's example suggests a more general principle. Letting '⇒' represent an as yet unspecified relation of consequence:

(CONTENT): Sentences A and B have the same content if and only if, for any set of sentences (auxiliary premises) S and sentence (conclusion) C: $S,A \Rightarrow C$ if and only if $S,B \Rightarrow C$.

[2] *Bs*, p. 104. Citations of Frege's published work refer to section and page numbers in the standard English translations.

It may seem implausible to attribute such a general principle to Frege on the basis of an example used in an informal presentation. However, Frege is simply adapting a piece of traditional logical wisdom to his own purposes: the distinction between the extension and comprehension of an idea or concept. This goes back at least to the Port-Royal *Logic* (1662) and arguably much further, to Porphyry's *Isagoge* and medieval commentators on it. For the Port-Royalists, 'the comprehension of an idea' consists in 'the attributes that it *contains* in itself',[3] in the sense in which the idea of *human* might be said to 'contain' the idea of *animal*. Ideas are pictured as arranged in a hierarchy, with some ideas containing others, higher in the hierarchy or tree. Logical relations of implication (etymologically, 'folding-in') are relations of containment of ideas – *human* contains *animal*; the first idea implies the second. On the other hand, ideas not only contain, but are contained in, other ideas. At the bottom of the tree we find the entities to which the ideas apply; these constitute the *extension* of the ideas. 'I call the *extension* of an idea the subjects to which this idea applies' (*Logic*, 40).

In his *Logic*,[4] a work which Frege read, Kant applied this distinction to *concepts*: 'Every concept, *as partial concept*, is contained in the representation of things; as *ground of cognition, i.e., as mark*, these things are contained *under* it. In the former respect every concept has a *content* (*Inhalt*), in the other an *extension* [*Umfang*]' (*LL*, p. 593). Kant states a principle of the inverse proportionality of extension and content: 'The content and extension of a concept stand in inverse relation to one another. The more a concept contains *under* itself, namely, the less it contains *in* itself, and conversely' (*LL*, p. 593). In early works, Frege appeals to this principle to argue that a predicate which applied to all objects would have

[3] Arnaud, A. and Nicole, P., *Logic or the Art of Thinking*, trans. J. Vance Buroker (Cambridge: Cambridge University Press, 1996), p. 39. My emphasis.

[4] Immanuel Kant, *Logik*, in *Kant's gesammelte Schriften*, ed. J. B. Jäsche, vol. XI (Berlin and Leipzig: de Gruyter, 1923); translated in *Lectures on Logic*, trans. J. Michael Young (Cambridge: Cambridge University Press, 1992), p. 39. Hereafter *LL*. The so-called *Jäsche Logik* of 1800, often held to be of dubious value in interpreting Kant, was assembled out of Kant's lecture notes by Benjamin Jäsche, with Kant's approval. It was known in the nineteenth century as 'Kant's Logic' and included in Kant's collected works. Frege cites it as representing Kant's views. I follow Frege here, since my intention is not to interpret Kant but to illuminate the development of Frege's thought.

maximal extension, and so no content;[5] another use of the principle is in 'Boole's logical calculus and the concept-script', *NS*, p. 16, fn **/ *PW*, p. 15, fn **).

Kant extended the distinction between content and extension from concepts to cognition in general, including judgements, speaking of the content of a cognition as a matter of its 'richness', 'logical importance' and 'fruitfulness' as the 'ground of many and great consequences', and of the extension as a matter of the 'horizon' of the cognition, the area within which it applied. (*LL*, pp. 549–50). In *Bs*, Frege adapts this notion of content to the case of sentences; content is a matter of what is *implied* by or *contained* in a given claim.[6]

Kant's talk of 'judgement' obscured a distinction that Frege carefully marked, between the *act* of judg*ing*, and the *content* which is judg*ed*, 'judgeable content' (*beurtheilbarer Inhalt*) (*Bs*, §2, p. 112):

A judgment will always be expressed with the aid of the symbol

|––

which stands to the left of the symbol or combination of symbols giving the content of the judgement. If we omit the small vertical stroke at the left end of the horizontal one, then the judgement is to be transformed into

[5] 'Dialogue with Pünjer on existence', *Nachgelassene Schriften*, ed. H. Hermes *et al.* (Hamburg: Felix Meiner, 1983), p. 71. Hereafter referred to as *NS*, with references to the English translations in *PW*.

[6] The importance of Frege's early notion of conceptual content was first made clear to me in Bob Brandom's 1982 Frege seminar. Beaney also emphasizes the *Bs* account of sameness of content. However, he wrongly claims this to be a major innovation on Frege's part (Michael Beaney, *Frege: Making Sense* (London: Duckworth, 1996), pp. 56–64 (hereafter *FMS*)). As we have seen, the fundamental idea is already present in Kant's *Logic*.

Dummett questions the value of taking seriously Frege's introduction of 'content' in *Bs*, arguing that it is 'superfluous to credit him with some rival theory' on the 'exceedingly thin' basis of *Bs*, §3 (Michael Dummett, *The Interpretation of Frege's Philosophy* (Cambridge, Mass.: Harvard University Press, 1981), pp. 298, 301). Hereafter *IFP*. However, a number of otherwise mysterious claims in Frege's early works become clear when 'content' is understood in the manner sketched here: (1) Frege's appeals to the inverse proportionality principle; (2) his claim, discussed below, that the axioms of the *Bs* 'have enough content' since they are 'adequate to the task' of proofs; and (3) his argument that his choice of logical primitives is superior to Boole's, since his primitives have a *simpler content*. He claims that 'the simpler a content is, the less it says' and that the material conditional is simpler than Boolean identity (material biconditional), conjunction, or exclusive disjunction. He sees Jevons's replacement of exclusive with inclusive disjunction as an 'improvement' because it 'diminishes the content of the sign'. Similarly, he speaks of adding 'an unnecessary condition to a judgement' as a *mere combination of representations* [*blosse Vorstellungsverbindung*] of which the writer does not state whether he acknowledges its truth.

Shortly after *Bs* Frege writes 'Through this mode of notation I meant to have a very clear distinction between the act of judging and the formation of a mere judgeable content.'[7] This suggests that we 'form' a judgeable content by combining 'representations', then judge by taking it to be true or false. While Frege addresses the distinction between content judg*ed* and act of judg*ing*, he is less clear here on the general 'ing–ed' ambiguity, between represent*ed* content and act of represent*ing*, in the notion of '*representation*'. Thus in *Bs* it is unclear what the precise status of the 'representations' which are to be combined into judgeable contents might be – sometimes they seem to be representations in the sense of psychological acts, sometimes in the sense of the objects of those acts.[8]

Frege's early account of conceptual content has often been taken to imply:

> (A) any two 'logical truths', or 'analytic truths', have the same conceptual content.

Such a result would be clearly undesirable, given Frege's logicist thesis that the truths of arithmetic are analytic, derivable from the basic laws of logic together with definitions. Given (A), his logicism would imply that the truths of arithmetic have the same content as the most trivial truths of logic, and that there is at most a psychological, but no logical, difference between judging that every natural number has a unique prime factorization and judging that the Moon is the Moon. Yet part of the point of Frege's logicism was to argue against Kant that analytic truths can extend our knowledge and so have genuine and distinctive content.

resulting in a 'diminution of content' ('Boole's logical calculus and the concept-script', *NS*, pp. 40–1, 43/*PW*, pp. 36, 38). All of this makes sense if we think of content in terms of consequences – the fewer the consequences, the less the content.

7 'On the aim of the conceptual notation', p. 94.

8 The development of Frege's conception of judgement, and its relationship to Kant's, is discussed further in M. Kremer, 'Judgment and truth in Frege', *Journal of the History of Philosophy*, 38 (2000), pp. 549–81.

(A) is commonly taken to follow from another consequence of (CONTENT):

> (B) two sentences A and B have the same content just in case they mutually entail each other: $A \Rightarrow B$ and $B \Rightarrow A$.

(B) is often thought to be a direct consequence of (CONTENT), but any *argument* from (CONTENT) to (B) must appeal to characteristics of the consequence relation $\Rightarrow$, and it is useful to make these explicit. One can argue that (B) follows from (CONTENT) using three general principles about consequence derivable from the natural deduction and sequent-calculus systems of logic devised by Gerhard Gentzen in 1935:[9]

(1) *identity*: every sentence is a consequence of itself ($A \Rightarrow A$).

(2) *weakening*: superfluous premises do not invalidate an inference. If a sentence A is a consequence of the set S ($S \Rightarrow A$), and T is a larger set than S ($S \subseteq T$), then A is a consequence of T ($T \Rightarrow A$).

(3) *cut*: if A is a consequence of S ($S \Rightarrow A$), and B is a consequence of S together with A ($S,A \Rightarrow B$), the 'lemma' A may be 'cut', and B is a consequence of S alone ($S \Rightarrow B$).

From these principles, together with (CONTENT), (B) follows.[10] On the one hand, suppose that A and B have the same content. Then, as $A \Rightarrow A$ (identity), $B \Rightarrow A$ (by (CONTENT), since B has all the consequences of A), and as $B \Rightarrow B$, also $A \Rightarrow B$ – that is, A and B are mutual consequences. On the other hand, suppose that A and B are mutual consequences, and suppose that $S,A \Rightarrow C$; then $S,A,B \Rightarrow C$ (weakening),

[9] Gerhard Gentzen, *The Collected Papers of Gerhard Gentzen*, ed. M. E. Szabo (Amsterdam: North-Holland, 1969), pp. 83–4. Gentzen took consequence to be a relation between *sequences* of premises and conclusions, allowing 'multiple conclusions'. Here, the premises are taken as a set, and there is only one conclusion of any inference.

[10] Beaney, *FMS*, p. 57, and Richard Mendelsohn, 'Frege's *Begriffsschrift* theory of identity', *Journal of the History of Philosophy*, 20 (1982), pp. 279–99, here p. 287, argue that (B) follows from (CONTENT) and identity; but from that principle, we can only conclude that sentences with the same content imply each other, but *not* conversely. Beaney thinks that (A) then follows from (B) (*FMS*, p. 63), but this requires further assumptions about logical truth.

and since $B \Rightarrow A$, we have that $S,B \Rightarrow A$ (weakening). Hence $S,B \Rightarrow C$ (cut). Similarly, if $S,B \Rightarrow C$, then $S,A \Rightarrow C$; so A and B have the same content (by (CONTENT)).[11]

Frege does not provide an explicit theory of consequence in *Bs*, so we can't be sure that he would have accepted all of Gentzen's principles. But, granting for the sake of argument that he would have, and so that he was implicitly committed to (B), we cannot infer a commitment to (A) without further ado. *Given* a conception of logical truth as that which is a consequence of the empty set of premises ($\varnothing \Rightarrow A$), we can argue from (B) to (A): if A and B are logical truths, then $\varnothing \Rightarrow A$ and $\varnothing \Rightarrow B$, so $A \Rightarrow B$ and $B \Rightarrow A$ by weakening, and so they have the same content by principle (B). However, it is doubtful that Frege would have accepted such an explication of logical truth. It is not clear what sense he could have made of a sentence being a 'consequence of the empty set'. For Frege, consequence, following from, is a *relation* between judgeable contents which enables one judgement to be *justified* on the *basis* of *others*. He did not have Gentzen's notions of inference from an assumption and the discharging of assumptions, which could support a conception of a 'proof' or 'reasoning' without premises. His systems of proof, both in *Bs* and in the later *Grundgesetze* (*Gg*), are devoid of rules such as conditional proof or *reductio ad absurdum* which rely on reasoning from assumptions. He represents apparent occurrences of such reasoning in mathematical practice not as involving assumptions, but instead explicitly asserted hypothetical sentences with the seeming assumptions as antecedents.

Moreover, it is a consequence of this Gentzen-style conception of logical truth, together with Gentzen's principle of cut, that logical truths can always be 'cut':

(C) if $S,A \Rightarrow B$ and A is a logical truth, then $S \Rightarrow B$.[12]

It is doubtful that Frege would have accepted (C). One might argue for (C) as follows: logical laws collect and make explicit patterns of inference linking premises and conclusions in valid reasoning. If we can infer B from premises S together with logical law A, A's only

[11] Neil Tennant, 'Frege's Content-Principle and relevant deducibility', *Journal of Philosophical Logic*, 32 (2003), pp. 245–58, presents a similar argument. He also considers what happens to (CONTENT) in the context of a form of relevance logic in which Gentzen's principle of cut does not hold unrestrictedly.

[12] One could also get the problematic result (A) from this principle directly.

function must be to exhibit the link between *S* and *B*. But this link must be there anyway, and so *B* must be a consequence of *S* without need of *A*. But this argument would imply that logical truths have no content and can in no way extend our knowledge, and this Frege must reject.

Frege might have accepted this argument for (C) in the case where the premises in *S* and the conclusion *B* are *not* logical laws. But Frege's account of the argumentative structure of *Bs* argues against his acceptance of (C) when *S* and *B* consist of logical laws. In *Bs*, he sets out to prove various logical laws from a set of *basic* logical principles.[13] He writes (*Bs*, §13, p. 136):

> It seems natural to deduce the more complex of these judgements from the simpler ones – not to make them more certain, which would in general be unnecessary, but to bring out the relations of the judgements to one another. Merely knowing the laws is obviously not the same as also understanding how some are implicitly contained in others. In this way we obtain a small number of laws in which (if we add the laws contained in the rules) is contained, though in embryonic form, the content of all of them. And it is an advantage of the deductive mode of presentation that it teaches us to recognize this kernel. Because we cannot enumerate all of the boundless number of laws that can be established, we can attain completeness only by a search for those which, *potentially*, imply all the others.

This passage makes little sense if any logical law follows from any other. For Frege, clearly, not all logical laws are on a deductive par.

In an unpublished paper, 'Boole's logical calculus and the concept-script,' Frege explains his choice of axioms for *Bs*: he only 'assumed such as appeared necessary for the proof'of the final proposition (133) of the book. He adds: 'that my sentences have enough content … follows from the fact that they were adequate to the task' – that they sufficed for the proof of (133) (*NS*, p. 43/*PW*, p. 38). Had he omitted some of his axioms, and been unable to carry out the proof, the resulting collection of axioms would *not* have had 'enough' content.

[13] This aspect of Frege's conception of logic is emphasized in Michael Detlefsen, 'Fregean hierarchies and mathematical explanation', *International Studies in the Philosophy of Science*, 3 (1988), pp. 97–116, and in Tyler Burge, 'Frege on knowing the foundations', *Mind*, 107 (1998), pp. 305–47.

Frege comments that, in carrying out his proofs, he had, at times, to employ 'diminution[s] in content' through the addition of 'superfluous conditions' as 'necessary transition points' (*NS*, p. 43/*PW*, p. 38). He has in mind cases like this: aiming to prove C, one proves $(A \supset B) \supset C$ and B; one then 'diminishes' the content of B by using an instance of axiom (1) of *BS*, $B \supset (A \supset B)$, to add the 'superfluous condition' A, arriving at $A \supset B$, from which one can infer C.[14] Even if both B and $A \supset B$ are logical laws, for Frege the latter has 'diminished' content compared to the former. Not everything which follows from B also follows from $A \supset B$; notably, B itself does not. In a similar vein, Frege says that he 'had to assume formulae which merely express the different ways in which you may alter the order of a number of conditions. Instead of giving a general rule that conditions may be ordered at random, I only introduced a much weaker axiom that two conditions may be interchanged, and then derived from this the permissibility of other transpositions' (*NS*, p. 43/*PW*, p. 38). He has in mind here axiom (8), $(d \supset (b \supset a)) \supset (b \supset (d \supset a))$, from which he derived a series of theorems, such as (12), $(d \supset (c \supset (b \supset a))) \supset (d \supset (b \supset (c \supset a)))$, about which he comments: 'Propositions (12)–(17) and (22) show how, when several conditions are present, their order can be altered' (*Bs*, §16, p. 148). It would have been simpler to introduce a general rule to this effect, as in his later *Gg*,[15] but that would have been to choose a stronger principle than necessary, *given* the other axioms of *Bs*. Given those axioms, (8) is sufficient, although from (8) alone one cannot prove (12) – axioms (1) and (2) are needed as well.

For Frege there is an ordering of the logical laws themselves, which the system of *Bs* lays bare. Frege would not admit that the first axiom of *Bs*, proposition (1), and its last theorem, proposition (133), are mutual consequences. (133) does not follow from (1) alone; the proof of (133) requires other axioms of *Bs* as well. Such a proof reveals the logical interconnections of the propositions of *Bs*, which

[14] A similar, but slightly more complicated case, occurs in the first sequence of proofs in *Bs*, culminating in (5), $(b \supset a) \supset ((c \supset b) \supset (c \supset a))$. To prove this, Frege begins with axiom (2), $(c \supset (b \supset a)) \supset ((c \supset b) \supset (c \supset a))$. He uses an instance of axiom (1) to 'weaken' this to (3), $(b \supset a) \supset [(c \supset (b \supset a)) \supset ((c \supset b) \supset (c \supset a))]$, which he combines with an instance of (2) to obtain (4), $[(b \supset a) \supset (c \supset (b \supset a))] \supset [(b \supset a) \supset ((c \supset b) \supset (c \supset a))]$. (5) follows from (4) and an instance of (1). Here the 'diminution of content' occurs as a 'necessary transition point' in the move from (2) to (3). (Here and elsewhere I modernize Frege's notation silently.)

[15] See the second rule of *Gg*, 'interchange of subcomponents' (*Gg*, vol. I, §48, p. 61).

Frege records in a table at the end of the work, indicating which propositions are used in the proofs of subsequent propositions. 'Consequence' is already for Frege a notion with epistemological import. Deducing consequences from basic logical laws is a process which generates content, insofar as the conclusions we deduce are contained in the basic laws collectively, but not individually. Later, when he was sure that he had proved that the laws of arithmetic are analytic truths, deducible from the basic laws of logic and definitions, Frege writes: 'Can it be said that the sentence '$3+7=10$' is deduced from the sentence '$2^2=4$'? Hardly. Is '$3+7=10$' a consequence of '$2^2=4$'? Apparently not ...'[16] Here again, he rejects the view that every logical or analytic truth is a consequence of every other. Thus, the supposed difficulty for his *Bs* conception of content, posed by principle (B) is void.[17] For although Frege might plausibly be held to accept principle (B), he clearly would have rejected the sorts of additional principles necessary to move from (B) to (A).

None of this, however, is to say that he had a worked out theory of the consequence relation. Rather, he relied on a working appreciation of the ways in which we count sentences as following from, or not following from, others. As a working mathematician, Frege knew that it was inappropriate to treat proposition (133) of *Bs* as following from proposition (1) – if he had offered as 'proof' of (133) the 'reasoning' '(1), therefore (133)' he would have failed in his goal of establishing that 'pure thought ... is able, all by itself, to produce from the content which arises from its own nature judgements which at first glance seem to be possible only on the grounds of some intuition' (*Bs* §23, p. 167) or, as he put it in 1884, commenting on the proof of (133), 'sentences which extend our knowledge can contain analytic judgments.'[18]

Nonetheless, Frege's practical sense of what counts as a consequence of what, and so of which sentences count as the same in content, presents serious difficulties in connection with providing any general account of the notions of logical law and logical consequence, and so, also, of conceptual content. It is natural to suppose that the logical laws are the sentences provable in *Bs* (or some

[16] 'On Mr Peano's conceptual notation and my own' (1897), p. 243.

[17] Hence there is no need to introduce, in addition to Frege's 'conceptual content', Beaney's epistemologically motivated notion of 'cognitive content' (*FMS*, p. 64).

[18] *Gl*, §91, p. 104.

suitably expanded version of it), and that *B* is a consequence of *A* if and only if the conditional $A \supset B$ is a logical law. This would imply, however, that any two logical laws are consequences of one another, and so have the same content. Some more refined notion of logical consequence is needed, perhaps agreeing with that given here when neither *A* nor *B* are logical laws. But, in *Bs*, Frege provided no such account. Still, he seems to have been aware that there was a problem here. When he claimed that his 'sentences have enough content' because they are 'adequate to the task' of proof, he added a *caveat*: 'in so far as you can talk of the content of sentences of pure logic at all' ('Boole's logical calculus and the concept-script', *NS*, p. 43/*PW*, p. 38). This hesitancy concerning talk of the content of 'sentences of pure logic' was overcome by his desire to show that *content* – 'richness', 'logical importance' and 'fruitfulness' as the 'ground of many and great consequences' – could arise from these sentences.[19]

Bs's notion of judgeable content is one root of the sense–reference distinction; more precisely, it is the ancestor of the later notion of

[19] My account of Frege's notion of content might seem to depend on a non-Fregean conception of logic, since it involves a consequence relation that can hold even when premises and conclusion are not truths. For Frege's considered view took valid inference to move from true premises to true conclusions. ('Foundations of geometry: second series III', pp. 336–7.)

However, it can be argued that Frege only came to such a conception of logic *after* drawing the sense–reference distinction. In some notes on Lotze's *Logik*, which Franz Hovens convincingly dates to the early 1880s (Frans Hovens, 'Lotze and Frege: The dating of the "Kernsätze"', *History and Philosophy of Logic*, 18 (1997), pp. 17–31), Frege wrote that 'the task of logic is to set up laws according to which a judgement is justified by others, irrespective of whether these are themselves true' ('17 Key Sentences on Logic,' *NS*, p. 190/*PW*, p. 175). In contrast, shortly after 'On sense and reference', he criticized 'content logicians' (*Inhaltslogiker*) for forgetting 'that logic is not concerned with how thoughts, regardless of truth-value, follow from other thoughts' ('Comments on sense and reference' (1892–5), *NS*, p. 133/*PW*, p. 122). It seems that it was the sense–reference distinction that forced on Frege the question of whether logic is primarily concerned with the realm of reference or the realm of sense. It is a further issue why he chose to answer this question by asserting that 'the laws of logic are first and foremost laws in the realm of reference and relate only indirectly to sense' (*NS*, p. 133/*PW*, p. 122). This is not the place for a developed answer to this question, but I will offer the following speculative suggestion: if logic is to yield us knowledge of *objects*, such as the numbers, it must operate at the level of reference. However this may be, even in his later period Frege surely recognized that thoughts stand in relations such that, were certain thoughts to be true, other thoughts would have to be true – relations specifying 'how thoughts, regardless of truth-value, follow from other thoughts'. His later view was simply that such relations are not the concern of *logic*.

the *thought*, the *sense* of a sentence. Much later in his career, in 1906, he returned to the issues of the proper characterization of consequence, and so also of the individuation of content, or as he now put it, the individuation of thoughts.

Frege's attempt to characterize consequence occurs at the end of his controversy with Hilbert over the foundations of geometry. Having rejected Hilbert's approach to proving the independence of the axioms of geometry, he tries to give a proper account of independence, and so also 'dependence' ('On the foundations of geometry: second series, III', p. 334):

> Let Ω be a group of true thoughts. Let a thought *G* follow from one or several of the thoughts in this group by means of a logical inference such that apart from the laws of logic, no sentence not belonging to Ω is used. Let us now form a new group of thoughts by adding this thought *G* to the group Ω. Call what we have just performed a logical step. Now if through such a sequence of steps, where every step takes the result of the preceding one as its basis, we can reach a group of thoughts that contains the thought *A*, then we call *A* dependent upon group Ω.

Frege limits the scope of this account, however ('On the foundations of geometry', p. 335):

> In taking a logical step from the thought-group Ω, we are applying a logical law. The latter is not to be counted among the premises and therefore need not occur in Ω. Thus there are certain thoughts, namely the laws of logic, which are not to be considered when dealing with questions concerning the dependence of a thought.

Here, he seems aware that the proposed account of consequence, if applied to logical laws, would make logical laws dependent upon one another, and takes pains to block this result.

In the same year, Frege twice took up the question of the individuation of thoughts, once in a letter to Husserl of 9 December 1906, and once in an unpublished manuscript, 'A brief survey of my logical doctrines'. The explanations he provides are similar in interesting ways, but also differ in important respects. He writes to Husserl:

> It seems to me that an objective criterion is necessary for recognizing a thought again as the same, for without it logical analysis is impossible. Now it seems to me that the only possible means of deciding whether sentence *A* expresses the same thought as sentence *B* is the following,

and here I assume that neither of the sentences contains a logically self-evident component part in its sense. If *both* the assumption that the content of *A* is false and that of *B* true *and* the assumption that the content of *A* is true and that of *B* false lead to a logical contradiction, and if this can be established without knowing whether the content of *A* or *B* is true or false, and without requiring other than purely logical laws for this purpose, then nothing can belong to the content of *A* as far as it is capable of being judged true or false, which does not also belong to the content of *B*; for there would be no reason at all for any surplus in the content of *B*, and according to the presupposition above, such a surplus would not be logically self-evident either. In the same way, given our supposition, nothing can belong to the content of *B*, insofar as it is capable of being judged true or false, except what also belongs to the content of *A*. Thus what is capable of being judged true or false in the contents of *A* and *B* is identical, and this alone is of concern to logic, and this is what I call the thought expressed by both *A* and *B*.[20]

In 'A brief survey,' on the other hand, he explains (*NS*, pp. 213–4/*PW*, p. 197):

Now two sentences *A* and *B* can stand in such a relation that anyone who recognizes the content of *A* as true must thereby also recognize the content of *B* as true and, conversely, that anyone who accepts the content of *B* must straightaway [*ohne weiteres*] accept that of *A*. (*Equipollence*). It is here being assumed that there is no difficulty in grasping the content of *A* and *B* … I assume that there is nothing in the content of either of the two equipollent sentences *A* and *B* that would have to be immediately accepted as true by anyone who had grasped it properly … So one has to separate off from the content of a sentence that part that alone can be accepted as true or rejected as false. I call this part the thought expressed by the sentence. It is the same in equipollent sentences of the kind given above. It is only with this part of the content that logic is concerned.

Both these explanations share with the *Bs* account of conceptual content a concern to isolate that *part* of a sentence's content which is 'of concern to logic'. Further, both explanations attempt to isolate this part through a sort of mutual consequence test – *Bs*'s account ultimately reduces to such a test, assuming Gentzen's principles

[20] *Wissenschaftlicher Briefwechsel*, ed. G. Gabriel *et al.* (Hamburg: Felix Meiner, 1976), pp. 105–6; translated as *Philosophical and Mathematical Correspondence*, trans. H. Kaal (Chicago: University of Chicago Press, 1980), pp. 70–1. Hereafter *WB*/*PMC*.

governing consequence. In the letter to Husserl, Frege spells this out in terms of a sort of *reductio ad absurdum* procedure, whereas in 'A brief survey' his approach is more direct: anyone who accepts *A* ought also to accept *B* (reading 'must' (*müssen*) here as having *normative* force). The most significant difference between the two accounts, though, comes with the word(s) 'straightaway' (*ohne weiteres*) in 'A brief survey' – this suggests that, in order to express the same thought, *A* and *B* must be mutual *immediate* consequences, whereas the procedure outlined in the letter to Husserl would count as equipollent sentences which are not *obviously* mutual consequences.

But most significant of all is the fact that each account explicitly omits from consideration sentences containing 'a logically self-evident component part', or something 'that would have to be immediately accepted as true by anyone who had grasped it properly'. Frege again seems aware that, without such a restriction, his explanations would entail that any two logical laws express the same thought, and takes steps to ward off this conclusion. Thus, in his attempts to characterize logical consequence, and to individuate thoughts, Frege in 1906 displays a concern with issues that have troubled us in our discussion of his 1879 account of judgeable content.

2 *BEGRIFFSSCHRIFT*: SUB-SENTENTIAL CONTENT

Frege's *Bs* account of conceptual content led him to reject the traditional distinction between subject and predicate as irrelevant to content. His initial example of sentences with the same content showed that the same thing can appear as subject or as predicate without changing the content. Frege replaced this traditional analysis of judgements with an analysis based on the mathematical notions of function and argument. In the Preface to *Bs*, he says that 'it is easy to see how regarding a content as a function of an argument leads to the formation of concepts' (*Bs*, p. 107). However, in *Bs* itself, the notions of function and argument are explained for linguistic expressions, rather than contents. If we start with a complex expression, with or without a judgeable content, we can view a part of this expression as replaceable by other expressions. This yields an analysis of the whole expression into a part which is held constant – the function – and a part which is left to vary – the argument

(*Bs*, §9, pp. 126–7). Two points are crucial about this explanation. First, there is more than one way to analyse a complex expression into function and argument. For example, the sentence 'Cato killed Cato' can be analysed into the argument 'Cato' and the functions '( ) killed Cato', 'Cato killed ( )' or '( ) killed ( )' (among others). Second, functions, unlike the expressions from which they are formed, are 'incomplete' – they have argument-places which need to be filled to form an expression with a complete content.

In papers written shortly after *Bs*, as well as in the *Grundlagen der Arithmetik* (*Gl*) of 1884, Frege extends this function-argument analysis to support an account of the formation of *concepts*. His plan is simply to transfer the replacement and omission model of *Bs* from the *expressions* of judgeable contents to the contents themselves. Thus, beginning with the content *Cato killed Cato*, we form the concept of suicide by viewing the *content* of the word 'Cato' as replaceable in both its 'occurrences'.[21] This plan requires that judgeable contents be constructed in a manner analogous to the sentences that express them, so that we can speak of the 'occurrences' of the contents of sub-sentential parts of sentences in the judgeable contents that the sentences express; and it requires a conception of the content of sub-sentential parts of sentences. All of this is governed by a principle of compositionality: the content of a complex expression is composed out of the contents of the parts of that expression, in a manner analogous to the way in which the expression is composed out of its parts. This compositionality of content has two consequences: first, the content of a part of an expression is a part of the content of the whole expression; and second, if two expressions have the same content, substituting one for the other within a larger complex expression will not change the content of the whole.

[21] 'Boole's logical calculus and the concept-script' (1882): 'I only allow the formation of concepts to proceed from judgements. If, that is, you imagine the 2 in the judgeable content

$2^4 = 16$

to be replaceable by something else, by (–2) or by 3 say, which may be indicated by putting an x in the place of the 2:

$x^4 = 16,$

the judgeable content is thus split up into a constant and a variable part. The former, regarded in its own right but holding a place open for the latter, gives the concept "4th root of 16"' (*NS*, p. 17/*PW*, p. 16). 'If from a judgeable content which deals with an object *a* and an object *b* we subtract *a* and *b*, we obtain as a remainder a relation-concept which is, accordingly incomplete at two points' (*Gl*, §70, p. 82).

The simplest form of sub-sentential expression is a proper name, such as 'Cato' in our example. What is the content of such a name? The answer implicit in Frege's account of concept-formation is: the object named by the name, in this case Cato himself.[22] The objects about which we judge are built into the judgeable contents themselves. Frege's model of judgeable content thereby simultaneously accounts for two kinds of norms governing our cognitive and linguistic acts of judging and asserting – on the one hand norms of logical consistency, inferential interconnection and generally responsibility of judgements to each other, and on the other hand, norms of truth, and generally responsibility to the world. Judgeable contents are individuated in terms of their consequences; hence, when one judges a given content to be true, it is determined what further contents one is committed to judge true as well. But the content which one judges true also contains as parts the objects about which one judges, and which determine the correctness of one's judgement. The concepts employed in judging are factored out of judgeable content by omitting the objects which figure in it. Concepts' dual character of content and extension derives from this factoring. A concept stands derivatively in consequence relations to other concepts, determined by the relations of the contents obtained by 'completing' them, and so can be said to have as content all concepts which 'follow' from it. At the same time, a concept determines a class of objects to which it applies, the class of objects which complete the concept to form a correct judgeable content; these make up the extension of the concept.

However, all is not well with this model. There seems to be an instability in trying to combine in one 'content' the two normative dimensions of truth, responsibility of judgement to the world, and inference, responsibility of judgements to one another. By building the objects about which we speak and think directly into judgeable contents, we risk identifying contents that we want to distinguish on inferential grounds. Frege is not unaware of this difficulty in *Bs*, where he attends to a special case of it, involving the concept of identity, a primitive logical sign in *Bs*.

[22] Compare the passages cited in note 21 above.

The problem arises as follows. Suppose that $a = b$, that is, a and b are the very same thing, and consider the sentences '$a = a$' and '$a = b$'.[23] These sentences have different consequences, and so must differ in content. From '$a = b$,' together with 'Fa', we can infer 'Fb'. From '$a = a$', no such conclusion follows. Yet, since $a = b$, the names 'a' and 'b' have the same content. Thus substituting one for the other should not change the content of the whole. But such a substitution transforms '$a = b$' into '$a = a$'. So these sentences must have the same content.

Frege responds to this dilemma in *Bs* by proposing a metalinguistic account of the identity sign – in identity contexts, names stand not for their contents but for themselves (*Bs*, §8, p. 124). Thus, '$a = b$' says that 'a' and 'b' have the same content. This affords a way out of the difficulty. We do not have to conclude that if $a = b$, '$a = b$' and '$a = a$' have the same content. *In this context*, even though $a = b$, 'a' and 'b' do not have the same content, as long as they are distinct names – since *in this context* the content of 'a' and 'b' is the names themselves.

Frege recognizes an objection to this solution, however. On this view, assertions of identity 'pertain to the *expression* and *not to the thought*', concerning only our form of expression, and not the things of which we would speak (*Bs*, §8, p. 124). This leads easily to the conclusion that in a perspicuous *Begriffsschrift*, we have no need for different names for one thing, or for a sign of identity. If the job of a name is to stand for its content, and we adhere to the principle, one sign for one job, the sign of identity will be superfluous.

Frege uses a geometrical example to respond to this objection. Beginning with a point A lying on a circle, another point is constructed as the point of intersection of a certain line and the circle. One then discovers that the first point and the second point are one and the same (*Bs*, §8, p. 125). Frege goes on to explain that in this example, 'the same point is determined in two ways': 'directly in intuition' and 'as the point B corresponding to' the given construction.

[23] In *Bs*, Frege uses the sign '≡' for identity; in his later writings he uses '='.

Frege's argument here echoes another distinction from Kant's *Logic*: that between the matter and form of a cognition.[24] Kant says that the matter of a cognition is the object, whereas the form is '*the way in which* we cognize the object'. He too provides an example (*LL*, pp. 544–5):

> If a savage sees a house from a distance … with whose use he is not acquainted, he … has before him in his representation the very same object as someone else who is acquainted with it determinately as a dwelling established for men. But as to form, this cognition of one and the same object is different in the two. With the one it is *mere intuition*, with the other it is *intuition* and *concept* at the same time.

Frege's example, like Kant's, involves an object being given both intuitively and conceptually. However, Frege's talk of different ways of 'determining' the same point modifies Kant's formulation in one key respect: *distinct* objects cannot be *determined* in the *same* way. In Kant's terms, where the form is the same, so must be the matter.

Frege uses his example to dispel the impression that the identity sign is dispensable: 'That the *same content* … is … given by *two modes of determination* is the content of a *judgment*' (*Bs*, §8, p. 125). The identity sign allows us to express such contents. When distinct names '*a*' and '*b*' are associated with distinct modes of determining the same object, the identity sentence '$a = b$' will bring this out. In such cases, Frege asserts, distinct names for one object, far from being superfluous, 'concern the very heart of the matter' (*das Wesen der Sache selbst betreffen*), and identity judgements are 'in Kant's sense, synthetic' – they extend our knowledge through having new and useful consequences (*Bs*, §8, p. 126). Frege thus officially introduces the identity sign in the following way: '$a = b$' asserts that '*a*' and '*b*' have the same content, so that each can be substituted for the other wherever it occurs (*Bs*, §8, p. 126). Thus the inferential content of '$a = b$' is preserved – from it, together with the sentence '*Fa*', we can infer the sentence '*Fb*'; and indeed one of the basic axioms of *Bs*, '$a = b \supset (Fa \supset Fb)$', codifies this inference.

[24] This point is also made by Gottfried Gabriel, 'Objektivität: Logik und Erkenntnistheorie bei Lotze und Frege', editor's introduction to H. Lotze, *Logik: Drittes Buch. Vom Erkennen (Methodologie)* (Hamburg: Felix Meiner, 1989).

There is, however, an initial difficulty with this way of explaining the function of '='. Frege first asserted that '$a = b$' states that the names 'a' and 'b' have the same content; later he claimed that '$a = b$' expresses the judgement that 'the same object is determined in different ways'. This seems to provide two ways of understanding the content of '$a = b$': (1) as a claim about the *names* 'a' and 'b', that they have the same content; (2) as a claim about the associated *modes of determination*, that they determine the same thing. It is not obvious that these are compatible. We will see that Frege's introduction of the sense–reference distinction is intertwined with an attempt to address this problem.

In *Bs*, Frege applies his identity sign to sentences, expressions of judgeable contents. He introduced his notion of conceptual content through a natural-language example of distinct sentences with the same content. The apparatus of *Bs* generates similar cases, involving sentences which Frege recognizes as mutually inferrable, since their equivalence is 'obvious' enough that we can count each as following directly from the other. The simplest such case is that of double negation – the distinct sentences 'A' and '$\sim\sim A$' express the same content.[25] Two of the axioms of *Bs* are '$\sim\sim A \supset A$' (proposition 31) and '$A \supset \sim\sim A$' (41) (*Bs*, §§18–19, pp. 156, 158). In the Preface, however, Frege says that these axioms 'can be combined into the single formula $\vdash(\sim\sim A = A)$' (*Bs*, p. 107). This new axiom states that '$\sim\sim A$' and 'A' have the *same content*, according to the *Bs* account of '='. One might argue that all Frege is really claiming is the *material* equivalence of 'A' and '$\sim\sim A$', since this is all that is stated by (31) and (41). However, the proposed new axiom is *not* a mere replacement for (31) and (41) but a real enrichment of *Bs*, which contains no principles from which one could derive '$\sim\sim A = A$'. There is no possibility in *Bs* of deducing '$A = B$' from the material conditionals '$A \supset B$' and '$B \supset A$'.[26]

[25] What follows makes it plausible that in *Bs* Frege took double negation to preserve content. There are no other clear examples in *Bs*. In 'Compound thoughts' (1923–6), Frege says that '$\sim\sim A$' and 'A' express the same thought, that 'A & B' and 'B & A' express the same thought, and that contraposition preserves the thought ('Compound thoughts', pp. 393, 399, 403).

[26] In contrast, the system of the later *Gg* has a much stronger identity axiom.

It has been frequently pointed out that the *Bs* theory of identity is beset with problems.[27] For example, the theory implies that in the *Bs* axiom '$a = b \supset (Fa \supset Fb)$' '$a$' and '$b$' are ambiguous, standing for themselves in '$a = b$', and for a and b in 'Fa' and 'Fb'. This makes conceptual trouble for the use of this axiom when 'Fa' is replaced by an identity context, an application which Frege needs to deduce the symmetry of identity (at proposition (55)).[28] One such application yields '$a = b \supset ((a = a) = (a = b))$' – in other words, if '$a$' and '$b$' have the same content, so do '$a = a$' and '$a = b$'. This demonstrates the bankruptcy of the proposed solution in the context of the formal system of *Bs* – the formal system itself treats identity as a relation between the objects named, not between names.

Avoiding these difficulties while retaining the *Bs* account of '=' would require a major overhaul of the formal system of *Bs*. But the *Bs* account of identity fails at a more fundamental level, for it does not really address the problem it was intended to solve. That problem was that, in treating the objects about which we judge as parts of the contents which we recognize as true, we end up conflating contents that we want to hold apart because they do not have the same consequences. The *Bs* account of identity addresses this issue only in the case of identity sentences, by allowing that, even when $a = b$, '$a = a$' and '$a = b$' can have different content. But the general *Bs* account of content still implies that, when $a = b$, 'Fa' and 'Fb' have the same content, and so are mutually inferrable. This, however, violates our intuitions about what follows from what. We do not suppose that, merely because $a = b$, one who asserts 'Fa' and denies 'Fb' thereby contradicts herself, as would happen if 'Fb' were a consequence of 'Fa'. Generalizing the *Bs* account of identity to handle all such problems would lead to the unfortunate result that, in all contexts, names stood for themselves rather than for their content.

These sorts of difficulties were very much in Frege's mind at the time of the composition of 'On sense and reference' ('Über Sinn und

[27] See, for example, Mendelsohn, 'Frege's *Begriffsschrift* theory of identity', Joan Weiner, 'Frege and the linguistic turn', *Philosophical Topics*, 25/2 (1997), pp. 265–88, and Dummett, *IFP*.

[28] There is also trouble for quantification into identity contexts, especially into 'mixed' contexts; Frege needs to be able to quantify into these contexts, for example in his definition of 'many-one procedure' (proposition (115)) and the subsequent theorems.

Bedeutung', *S&B*, 1892) and *Grundgesetze der Arithmetik* (*Gg I*, 1893). However, in the years following the publication of *Bs*, Frege continued to work with the picture of content we have sketched here. In his 'logicist manifesto', *Grundlagen der Arithmetik* (*Gl*) of 1884, the *Bs* picture is assumed throughout, and plays an important role in his account of the Kantian analytic–synthetic distinction, with which he frames the project of the book.[29] However, by the time of the publication of his second great logicist work, *Gg*, the basic picture of content adumbrated so far had been replaced by the famous theory of sense and reference. In my view the sources of this fundamental reworking of Frege's thought are to be found in the development of his logicist project in *Gl*.

3 OBJECTIVITY, OBJECTHOOD AND THE CONTEXT PRINCIPLE IN *GRUNDLAGEN*[30]

Frege wrote *Bs* with the aim of establishing the epistemological status of arithmetic. In part III of that work, he proved some results in a 'general theory of sequences', which he hoped to be able to apply to the sequence of natural numbers. Five years later, in *Gl*, he developed the philosophical basis of his claim that, *contra* Kant, arithmetic is analytic, and that its truths can be proved from logical laws and definitions alone. In *Gl*, he claimed only to have made this plausible informally; the formal demonstration, making use of the apparatus of *Bs*, was reserved for *Gg*, the first volume of which appeared nine years after *Gl*. In *Gl* the need for developing a more careful account of judgement and content began to reveal itself; and the project of *Gg* led to the full-fledged theory of sense and reference.

Frege prepares the ground for his own account in *Gl* with a critique of other philosophies of arithmetic. In the course of his argument, he emphasizes both the *objectivity* of mathematical truths and the *objecthood* of the numbers which mathematics studies. He thus sets himself against two tendencies in the philosophy of mathematics, against which he polemicized throughout his subsequent

[29] For an illuminating discussion of this see Jamie Tappenden, 'Extending knowledge and "fruitful concepts"', *Noûs*, 29 (1995), pp. 427–67.

[30] A more extended discussion of these issues is found in Ricketts, chapter 6, this volume, §§7 and 8. I have followed Austin in translating 'Bedeutung' as 'meaning' in *Gl* since Frege had not then given the term its special technical significance.

career: psychologism and formalism. The former takes mathematical terms to mean *ideas*, while the latter avoids ascribing any meaning to them whatsoever.

Frege's argument turns on 'three fundamental principles' (*Gl*, p. x):

always to separate sharply the psychological from the logical, the subjective from the objective

never to ask for the meaning of a word in isolation, but only in the context of a sentence

never to lose sight of the distinction between concept and object.

The first principle rejects psychologism; the third principle, Frege says, implies that 'a widely held formalist theory ... is untenable.'[31] The second of the three principles, the 'Context Principle', is the lynchpin on which the others depend.[32] Frege claims that, if the Context Principle is violated, 'one is almost forced to take as the meanings of words mental pictures in the individual mind, and so to offend against the first principle as well'. He is less explicit about the relation between the Context Principle and the concept–object distinction, but it is there nonetheless. In *Gl*, Frege gets at the distinction between concept and object through a distinction between *names* and *concept-words* (predicates), itself drawn with the help of the Context Principle. It is only by considering how a word functions in a sentence that we can determine its logical place as name or predicate, and so determine the place of its content as concept or object.

Frege's distinction between the psychological/subjective and the logical/objective leads him to reconsider the Kantian vocabulary of 'representation' (*Vorstellung*) and 'content' employed in *Bs*. He now recognizes the ing–ed ambiguity of 'representation': 'in compliance with the first principle, I have used the word 'representation' always in the psychological sense, and have distinguished representations

[31] This is the theory that mathematical existence is simply consistency. Frege's point is that consistency is a property of the concept with which one defines an object, and does not guarantee existence of the object. The formalist confuses concept and object here. (*Gl*, §95, p. 106.)

[32] My understanding of the Context Principle is indebted to Jim Conant's detailed analysis of Wittgenstein's use of it in 'The method of the *Tractatus*', in E. Reck (ed.), *From Frege to Wittgenstein: Perspectives on Early Analytic Philosophy* (Oxford: Oxford University Press, 2002).

from concepts and from objects'[33] (*Gl*, p. x). In a footnote to a discussion of the view that 'number is the representation of the position of an item in a series' he writes (*Gl*, §27, p. 37):

My arguments would be beside the point if he meant by representation an objective idea (*Idee*); but in that case what difference would there be between the representation of the position and the position itself?

A representation in the subjective sense is what is governed by the psychological laws of association; it is of a sensible, pictorial character. A representation in the objective sense belongs to logic and is in principle non-sensible, although the word which means an objective representation is often accompanied by a subjective representation, which nevertheless is not its meaning. Subjective representations are often demonstrably different in different men, objective representations are the same for all. Objective representations can be divided into objects and concepts. I shall myself, to avoid confusion, use 'representation' only in the subjective sense.

Here, Frege distinguishes between the *act* of represent*ing*, the 'subjective representation', and the content represent*ed*, the 'objective representation.[34] In *Bs* Frege marked the distinction between the content of judgement and the act of judging in his logical notation, through his 'judgement-stroke.' However, he was not yet completely clear on the subjective–objective distinction drawn in *Gl*, as is shown by his treatment of judgeable content as a 'mere combination of representations', 'formed' by a mental act. *Gl*, in contrast, emphasizes the objectivity of content.

If arithmetic is an objective science, it must have its 'objective representations', objects and concepts. Frege insists on a sharp distinction between objects and concepts, and argues that numbers are not concepts, but objects. The notion of concept here is modelled on the linguistic idea of 'function' of *Bs*. That account of function, however, presupposes a category of 'complete' expressions which are

[33] Austin translates '*Vorstellung*' as 'idea' in order to bring out this psychological aspect of '*Vorstellung*' in Frege's usage; but this obscures Frege's point in the footnote discussed below.

[34] Aware of the fact that 'representation' and associated words might be understood in the 'objective' sense, Frege repeated such cautions in his later writings. For example, in 'On sense and reference' (hereafter *S&B*) he writes: 'We may include with representations intuitions ... One may on the other hand understand intuition as including any object in so far as it is sensibly perceptible and spatial.' (*S&B*, p. 29, fn. 3.)

not functions, and which serve both as the 'wholes' within which we can omit and vary parts to obtain functions, and as the 'parts' which we omit and vary.[35] Similarly, the account of concepts in *Gl* presupposes both judgeable contents and objects as parts of those contents. Yet the Context Principle might seem to undercut the needed distinction by making all sub-sentential expressions equally 'incomplete', and all non-judgeable contents equally dependent on judgeable contents.

To resolve this difficulty we must attend to Frege's actual use of the context principle in *Gl*. When he writes, 'Only in a sentence have the words really a meaning … It is enough if the sentence taken as a whole has a sense; it is this that confers on its parts also their content' (*Gl*, §60, p. 71), it is tempting to take this in one of two extreme ways: (1) the content of a word is determined by the sense of *any* sentence in which it occurs; (2) the content of a word is determined by the sense of *all* sentences in which it occurs. Either of these possibilities would make it difficult to sustain a real distinction between concept and object. Neither of these suggestions, however, reflects Frege's intention.

Frege approaches the nature of numbers through the Context Principle: 'It should throw some light on the matter to consider number in the context of a judgement which brings out its basic use' (*Gl*, §46, p. 59). He considers 'statements of number' such as 'there are two senators from Indiana' and concludes that 'a statement of number contains an assertion about a concept' (we can say of the concept *senator from Indiana* that two individuals fall under it). This might lead to the view that numbers are 'second-order concepts' (the number two is the concept under which such concepts as *senator from Indiana* and *prime less than 5* fall). However, Frege denies that numbers are such concepts. Rather, they are 'self-subsistent objects that can be recognized as the same again' (*Gl*, §56, p. 68). Frege distinguishes concepts and objects by the *kind* of questions one can ask about them, the *kind* of sentences in which they occur. 'With a concept, the question is always whether anything,

[35] Higher-order functions are generated by allowing lower-order functions to be the arguments which are omitted and replaced; but recognizing these lower-order functions requires that at some level there are arguments that are themselves not functions.

and if so what, falls under it. With a proper name, such questions make no sense' (*Gl*, §51, p. 64). On the other hand, the 'self-subsistence', and thus objecthood, of numbers 'comes out at every turn, as for example in the identity 1 + 1 = 2' (*Gl*, §57, p. 69).

Frege's use of the Context Principle shows that contexts of the form '*a* is *F*' are primary in establishing the content of a predicate like '*F*', whereas contexts of the form '*a* = *b*' are primary in establishing the content of a name like '*a*'. The 'self-subsistence' of objects *consists* in their being 'recognizable as the same again', as expressed in statements of identity. Thus Frege can claim: (*Gl*, §60, p. 72):

> The self-subsistence which I am claiming for number is not to be taken to mean that a number word signifies something when removed from the context of a sentence, but only to preclude the use of such words as predicates or attributes, which appreciably alters their meaning.

This interpretation of Frege's use of the Context Principle is borne out by his answer to the Kantian question: 'How then are numbers given to us, if we cannot have any representations or intuitions of them?' He appeals to the Context Principle: 'Since it is only in the context of a sentence that words have any meaning, our problem becomes this: to define the sense of a sentence in which a number word occurs.' But he immediately narrows this problem: 'we have already settled that number words are to be understood as standing for self-subsistent objects. And that is enough to give us a class of sentences which must have a sense, namely those which express our recognition of a number as the same again.' The problem thus becomes 'to construct the content of a judgement which can be taken as an identity such that each side of it is a number' (*Gl*, §§62–3, pp. 73–4).

However, the use of *identity* contexts as the crucial marker of namehood, and so objecthood, puts considerable pressure on the *Bs* account of identity. For it requires *both* that identity be a relation between objects, rather than between names, *and* that non-trivial identity statements be possible. If identity were only a relation between names, the use of names in identity contexts could hardly 'confer content' on them in such a way as to 'give' us objects. But equally, if identity statements were always trivial, like '*a* = *a*', an ability to form such sentences and judge them to be true would have no interesting consequences. Yet the *Bs* account preserved the non-trivial character of identities only by treating identity as a relation between names and not objects.

In *Gl* these difficulties surface in Frege's attempts to define the numbers. His stated aim to 'construct the content of a judgement which can be taken as an identity such that each side of it is a number' shows that he understands identity as a relation between objects which are parts of the content of an identity judgement.[36] On the other hand, he emphasizes the importance of non-trivial identities, arguing that if we assumed that every object could be given in only one way,

All identities would then amount simply to this, that whatever is given to us in the same way is to be reckoned as the same. This, however, is a principle so obvious and so sterile as not to be worth stating. We could not, in fact, draw from it any conclusion which was not the same as one of our premises. Why is it, after all, that we are able to make use of identities with such significant results in such divers fields? Surely it is rather because we are able to recognize something as the same again even although it is given in a different way.[37]

[36] This is also clear from his acceptance of Leibniz's definition of identity (*Gl*, §65, p. 76.): 'Things are the same as each other, of which one can be substituted for the other without loss of truth [*Eadem sunt, quorum unum potest substitui alteri salva veritate*]'. Beaney sees use-mention confusion here, and assumes that Frege is thinking of substituting *expressions* for one another (*FMS*, p. 155). However, Frege's discussion shows that he is thinking of substituting *objects* within judgeable contents. He writes (Gl, §65, p. 77): 'In order … to justify our proposed definition … we should have to show that it is possible, if line *a* is parallel to line *b*, to substitute the direction of *b* everywhere for the direction of *a*.' While Austin and Beaney (*FMS*, p. 101; *The Frege Reader* (Oxford: Blackwell, 1997), p. 112) insert quotation marks around 'the direction of *b*' and 'the direction of *a*', these are not present in the German text. A little later Frege remarks that, if we introduce any new 'assertion about directions', we will need to 'make it a rule always to see that it must remain possible to substitute for the direction of any line the direction of any line parallel to it'. Here too identity appears as a relation between objects, not names.

[37] Beaney emphasizes the importance of this passage (*FMS*, p. 102). It is worth noting that, in context, this passage equivocates on the key notion of a 'way of being given'. This shows that at this stage Frege had not yet achieved complete clarity on this issue. The passage occurs as part of Frege's discussion of a proposed definition of 'direction' in terms of parallelism:

the direction of $a =$ the direction of b if and only if $a//b$.

His worry is that, while this settles when the direction of one line is identical with that of another, it does not establish a *general* criterion which would settle for any q whether $q =$ the direction of a. Hence, it does not tell us *what* the direction of a line is, does not 'give us' the object. Frege notes that, if we had the

This passage clearly harks back to the *Bs* account of content and identity, pointing out that identities often have non-vacuous *content*, serving as the ground for novel consequences. At this point, however, Frege had no account of how non-trivial identities are possible, compatible with a treatment of identity as a relation between objects. Such an account would have to wait for the theory of sense and reference.

4 TRANSITION TO 'ON SENSE AND REFERENCE'[38]

Difficulties with identity and the notion of object were on Frege's mind in the period between *Gl* and *S&B*, judging from the evidence of his unpublished writings. In his *Nachlass*, a folder headed 'On the concept of number' was found, the contents of which were separated by the editors of *NS* into two parts, since the second part was 'obviously a preliminary draft' (*NS*, p. 96/*PW*, p. 87) of 'On concept and object', while the first contained criticisms of the views of Otto Biermann. The editors assign to these manuscripts a date of 1891/2,

'concept of direction', we could complete our definition by stipulating that, if q is not a direction, $q \neq$ the direction of a. He considers the stipulation(*) q is a direction if and only if q is introduced by means of the proposed definition. He replies: (*Gl*, §67, p. 78): 'If … we were to adopt this way out, we should have to be presupposing that an object can only be given in one single way; for otherwise it would not follow, from the fact that q *was* not introduced by means of our definitions, that it *could* not have been introduced by means of it.' The passage quoted in the main text follows this immediately.

Yet this is an equivocation. Frege says that (*) presupposes that an object can be given in only one way. This is true *if* he means 'a direction can only be given *as a direction*'. He then concludes that, if an object can only be given in one way, all identities become sterile instances of the obvious tautology that what is given in the same way is the same. But *this* conclusion follows only if he means that the only way in which the direction of a can be given is *as the direction of a*, and this in no way follows from (*). It is fully compatible with (*) that the same object be given as the direction of a and as the direction of b, in which case the true identity 'the direction of a = the direction of b' will be neither more nor less trivial than '$a//b$'.

The equivocation involved here is between a conception of the 'way in which an object is given' which, like Kant's 'form', permits several objects to be 'given in the same way', and a conception which, like *Bs*'s 'mode of determination', requires that at most one object can be given (determined) in any particular way. That Frege was capable of this equivocation reveals that his own conception of what it is for an object to be 'given' was not yet completely fixed (*Gl*, §67, pp. 78–9).

[38] This section is speculative and can be skipped without significant loss of continuity.

on the grounds that 'On concept and object' appeared in 1892, noting that 'the section dealing with Biermann may have been written earlier' since the pre-1891 terminology of 'content' is used there (*NS*, p. 81/*PW*, p. 72).

Careful attention to these texts reveals some interesting points. First, they seem to form one continuous piece of writing, as indicated by Frege's placing them in one folder. The supposed draft of 'On concept and object' begins 'I turn now to consider …' – not how one would expect a draft of a free-standing essay to begin. Second, while the terminology of sense and reference does occur there, the terminology of judgeable content is also used, even after the point at which Frege introduces the sense–reference distinction (*NS*, p. 108/*PW*, p. 99; compare the long footnote at *NS*, p. 109ff./*PW*, 100ff.). Third, the first part begins as if introducing a larger work (*NS*, p. 81/*PW*, p. 72):

In my *Grundlagen* (§68) I called the concept *F* equal in number to the concept *G* if it is possible to correlate one-to-one the objects falling under *F* with those falling under *G* and then gave the following definition:

The number belonging to the concept *F* is the extension of the concept 'equal in number to the concept *F*'.

The following discussion will show that this definition gives the right results when applied, by deriving the basic properties of numbers from it. But first we need to clarify a few points and meet some objections.

Of what projected work is this the beginning? In *Gg* Frege writes that 'internal changes in my *Begriffsschrift* … forced me to discard an almost completed manuscript' (*Gg*, vol. I, p. ix). Frege traces these 'internal changes' to 'a thoroughgoing development of my logical views', highlighting in particular the sense–reference distinction. I suggest that the manuscripts headed 'On the concept of number' constitute part of the introduction to the 'almost completed' version of *Gg* which Frege had to discard. This makes sense of the opening paragraphs of the manuscript: the task of *Gg* was precisely to give a formal proof of the correctness of *Gl*'s philosophy of arithmetic. The points to be clarified are found in the criticisms of Biermann, and the objections to be met are Kerry's criticisms of the concept–object distinction. These manuscripts can plausibly be dated to the late 1880s: Biermann's book appeared in 1887, and Kerry's articles in 1886 and 1887. Apparently, Frege abandoned this introduction to *Gg* upon drawing the sense–reference distinction, after a half-hearted attempt to simply incorporate the distinction into the manuscript; eventually, he reworked a part of the manuscript into 'On concept and object'.

How, in writing this manuscript, was Frege driven to the sense–reference distinction, though? There are some fascinating hints. His critique of Biermann and Kerry must have focused his thought on the notion of *object*, which we saw to be intimately related to the problem of identity. Kerry's criticism of the concept–object distinction challenged the very core of Frege's claim that numbers are objects and not concepts. His final point against Biermann, that 'there is only one number called 0, there is only one number called 1, only one number 2, and so on' (*NS*, p. 94/*PW*, p. 85), drew him directly to the problem of identity, leading him to write (*NS*, pp. 94–5/*PW*, p. 85):

> There are various designations for any one number. It is the same number which is designated by '1 + 1' and '2'. Nothing can be asserted of 2 which cannot also be asserted of 1 + 1; where there appears to be an exception, the explanation is that the signs '2' and '1 + 1' are being discussed and not their content. It is inevitable that various signs should be used for the same thing, since there are different possible ways of arriving at it, and then we first have to ascertain that it is really the same thing we have reached.

Here Frege attempts a striking generalization of *Bs*'s metalinguistic theory of identity, suggesting that *all* contexts in which substitution of identicals fails are metalinguistic. This would entail that in indirect discourse contexts such as 'Caesar said that Tully was a great orator', embedded names like 'Tully' stand for themselves rather than for their usual contents (in this case Cicero). This both anticipates and yet differs importantly from his later view that in such contexts words stand for their *ordinary sense*.[39] The importance to Frege of this issue is shown by the fact that in the Introduction to *Gg*, Frege cites only one thing in support of the sense–reference distinction: 'only in this way can indirect discourse be correctly understood' (*Gg I*, p. x) – a remarkable fact, given that indirect discourse is not even represented in the formal system of *Gg*.

[39] The view suggested in Frege's response to Biermann was later put forward by Carnap and criticized by Church, whose own position resembled Frege's later, settled view.

It is not clear exactly when or how the drawbacks of a metalinguistic account of indirect discourse presented themselves to Frege.[40] But it is plausible that in his first attempts to draft an introduction to *Gg*, he began to see how the problem of identity required something like the sense–reference distinction. But there was another route to the distinction as well, connected to the eventual introduction to the published version of *Gg*. While Frege opens *S&B* with the problem of informative identities, and often reverts to this problem in explaining the need for the sense–reference distinction ('On Mr Peano's conceptual notation and my own', p. 241; 'Introduction to logic', *NS*, pp. 208–9/*PW*, p. 192; 'Logic in mathematics', *NS*, p. 243/*PW*, p. 225; Frege to Jourdain, January 1914, *WB*, pp. 127–8/*PMC*, pp. 79–80; Frege to Peano, undated, *WB*, p. 196/*PMC*, p. 127; Frege to Russell, 28 December 1902, *WB*, pp. 234–5/*PMC*, pp. 152–3), he sometimes also characterizes the distinction in terms of his having 'split up' or 'analysed' (*zerlegt*) into thought and truth-value what he had previously combined in his former notion of judgeable content ('On concept and object', p. 187; *Gg*, vol. I, p. x; Frege to Husserl, 24 May 1891, *WB*, p. 96/*PMC*, p. 63). This points to a second motivation for the distinction, having to do with not with sub-sentential content and the problem of identity, but with sentential, judgeable content and Frege's account of judgement. This motivation arises with the development of Frege's attack on psychologism.

In *Gl*, and in other writings of about the same time, Frege's discussions of psychologism focused on distinguishing between *ideas* (subjective representations) and *concepts* and *objects* (objective representations), with Frege arguing, for example, that while ideas develop and change, concepts do not (*Gl*, pp. v, vii; 'On the law of inertia' (1891), pp. 132–5). In contrast, Frege's polemic against 'psychological logicians' in the Introduction to *Gg* focuses on the status of the *laws of logic* (*Gg*, vol. I, pp. xiv–xxvi). He distinguishes between psychological laws, which describe *how* people think, and logical laws, which are laws of truth, and 'prescribe universally the way in which one ought to think' (*Gg*, vol. I, p. xv). This form of polemic first occurs in Frege's writings in an unpublished 'Logic'. The date of

[40] Church pointed out that in translating from one language to another, we translate the words occurring in indirect discourse contexts, which contradicts the metalinguistic account.

this piece can only be established with certainty as between 1879 and 1891;[41] but internal evidence suggests that it was written fairly close to 1891. In the first place, the resemblance of its argumentative strategy to that of the Introduction to *Gg*, and to nothing else written by Frege before *Gg*, argues for a date close to the time of composition of *Gg*. Moreover, in the course of his argument against psychologism, Frege employs, as an example of an unchanging law, the law of inertia (*das Trägheitsgesetz*) ('Logic', *NS*, pp. 4–5/*PW*, pp. 4–5). In 1891 Frege reviewed a book published in 1886, under the title 'On the law of inertia' ('Über das Trägheitsgesetz'). This review devotes several pages to a discussion of psychologism. This suggests that Frege wrote the undated 'Logic' at a time when he was working on 'On the law of inertia', thus between 1886 and 1891.[42]

The 'Logic' opens with a characterization of judgement and assertion: 'Inwardly to *recognize something as true* is to *make a judgement*, and to give expression to this judgement is to make an assertion' (*NS*, p. 2/*PW*, p. 2). Arguments against psychologism, reminiscent of *Gg*, follow, leading to the conclusion that 'it is the business of the logician to conduct an unceasing struggle against psychology…' (*NS*, p. 7/*PW*, p. 6). Frege then turns to a brief discussion of the notion of 'judgeable content', after which the manuscript breaks off (*NS*, pp. 7–8/*PW*, pp. 7–8).

In attempting to characterize his idea of judgeable content, Frege emphasizes its objectivity: 'such a content is not the result

[41] It contains a reference to *Bs*, published in 1879, and uses the terminology of 'judgeable content' abandoned in 1891.

[42] Hans Sluga ('Frege on the Indefinability of Truth', in Reck (ed.), *From Frege to Wittgenstein*, p. 82) thinks that this 'Logic' dates from before 1884, since its projected table of contents includes a discussion of 'judgments in which something is recognized as the same again' under the heading 'definition of objects'. Sluga sees here a reference to the strategy, rejected in *Gl*, of contextually defining numbers via fixing the sense of numerical identities. However, even explicit definitions must make use of identities; moreover, as Burge has pointed out, there is evidence in the catalogue of Frege's *Nachlass* that, even after 1884, he continued to toy with the contextual definition strategy. See Tyler Burge, 'Frege on extensions of concepts, from 1884 to 1903', *Philosophical Review*, 93 (1984), pp. 3–34, here pp. 12–13; and Albert Veraart, 'Geschichte des wissenschaftlichen Nachlasses Gottlob Freges und seiner Edition. Mit einem Katalog des ursprünglichen Bestands der nachgelassenen Schriften Freges', in M. Schirn (ed.), *Studien zu Frege*, vol. I: *Logik und Philosophie der Mathematik* (Stuttgart: Frommann, 1976), item 47, p. 95.

of an inner process … but something objective … something that is exactly the same for all … who are capable of grasping it' (*NS*, p. 7/*PW*, p. 7). Thus not only is his characterization of judgement as the 'inward recognition of something as true' close to his later slogan that 'the act of judgement' is 'the recognition of the truth of a thought', he also anticipates here his account of the necessary preceding stage: 'the grasp of a thought – thinking' ('Thoughts', p. 355). In the 'Logic', Frege appeals to the phenomenon of yes-no questions to further explain his idea of content: 'A judgement is often preceded by questions … We grasp the content of a truth before we recognize it as true, but we grasp not only this; we grasp the opposite as well' (*NS*, p. 8/*PW*, p. 7). In his late essay 'Negation', Frege appeals to just this point to argue against the view that 'a thought has being by being true', insisting that 'being-true cannot be reckoned to the sense of a propositional question' ('Negation', pp. 373–4). His argument points precisely to the need to distinguish thought from truth-value, and the correlative need to distinguish two aspects in the act of judging – grasp of thought and recognition of truth. In a footnote to 'Negation', Frege writes (p. 381):

> We are probably best in accord with ordinary usage if we take a judgement to be an act of judging, as a leap is an act of leaping. Of course this leaves the kernel of the difficulty uncracked; it now lies in the word 'judging'. Judging, we may say, is recognizing the truth of something; what is recognized as true can only be thought. The original kernel now seems to have cracked into two; one part of it lies in the word 'thought' and the other in the word 'true'. Here for sure we must stop.

This cracking of the kernel, the act of judging, parallels precisely the splitting of the 'judgeable content' into thought and truth-value.

Consequently, at the end of the 'Logic' manuscript, Frege was already very close to his later distinction between truth-value and thought, and the corresponding distinction between grasp of thought and recognition of truth. He seems, in fact, to have broken off the manuscript just at the point where he realized the necessity for making such a distinction, in the midst of his account of 'judgeable content'. In any case, such a distinction is needed in the context of the argument against psychologism which he began to sketch in the 'Logic' and which he developed more fully in *Gg*. This argument requires that we maintain a three-fold distinction between (1) thinking, the psychological act of grasping a thought, (2) judging, the act of recognizing a thought as true, and (3) the actual truth of the thought. Psychological 'laws of thought' concern (1) and (2), whereas logical 'laws of thought' concern primarily (3), but give rise to prescriptions concerning (1) and (2), *given* that the aim of thought, and especially judgement, is truth. That such distinctions are necessary was already pointed out in *Gl*, where Frege warns against confusing the 'being-thought [*Gedachtwerden*] of a proposition with its truth' (*Gl*, p. vi; §77, p. 91). Yet the *Bs* account of content presupposed in *Gl* fails to accommodate this important distinction.

The problem is this: according to the *Bs* picture of content, a judgeable content contains as parts the objects which the content is about, and the concepts which it predicates of those objects. But on this picture, a judgeable content could not *exist* without being *true* – the *Bs* account supports the thesis against which Frege argues in 'Negation', that the being of a thought is its truth. Consider, for example, the judgeable content that Desdemona loves Othello.[43] This content is made up of the objects Desdemona and Othello, linked by the unsaturated concept expressed by the predicate '( )loves( )'. Given that this content *exists*, it would seem that Desdemona must, in fact, love Othello. For what else could her being linked to him by the concept of loving amount to? If in grasping the thought that Desdemona loves Othello, I grasp a complex entity consisting of Desdemona and Othello linked in this way, how can my grasp not translate into a recognition of truth? On the *Bs* model, to understand the question 'Does Desdemona love Othello?' would be to answer it 'yes'. Such reflections may have led Frege to break off his 'Logic' just at the point when the need to 'crack the kernel' of judgeable content showed itself.

Thus, in Frege's unpublished writings from the period immediately prior to 'On sense and reference', two sets of concerns cause him to break off his work. On the one hand, concerns about identity and the notion of object interfere with his writing of an apparent introduction to *Gg*; on the other hand, concerns about the notion of

[43] The example is Russell's; I choose it to highlight the similarity of the problem here with one which confronted Moore and Russell's early account of judgement and truth.

judgeable content interfere with his writing of a 'Logic'. The theory of sense and reference will address both sets of concerns.

5 'ON SENSE AND REFERENCE': INTRODUCTION

We come now to the official presentation of Frege's views in his most famous essay, 'On sense and reference' (*S&B*, 1892). Frege begins *S&B* with essentially the picture of 'content' sketched in §§1 and 2 above. But he now recognizes the instability of this picture – the contents assigned to names on the one hand and to sentences on the other cannot fit together in the way demanded by the principles of substitution and compositionality. The metalinguistic account of identity in *Bs* was an attempt to get around this problem. *S&B* opens with a reconsideration of this account.

Frege begins by asking whether identity is a relation, and if so whether it relates objects or names.[44] He states that in *Bs* he had taken it to relate names, and says that he will expound the reasons for this view. The rest of the long opening paragraph of *S&B* is often thought to contain not only the argument for the earlier view but also some or other argument against it; but a careful comparison of *Bs* and *S&B* shows that Frege is, just as he says, simply recapitulating the argument for the *Bs* account. We will follow this paragraph closely (*S&B*, pp. 157–8).[45]

First, Frege outlines the problem posed by identity statements. If identity relates objects, it seems that '$a = b$' and '$a = a$' must have the same content, since they state the same relation to hold between the same objects. This argument turns on the idea that name-content is part of sentence-content – the two identity sentences here are composed out of the same parts in the same way. However, Frege points out, the two sentences have manifestly different 'cognitive value' (*Erkenntniswert*). '$a = a$' is a trivial consequence of the law of identity – it has no interesting consequences. '$a = b$' may require investigation to discover; it can be a fruitful discovery, rich in

[44] In a footnote he states 'I understand 'a=b' to have the sense of 'a is the same as b' or 'a and b coincide'. This might be taken to endorse the view that identity is a relation between objects – except that the footnote occurs before the question of the nature of identity is posed.

[45] In this discussion I am indebted to William Taschek, 'Frege's puzzle, sense and information content', *Mind* 101 (1992), pp. 767–91.

consequences. 'The discovery that the rising sun is not new every morning, but always the same, was one of the most fertile astronomical discoveries. Even today the reidentification of a small planet or comet is not always a matter of course' (*S&B*, p. 157). Clearly, these sentences differ in content, in their *consequences*. In later writings, Frege sometimes uses the terminology of 'content' explicitly to make this point. For example, in correspondence with Peano, he writes that 'anyone can see that the thought of second sentence ['The evening star is the same as the morning star'] is different, and in particular that it is *essentially richer in content* [*wesentlich inhaltreicher*] than that of the former ['The evening star is the same as the evening star']' (Frege to Peano, undated, *WB*, p. 196/*PMC*, p. 127).

The *Bs* theory was meant to provide for this worry. It assigns different content to the two sentences by treating '$a = b$' as saying that the signs or names 'a' and 'b' stand for the same thing. However, Frege remarks, 'this is arbitrary'. He thus raises the worry expressed in *Bs*, that the metalinguistic theory rescues the wrong *sort* of content for identity-statements, content which 'pertains merely to the *expression* and *not to the thought*' (*Bs*, §8, p. 124). 'The sentence $a = b$ would no longer refer to the subject matter, but only to its mode of designation; we would express no proper knowledge by its means' (*S&B*, p. 157).[46]

In *S&B*, Frege clarifies this worry by asking what our talk of 'names' or 'signs' here amounts to. He asks whether 'the sign "*a*" is distinguished from the sign "*b*" only *as an object* (here, by means of its shape)' or '*as a sign* (*i.e.* ... by the manner in which it designates something)' (*S&B*, pp. 157–8, my emphasis). When we speak of 'signs', we might mean 'mere' signs, natural objects which we may put to some specific linguistic use, but which are not individuated by any features of that linguistic use. In that case, as Frege puts it, the use to which a sign is put is arbitrary – we could equally put that sign to some other use or put some other sign to that use. But we might also mean signs-in-use, individuated 'as signs', not *only* by such features as shape or size, but *also* by the linguistic use to which they are put.

[46] The phrase 'refer to the subject matter' translates 'die Sache selbst ... betreffen', an echo of *Bs*'s 'concern the very heart of the matter', 'das Wesen der Sache selbst betreffen'.

Frege had already implicitly deployed such a distinction in his critique of formalism in *Gl*, where he wrote that 'an empty sign (*Zeichen*) … without some content … is merely ink or print on paper, as which it possesses physical properties', but 'really … would not be a sign at all' (*Gl*, §95, p. 107). Shortly after *Gl*, Frege made an explicit distinction between 'figure' ('*Figur*') and 'sign' ('*Zeichen*'), the former possessing 'geometrical, physical and chemical properties', whereas the latter essentially has the 'purpose of designating' ('On formal theories of arithmetic' (1885), p. 115). In his critique of formalism, Frege emphasized the difference between an *empty* figure to which no content is assigned and a true *sign*, complete with content. But in later writings, he also applied this sort of distinction in the case of what could be called figures-in-use. For example, in 'Compound thoughts' (1918–19) he writes (p. 393):

> As a mere thing, of course, the group of letters 'and' is no more unsaturated than any other thing. It may be called unsaturated in respect of its employment as a sign [*Zeichen*] meant to express a sense, for here it can have its intended sense only when situated between two sentences: its purpose as a sign requires completion by a preceding and succeeding sentence.

Similarly, he emphasizes that a sentence (*Satz*) is a 'group of signs that expresses a thought', so that if a different thought were associated with the same group of sign-designs, 'it would not even be the same sentence' ('On the foundations of geometry, second series, II', p. 308; 'On the foundations of geometry: first series, I,' p. 277). More generally we can say that the same figure can be used in different ways, yielding different signs; we have the *same* sign if and only if we have the same figure, put to the same use.[47]

Adapting a suggestion of Wilfrid Sellars ('Abstract entities') to this context, let us use 'asterisk quotes' to form names of the 'figures' enclosed in them; and let us reserve ordinary quotation for naming 'signs', figures-in-use. So, for example, *and* occurs twice in *Sand and water make mud*, while 'and' occurs only once in 'Sand and water make mud'. Now suppose that '$a = b$' says that the

[47] This distinction is orthogonal to Peirce's type-token distinction; both figures and signs can be either tokens or types. Frege was well aware of such issues, but they need not detain us here.

figures *a* and *b* are used to stand for the same thing. Here 'used' must mean something like 'used by me/us'. We could then rewrite '$a = b$' as something like

$$\exists x(U(i,{}^\star a^\star,x) \,\&\, U(i,{}^\star b^\star,x))$$

where '$U(x,y,z)$' says that x uses y as a sign for z – there is something which *I* use both *a* and *b* to stand for. This expresses no 'proper knowledge' about the objects *a* and *b*, but only linguistic knowledge about the use to which *I* put the figures *a* and *b*. Why? Consider what *follows* from the identity statement '$a = b$' so understood. The answer is – nothing, except more statements about the use of figures. In particular, from '$a = b$' so understood, together with the premise '*Fa*', the conclusion '*Fb*' does *not* follow. For there are possible worlds in which I use both the expressions *Augustus* and *Julius* to refer to Tiberius, while in those worlds Augustus, and not Julius, remains the first emperor.[48] It is not quite true that on the *Bs* view of identity so understood, there is *no* difference in content between '$a = a$' and '$a = b$' – there is *a* difference, but it concerns only 'our mode of expression', not the objects *a* and *b* themselves. But this amounts, for present purposes, to no difference worth speaking of. This is why Frege says that 'if the sign '*a*' is distinguished from the sign '*b*' only as an object … not as a sign …, the cognitive value of $a = a$ becomes essentially equal to that of $a = b$, provided $a = b$ is true'.

The moral of the argument so far is not to refute the *Bs* account but to show that it has been misconceived. What we need to do is to view '$a = b$' not as specifying a relation between mere figures, but between signs, individuated 'as signs', thus, in part, by the linguistic use to which they are put. The question is *exactly* how to understand this. We had better not understand the 'use' of a name purely in terms of standing for a particular object. If we do that, we will treat the two signs '*a*' and '*b*' as essentially the same, differing only by an arbitrary choice of figure, when '$a = b$' is true, and we will be left with precisely the view that '$a = b$' never expresses any proper knowledge. As the two signs would differ only as figures, the only

[48] Compare Abraham Lincoln's answer to the question 'If you called a tail a leg, how many legs would a donkey have?' – 'Four. Calling a tail a leg doesn't make it one.' My appeal to 'possible worlds' here is anachronistic but inessential.

knowledge we would express by '$a = b$' would be the knowledge that the figures **a** and **b** have been put to the same use. So we have to individuate names not only through their use as standing for an object, but through some other feature of their use as well. Frege now reiterates the same reworking of the Kantian form–matter distinction as that given in *Bs*: a sign is to be individuated by the '*manner* in which it designates something', the 'mode of presentation [*Art des Gegebenseins*, mode of being given] of the thing designated'.[49] As in *Bs* he provides a geometrical example, involving two 'different designations for the same point', which 'likewise indicate the mode of presentation'. Hence 'the statement [of identity] contains actual knowledge'. (*S&B*, p. 158).

We can now view the *Bs* account of identity in this way: '$a = b$' is to be explained as $\exists x(D('a', x)\ \&\ D('b', x))$, where '$D(x,y)$' says that sign x designates y. 'a' and 'b' are here individuated as signs, (in part) by their associated modes of presentation, and so we need not include any mention of the user. Moreover, Frege assumes that we cannot have two objects given in the same way, by the same 'mode of presentation'. It follows that if we were to use the figure *Julius* to refer to Tiberius, we would have a different mode of presentation, and hence a different *sign*. We literally could not use the *sign* 'Julius' to refer to Tiberius. Thus the relation of *sign* to object is no longer a matter of arbitrary choice. Hence it is possible to express proper knowledge through an identity statement understood in this way. Yet we are still working with the *Bs* account; all we have done is to carefully spell out that account so as to avoid some obvious difficulties with it.

Having explained the reasons lying behind the *Bs* view of identity, Frege immediately moves into a discussion of the sense–reference distinction, to which we will turn in a moment. But it is noteworthy that his opening paragraph leaves some big questions dangling. Is identity, after all, a relation between names? And how *exactly* do we account for the specific cognitive value of identity statements (not merely the possibility of their expressing real knowledge, but the specific content which makes them into useful discoveries)? Frege

[49] While in *S&B* he introduces the terminology of 'mode of presentation', after *S&B* he sometimes reverts to talk of 'modes of determination' (Frege to Jourdain, 1914, *WB*, p. 128/*PMC*, p. 80; Frege to Russell, 28 December 1902, *WB*, p. 234/*PMC*, p. 153). The terminology of *S&B*, literally 'mode of being given', 'Art des Gegebenseins', harks back to the *Gl*'s question, 'How are numbers given to us?'

returns to these issues only very briefly at the end of the essay and even then they are not entirely resolved. The explanation of this must wait on a full-fledged spelling out of the advances in Frege's conception of content, of the act of judgement, and so of knowledge itself. These advances come on two fronts: the treatment of the content of names, and the treatment of the content of sentences. Together they are designed to provide us with a unified picture which both builds on and replaces the original *Bs* account.

6 'ON SENSE AND REFERENCE': NAMES

Frege follows up his discussion of the *Bs* theory of identity by introducing the terminology of sense and reference. He says that we can associate with each name, in addition to the object named, its 'reference', a 'sense' in which 'the mode of presentation is contained' (*S&B*, p. 158). This move reflects the *Bs* suggestion that the content of a judgement of identity is 'that the same content … is actually given by two modes of determination' (*Bs*, §8, p. 125). That idea, when pressed, yielded the view that the parts of the content of '$a = b$' are *not* the names 'a' and 'b', but the modes of determination with which they are associated. '$a = b$' turns out to be *about* these modes of determination; it is they rather than the objects named which are what the names 'a' and 'b' stand for in this context. *This* thought is not retained in the mature picture. What is retained is the idea that it is the modes of determination/presentation of the objects named which are the parts of the content, or *sense*, of the whole sentence.

In introducing the sense–reference distinction for names, Frege began with the *Bs* model in which the content of a name is the object that it names, while the content of a sentence is individuated in terms of its consequences. This model is in tension with the principles of compositionality and substitution, as is revealed by the problem of identity. The *Bs* theory of identity resolved this tension through a special reinterpretation of the identity sign. Careful examination of this theory has now shown that there is a semantically significant aspect of names over and above the 'content' or object named, which is not reducible to the arbitrary choice of a particular figure to name that object. This feature Frege identifies as the mode of presentation of the object. Since the 'content' of names, as object named, turns out to be inappropriate to serve as a constituent of the 'content' of sentences as individuated by consequences, Frege searches for a feature of names which can be more suitably pressed into that role. Within the confines of the basic *Bs* picture, nothing *could* be pressed into that role other than the mode of determination/presentation of the object named. This is a reflection of the fact that, for the Frege of *Bs*, names refer 'directly' to objects, although they may do so in different ways.[50]

Frege never goes beyond the metaphor of 'mode of presentation' to provide a more detailed theory of sense. He does, however, make several remarks which clarify the role of the notion of name-sense in his account of language.[51] The sense of a name, he tells us, is grasped by anyone sufficiently familiar with the language to which the name belongs. Ideally, to each sign in a given language there will correspond exactly one sense, and to each sense exactly one reference. However, in natural language, ambiguity is possible – the same word can have different senses in different contexts.[52] Similarly, it is possible to form expressions which possess a sense but no reference, for example 'the least rapidly converging series' (*S&B*, p. 159). Frege claims that in a proper scientific *Begriffsschrift*, every proper name must have a reference; in *Gg* he attempts to prove this of his own system of notation[53] (*Gg*, vol. I, p. xii; §§28–31, pp. 45–50). It is therefore a 'fault'of language to permit the formation of names with sense but no reference, just as it is a fault of language to allow ambiguity (*S&B*, p. 168).[54]

[50] The supposed opposition between direct and Fregean theories of reference is not so clear as it is often taken to be.

[51] However, he waffles on several of these claims. I discuss the significance of this below.

[52] Here, clearly, words must be understood as individuated *as objects* rather than *as signs*.

[53] Frege reiterated the possibility of sense without reference several times in his correspondence with Russell, after the latter's discovery of the paradox had led him to conclude that 'my explanations in sect. 31 [of *Gg I*] do not suffice to secure a reference for my combinations of signs in all cases'. (Frege to Russell, 22 June 1902, *WB*, p. 213/*PMC*, p. 132.)

If he had rejected the possibility of sense without reference, it would have followed that in *Gg* he had neither expressed any thoughts nor carried out any reasoning. Clearly, he may have found this evaluation of his life's work unattractive.

[54] This analogy is weak, however. If there *are* senses without reference, it can hardly be a 'fault' to express them. Rather, a language which could not express them would be impoverished.

Frege emphasizes that the sense of a name is expressible in different languages, and is objective. Thus, it is distinguished from the representation (*Vorstellung*) called up by the name. The latter is subjective, private and psychological, the former is objective and communicable. He provides an image to elucidate this distinction (*S&B*, pp. 160–1):

> 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 lens [*objective Glas*] in the interior of the telescope, and by the retinal image of the observer. The former I compare to the sense, the latter is like the representation or intuition [*Vorstellung oder Anschauung*]. The optical image in the telescope is indeed one-sided and dependent on the standpoint of observation; but it is still objective, inasmuch as it can be shared by several observers … But each one would have his own retinal image.

This analogy is meant to highlight the shareability and objectivity of the notion of sense, but it has another important feature. The 'real image' in the telescope is not a 'third thing' intervening between the observer and the star which she sees 'through' it. Similarly, sense as 'mode of presentation' need not be seen as a 'third thing' intervening between speaker and reference. The idea that the 'way in which the object is given' is a semantically significant feature of a name need not detract from the idea that objects are given to us 'directly'.[55] Thus worries, such as troubled Russell, that the notion of sense leads to sceptical doubts like those generated by representationalist theories of knowledge of classical British empiricism may be misplaced.[56] Yet this way of thinking about sense is in tension with Frege's repeated commitment to the possibility of sense without reference. If sense is simply the way in which an object is given to us, and not an entity intervening between us and the object, the possibility of sense without reference becomes mysterious. The interpretation of Frege developed by Gareth Evans and John McDowell takes off from this observation. According to Evans and McDowell, Frege's 'better self' would reject the possibility of sense without reference. While

[55] Alasdair MacIntyre made this point to me in conversation.

[56] For more on Russell's reaction to Frege see Peter Hylton, chapter 13, this volume, and Michael Kremer, 'The argument of "On Denoting"', *Philosophical Review*, 103 (1994), pp. 249–97.

there is much to this line of thought, the full force of this issue can only be appreciated after we have developed the other side of Frege's sense–reference theory, his account of the sense and reference of sentences. As we will see, the possibility of sense without reference plays a crucial role in Frege's arguments concerning the reference of sentences, yet sits poorly with the resulting account of judgement and truth.

7 'ON SENSE AND REFERENCE': SENTENCES

After elucidating the sense–reference distinction for names, Frege turns to the 'content' of declarative sentences, which we use in assertion to express acts of judgement.[57] He begins by stating that a sentence contains (*enthält*) a 'thought' (*Gedanke*), by which he means 'not the subjective performance of thinking but its objective content (*Inhalt*), which is capable of being the common property of many thinkers' (*S&B*, p. 162, fn. 5). I take this to be a reference to his earlier doctrine of 'judgeable content', individuated in terms of consequences.[58] At the same time, Frege continues to use 'content' informally to refer to the reference, or object named, when speaking of a proper name. Thus, in 'On Euclidean geometry' (1899–1906?), he speaks both of the sense of a sentence as its content and of the 'confusion of numerals and numbers' as an example of the confusion of a 'sign and its content' (*NS*, p. 182/*PW*, p. 167; examples could be multiplied). Here the continued influence of the *Bs* model is palpable. Frege then argues for the following claims (*S&B*, pp. 162–4):[59]

[57] The following discussion is heavily indebted to Burge, 'Frege on Truth', in L. Haaparanta and J. Hintikka (eds.), *Frege Synthesized: Essays on the Philosophical and Foundational Work of Gottlob Frege* (Dordrecht: Reidel, 1986). and to the writings of Tom Ricketts, especially 'Objectivity and objecthood', in *ibid.*, pp. 65–95.

[58] In his later writings, Frege frequently speaks of the 'thought-content' (*Gedankeninhalt*) of sentences, and uses 'content' and 'thought' interchangeably ('Foundations of geometry, second series, I', p. 294, among many other examples). In particular, he often speaks of sentences as having 'generality of content' insofar as they have a range of particular sentences as consequences (*ibid.*, p. 307; again, among many examples.) This phrase, which harks back to his earlier account of judgeable content in terms of consequences, first occurs in his writings in the early 1880's, in 'Boole's logical Calculus and the Concept-script'. (*NS*, 11, fn. ***/*PW*, 11, fn. ***.)

[59] My presentation follows closely that of Burge, 'Frege on Truth'.

(1) the thought expressed is the sense of a sentence, not its reference;
(2) a sentence must in some cases have a reference as well as a sense;
(3) the reference of a sentence is a 'truth-value';
(4) truth-values are objects, and sentences, the names of those objects.

His arguments show that he has decided in advance that the principle of substitution must apply to both sense and reference, insofar as these notions can be applied to complex linguistic expressions. His procedure follows the pattern exhibited for names – starting with the picture of names as standing for objects and sentences as expressing thoughts, Frege sought something 'on the level of thoughts' associated with names, and now he seeks something 'on the level of objects' associated with sentences, in each case guided in part by the principle of substitution. But his argument also shows the special character of sentences as those expressions through which we express acts of judgement, therefore acts at least potentially of knowledge; this dimension of sentences plays a dominant role in the argument especially for (2)–(3). This belies any simple version of the common charge that Frege made a decisive mistake in rejecting the fundamental categorical distinction between sentences and names, judgeable and unjudgeable contents, central to his earlier thought.[60]

Frege argues for (1) in two steps. First, he argues (1a) that the thought expressed by a sentence is not its reference; second, he concludes immediately (1b) that the thought is the sense. This conclusion reflects a determination to make do with the two categories of sense and reference. This should not be surprising – name-sense was introduced as something which could serve as the name-counterpart of sentential content/thought. We see here a fundamental reorganization of the *Bs* picture: the *Bs* view that the object named, the 'content' of the name, is part of the thought expressed, the 'content' of the sentence, is replaced by the parallel, yet fundamentally different, view that the sense of the name (the mode of presentation of the object named) is part of the sense of the sentence (the thought expressed).

[60] This point is made by Burge, *ibid*.

Frege's argument for (1a) turns on the substitution principle:

(i) there are cases in which names '*a*' and '*b*' have the same reference, but sentences '*Fa*' and '*Fb*' do not express the same thought;
(ii) if '*a*' and '*b*' have the same reference, and '*Fa*' and '*Fb*' have a reference, '*Fa*' and '*Fb*' have the same reference (by the substitution principle);
(iii) therefore, the thought expressed is not the reference of the sentence which expresses it.

The crucial step in this argument is (i). This claim fits well with our discussion of the problem of identity – if substitutions of names with the same reference always preserved the thought expressed, there would be no more content to '$a = b$' than to '$a = a$', since the supposed additional content of allowing such inferences as that from '*Fa*' to '*Fb*' would reduce to a matter of the trivial verbal reformulation of the same thought. Frege argues to the contrary that '*Fa*' and '*Fb*' in such cases need not express the same thought because one who did not know that $a = b$ might hold the one to be true and the other to be false.

This might seem to be a simple application of the indiscernibility of identicals, but such an argument would fail. One who holds '*Fa*' to be true and '*Fb*' false might also hold '*Fb*' true and '*Fa*' false, albeit unwittingly, if these express the same thought. What is crucial here is the claim that one who did not know that $a = b$ might hold '*Fa*' true while *failing* to hold '*Fb*' true. Frege puts it this way in a letter to Russell (Frege to Russell, 21 May 1903, *WB*, p. 240/*PMC*, pp. 157–8):

> Now the thoughts contained in these sentences are evidently different; for after having recognized the first as true, we still need a special act to recognize the second as true. If we had the same thought there would be no need for two acts of recognition, but only for a single one.

Thus '*Fa*' and '*Fb*' do not express the same thought because one who holds '*Fa*' to be true need not immediately, and without a special cognitive act, hold '*Fb*' true as well.[61]

[61] Compare the criteria for the individuation of thoughts discussed at the end of §1.

Having established (1) that the thought expressed by a sentence is its sense, not its reference, Frege goes on to argue (2) that at least some sentences have a reference as well as a sense. His argument for (2) can be summarized roughly as follows.[62]

(i) There are contexts, such as fiction and poetry, in which we are not interested in the references of our words, but only their sense; in such contexts we require only that our sentences express a thought, not that they have a reference.
(ii) We are led to ask after the references of our words only in the context of the search for *truth*.
(iii) Therefore, there must be something associated with each sentence, which (a) depends on the references of the words making up the sentence, and (b) accounts for our interest in the references of words in the context of the search for truth.
(iv) This feature of a sentence can be called its reference, since it depends on the reference of the words making up the sentence.

The contrast Frege draws between poetry and fiction, on the one hand, and the search for truth, on the other, allows him to claim that there is something distinctive about the second case, which requires the introduction of reference for sentences as well as names. Frege claims that in contexts of fiction and poetry we care only about the thoughts expressed by our sentences. He argues that these thoughts do not depend on the references of our words, so that even if those words were to lack reference altogether, our sentences would still express thoughts, so long as our words had a sense. For example, the thought expressed by the sentence 'Odysseus was set ashore at Ithaca while sound asleep' 'remains the same whether "Odysseus" refers to something or not' (*S&B*, p. 163). If we were to discover that Odysseus

[62] This argument is repeated by Frege on several occasions; my summary draws on several versions. (Frege to Russell, 28 December 1902, *WB*, pp. 234–5/*PMC*, pp. 152–3; Frege to Russell, 21 May 1903, *WB*, p. 240/*PMC*, pp. 157–8; Frege to Russell, 13 November 1904, *WB*, p. 247/*PMC*, p. 165; 'Introduction to logic' (1906), *NS*, pp. 210–11/*PW*, p. 194; 'Logic in mathematics' (1914), *NS*, pp. 250–1/*PW*, p. 232). The possibility of sense without reference is inscribed into the heart of Frege's argument that some sentences have a reference as well as a sense. Yet this very possibility raises serious difficulties for his theory, as we will see below.

in fact existed, this would not change the thoughts expressed by the sentences in the *Odyssey* one bit. However, it would put us in a position to ask whether the sentences in the story were true – a question which we would not even care to raise so long as we were just taking the story *as* a story, as fiction or poetry.

This question of truth requires that our words have reference as well as sense, according to Frege. And it is this question which drives him to conclude that at least some *sentences* must have a reference as well as a sense. He concludes almost immediately (3) that this reference must be the 'truth-value' of the sentence, *true* or *false*, since it is only when we are inquiring after the truth-value that we are led to take an interest in references.[63] Finally, he asserts (4) that the two truth-values are *objects*, '*the* True' and '*the* False', of which sentences are *proper names*, since sentences, like names and unlike predicates and other functional expressions, are 'complete'.

This way of speaking may sound artificial, as Frege recognized, and is often taken to be a serious retrograde step, going back on *Bs*'s recognition of the crucial difference between sentences and names, judgeable and unjudgeable contents. However, Frege is engaged here in a fundamental rethinking of his account of judgement, and when this is appreciated it can be seen that in fact Frege retains a basic distinction between sentences and names insofar as it is only names of truth-values which can be used to make assertions, and so to express judgements.

8 'ON SENSE AND REFERENCE': JUDGEMENT AND IDENTITY[64]

Judgement, Frege tells us, can be regarded as the 'advance from a thought to a truth-value' (*S&B*, pp. 164–5). This is to be contrasted *not* with the earlier formula that judgement is the recognition of the truth of a content, which Frege repeats numerous times in his later writings with 'thought' replacing 'content', but with the earlier

[63] Frege also argues for this conclusion on the grounds that it is only the truth-value which is preserved under arbitrary substitutions of names with the same reference. (*S&B*, p. 164.) This argument is not convincing, however, as Burge points out – such items as 'Russellian propositions' for example, would seem to be preserved under the substitutions in question.

[64] In the following I draw especially on the work of Tom Ricketts.

conception of the *grammar* of that formula. The earlier formula seems to imply that truth is a *property* of contents or thoughts, and that judgement is a special case of the more general act of recognizing something as *F* – the special case of recognizing something as true. In the 'Logic' of the 1880s, in which Frege introduces his account of judgement as the recognition of truth, he writes of the importance of 'the *property* [*Eigenschaft*] "true"' for logic (*NS*, p. 4/*PW*, p. 4, my emphasis). But in *S&B* Frege tells us that the relation of thought to truth-value is not that of subject to predicate, but rather that of sense to reference (*S&B*, p. 164). Thus, the act of recognizing something as *F*, where *F* is an ordinary property of things, is a special case of judging, of recognizing a thought as true, rather than the other way around. Hence we cannot conceive of judging as recognizing that a thing (thought) has a property (truth). Rather judgement is 'something quite peculiar and incomparable', which we get at metaphorically as an 'advance from a thought to a truth-value' (*S&B*, pp. 164–5). To say that this is an advance from sense to reference, not from a thought to one of its properties, is a way of gesturing at the unique status of judgement, and so also of truth. To say that *the True* is an object is another way of gesturing at the unique status of truth by denying that it is a concept, the reference of an ordinary predicate. In the later 'Logic' of 1897, Frege avoids talk of a property of truth, writing instead that 'the *word* [*Wort*] "true" can be used to indicate ... a goal for logic' and emphasizing the 'peculiarity' of this grammatical predicate. (*NS*, pp. 139–40/*PW*, pp. 128–9, my emphasis).

In support of the claim that truth is not related to thought as subject to predicate, Frege points out that the ostensible subject–predicate sentence 'the thought that *p* is true' in fact says nothing more than '*p*' (*S&B*, p. 164). These two sentences, used assertorically, make exactly the same truth-claim; on the other hand, when uttered without assertoric force, neither makes a truth-claim at all. Hence, the essential truth-claim is not made by using the predicate 'true' but rather in the act of asserting. Frege concludes that 'the relation of the thought to the True may not be compared with that of subject and predicate', since 'subject and predicate ... are just elements of thought; they stand on the same level for knowledge. By combining subject and predicate one reaches only a thought, never passes from sense to reference, never from a thought to its truth-value' (*S&B*,

p. 164). The point of all this is to rethink both the nature of truth and of judgement.

If judgement is an advance from thought to truth-value, it is also an advance from sense to reference; but sub-sentential expressions such as names, as well as sentences, are said to have sense and reference. We can now see that Frege has achieved a *reorganization* of the basic *Bs* picture while retaining essentially the same *elements*. The reorganization comes through the placing of what were essentially 'name-content' and 'sentence-content' on different levels, the levels of reference and sense; but other aspects of the original picture are picked up as correlates of these so that both names and sentences have both sense and reference. What makes it reasonable to say that we have two *levels*, that truth-value is to sentence as object named is to name, or that mode-of-presentation is to name as thought is to sentence, is that the principle of substitution holds at both levels. This principle unifies the 'levels' and makes it reasonable to use one term ('sense,' 'reference') across each level. We have seen the principle of substitution for reference at work in Frege's arguments that thoughts cannot be the references of sentences. Frege later explicitly adopted the principle of compositionality for senses, and from this the principle of substitution for senses follows.[65]

In *S&B*, Frege seems to commit himself to the principle of compositionality for reference as well, suggesting that 'judgements are distinctions of parts within truth-values' (*S&B*, p. 165). Frege later apparently contradicts this, noting that the reference of a part of an expression need not be a part of the reference of the whole expression – while 'Sweden' is a part of 'the capital of Sweden', Sweden is not a part of Stockholm, for example ('Notes for Ludwig Darmstaedter' (1919), *NS*, p. 275/*PW*, p. 255). However, careful examination of his suggestion in *S&B* shows that already there he has hedged his bets so as to avoid this objection. He first says that the 'distinction of parts' within a truth-value 'occurs by a return to the thought' so that there will be a 'mode of analysis' of the True (the False) for each true (false) thought. He notes, however, that

[65] There is a much controverted issue concerning whether compositionality for senses conflicts with Frege's thesis that the same thought can be expressed in radically different ways. I do not myself think there is a conflict – metaphorically, the same thing can be divided into parts in many ways – but I will not discuss this issue here.

he has 'used the word "part" in a special sense', differing from its ordinary use. Normally, given the whole and the part, we would be able to determine a unique remainder, but given a truth-value and an object we cannot uniquely determine a concept – '2 is prime' and '2 is even' are both true, yet '( ) is prime' and '( ) is even' designate distinct concepts. Hence 'a special term would need to be invented' for this notion of part-hood (*S&B*, p. 165).

In Frege's later writings he does not continue to speak of word-referents as parts of truth-values; but the deepest point of the metaphor, and so of the idea of compositionality for reference, is retained. This is the point that in judgement, the recognition of a thought as true, we are not directed merely to a thought and a truth-value, but also to the referents of the parts of the sentence, the objects which the sentence is about and the concepts which are applied to those objects.[66] These are what he metaphorically calls the 'parts' of the truth-value. Judgement is that act in which we are directed to truth, and it is *thereby* that act in which we are led to ask for the reference of our words. Judgement requires that we grasp thoughts as *articulated* in a way that directs us to objects and concepts. In Frege's later writings he makes the same point by insisting that our sentences are *about* the objects and concepts which are the referents of the words we use. Only if our sentences are about objects and concepts can we express truths through them.

In later writings ('Logic' (1897), *NS*, pp. 139–40/*PW*, pp. 128–9; 'Thoughts' (1918–19), pp. 352–3) Frege builds on his argument that truth is not a property, to show that truth is indefinable.[67] He argues as follows. If truth were definable, we would have available some such definition as:

(the thought that p) is true = φ(the thought that p)

where φ is some perhaps complex defining phrase, for example

(the thought that p) is true = (the thought that p) corresponds to a fact.

[66] This is for the simplest case; in more complex cases we may be directed to the concepts which the sentence is about and the higher-level concepts which are applied to them.

[67] Ricketts's 'Objectivity and objecthood' is especially helpful on this argument. See also his 'Logic and truth in Frege', *Aristotelian Society, Supplementary Volume*, 70 (1996), pp. 121–40.

To use such a definition in order to determine whether the thought that *p* was true, we would have to judge whether or not φ(the thought that *p*) – we would need to recognize *this* as true, or reject it as false. This would require yet another application of the definition, yet another judgement, and so on, resulting in an infinite regress. Consequently, the purported definition cannot be used, and so is no definition at all; and truth has to be recognized as indefinable.

This has one important consequence which Frege does not explicitly draw, however. Just as truth is not a property of thoughts, so there is no 'relation of reference' between senses and referents.[68] For if there were such a relation, truth would be a *definable property* of thoughts:

> the thought that *p* is true = the thought that *p* refers to the True.

But if there is no *relation* of reference between senses and referents, to say that the relation of thought to truth-value is that of sense to reference rather than of subject to predicate is at best to say something metaphorical and elucidatory of judgement and truth, both of which are indefinable, simple and quite 'peculiar and incomparable'.[69] This highlights the fundamental place of judgement in all our talk of sense and reference. It is only in the context of judgement that the 'relation' of sense to reference is approached, not only for sentences, but for the words that make them up. It is only in the context of judgement, the striving for truth, that we become interested in the *referents* of our words. Here there is an echo of the context principle of *Gl*[70] – it is only in the act of grasping a complete thought and judging as to its truth that we take 'the step from the level of thoughts to the level of reference (the objective)' (*S&B*, p. 164). While Frege had earlier emphasized the objectivity of judgeable content, the point here is that it is only in judging that we move to the objective level of truth-evaluability, the level at which our thought makes contact with the referents of our words. The echo of the Context Principle here shows that Frege has not simply

[68] This point is made by Ricketts in 'Objectivity and objecthood'.
[69] Joan Weiner's work on the 'elucidation' of primitive terms is illuminating here.
[70] I owe this important point to Burge, 'Frege on truth'.

abandoned his insight into the difference between sentences and names, judgeable and unjudgeable contents, but has relocated it in the context of a richer theory of the act of judgement.

In *S&B*, Frege bases his argument on the claim that '*p*' and 'the thought that *p* is true' say the same thing. In other writings, he extends this claim in two ways, each of which might seem problematic. In the unpublished 'Logic' of 1897, he claims that '*p*' and 'it is true that *p*' express the same thought (*NS*, p. 153/*PW*, p. 141; also Frege to Russell, 13 November 1904, *WB*, p. 245/*PMC*, p. 163); and in 'Logic in mathematics' (1914) he comes close to claiming that '*p*' and '"*p*" is true' are similarly equivalent. We discuss these in turn.

Frege's claim that '*p*' and 'it is true that *p*' say the same thing might seem to be in conflict with his account of 'indirect discourse' contexts such as 'Jones believes that *p*'. In such contexts, according to Frege, words do not stand for their ordinary referents, but rather for their ordinary senses, which are their 'indirect referents' (*S&B*, p. 159). Thus in 'Jones believes that Smith is a fool' the name 'Smith' stands for its ordinary sense, and the subordinate sentence 'Smith is a fool' stands for the thought that Smith is a fool. One might suppose that it is the business of the word 'that' to effect this transformation, so that '*p*' in any clause of the form 'that *p*' stands for a thought.[71] In that case, we would have to conclude that '*p*' in 'it is true that *p*' stands for a thought, and it would be hard to see what work 'it is true that' could play other than to predicate truth of the thought.

Yet a careful consideration of Frege's long discussion of such cases in the second half of *S&B* should dispel this worry. Frege nowhere says that 'that' (or even 'the thought that') automatically generates an indirect discourse context. Rather, he attends to the specific behaviour of each context in which a subordinate clause is involved. Of particular interest is his treatment of contexts of the form '*S* fancies that *p*'. He claims that in such contexts we have the simultaneous expression of two thoughts: '*S* believes that *p*' and 'not-*p*'. He comments (*S&B*, p. 175):

[71] This appears to be the view of Terence Parsons in 'What do quotation marks name? Frege's theories of quotations and that-clauses', *Philosophical Studies*, 42 (1982), pp. 315–28.

In the expression of the first thought, the words of the subordinate clause have their indirect reference, while the same words have their customary reference in the expression of the second thought. This shows that the subordinate clause in our original complex sentence is to be taken twice over, with different referents: once for a thought, once for a truth-value.

Here the little word 'that' does not prevent the words in the subordinate clause from having their customary reference (albeit in addition to their indirect reference).

Frege's reasoning here is guided by two principles. First, if we cannot substitute words with the same customary reference within a larger sentential context, *salva veritate*, this shows that the words have their indirect reference; but if we *can* perform such substitutions then we should take the words to have their customary reference – this is the *default* position which failures of substitution upset; the doctrine of indirect reference is introduced only to account for such failures. Second, if we find that a truth-functional context involving '*p*' (such as 'not-*p*') is a *consequence* of a context 'φ(that *p*)' we should take it that the words occurring in this context have their customary referents. This principle reflects the ancestry of his notion of thought in his earlier consequence-driven conception of content. The two principles together account for his diagnosis of '*S* fancies that *p*'. But they also imply that in the context 'it is true that *p*' words have their customary references, and no other. We can always substitute words with the same customary reference in this context *salva veritate*, so we have no grounds for taking the words to have other than their customary reference; and, since 'it is true that *p*' has '*p*' as a consequence, we have a reason to take the words to have their customary reference. Even in the context 'the thought that *p* is true' we do not have a reference to a thought. This sentence says no more than '*p*'.

In 'Logic and mathematics', Frege assimilates the case of '"*p*" is true' to that of 'the thought that *p* is true': 'to say of a sentence, or thought, that it is true is really quite different from saying of sea water, for example, that it is salt. In the latter case we add something essential by the predicate, in the former we do not.' He concludes that this shows 'that truth is not a property of sentences or thoughts', and 'confirms that a thought is related to its truth-value as the sense of a sign is to its reference' (*NS*, p. 252/*PW*, p. 234). This discussion strongly suggests that, like 'the thought that *p* is true', '"*p*" is true' is equivalent in sense to *p*.

Yet it may seem that Frege is wrong to assimilate 'the thought that *p* is true' and '"*p*" is true' in this way. For he holds that truth is primarily ascribed to thoughts and only derivatively to the sentences that express them. ('Thoughts', pp. 353–4). This suggests the following schema, using '≈' to indicate sameness of sense:

> the thought that *p* is true ≈ *p*
>
> '*p*' is true ≈ '*p*' expresses the thought that *p*, and (the thought that *p*) is true
>
> ≈ '*p*' expresses the thought that *p*, and *p*.

Here it seems that '"*p*" is true' has content over and above '*p*', shown in the clause '"*p*" expresses the thought that *p*'.

However, bearing in mind the figure–sign distinction drawn above, we must ask whether by '*p*' we are to understand a figure or a sign. In the first case, the above schema becomes:

> **p** is true ≈ **p** expresses the thought that *p*, and (the thought that *p*) is true
>
> ≈ **p** expresses the thought that *p*, and *p*.

Here there is indeed additional content in the clause '**p** expresses the thought that *p*,' since the use of a particular complex figure to express a particular thought is a matter of arbitrary choice. Indeed, there is an intelligible sense in which a *property* of truth for propositional *figures* has been explained here, although all the *interesting* work is done by the *expression* relation between figures and thoughts. But in the second case (directly represented by the first schema above, given our decision to use quotation to name signs, not figures) the supposed additional content can be discounted; for the sentential *sign* '*p*' expresses the thought that *p essentially*. Frege states in 'Thoughts' that he uses the word 'sentence' ('*Satz*') in *this* sense, so that 'the sense necessarily goes with the sentence' ('Thoughts', p. 362; see also 'Foundations of geometry, first series, I', p. 277). Hence there is no conflict between his claim that truth is ascribed primarily to thoughts and derivatively to sentences, and his claim that ''*p*' is true' and '*p*' express the same thought.

I claimed above that, as a consequence of Frege's thesis that truth is not a property of thoughts, there is no real relation between sense and reference. However, if this is so, what can we make of such claims as that name '*a*' refers to object *b*, or that the sense of '*a*' refers to *b*? I suggest that we take our cue from Frege's argument that truth is not a property, *since* '"*p*" is true', 'the thought that *p* is true' and '*p*' say the same thing. In order to simplify the following exposition, let us adapt another of Sellars's devices, and use 'dot-quotation' to abbreviate 'the sense of the name '__',' so that '•Caesar•' abbreviates 'the sense of the name 'Caesar'.' Now consider '*a*' and 'the reference of •*a*•.' It is plausible that these expressions have not only the same reference, namely *a*, but also present this reference in the same way, that is in the way that '*a*' presents *a*. Thus, we can view these two expressions as possessing not only the same reference but also the same sense. Furthermore, if we assume that the name '*a*' is individuated *as a sign* and not as a figure, '*a*' will have its sense essentially, and so we can conclude that all three expressions '*a*', 'the reference of "*a*"' and 'the reference of •*a*•' express the same sense. Finally, since '"*a*" refers to *b*' can be rephrased as 'the reference of "*a*" = *b*,' this expresses the very same thought as '*a* = *b*'.

Frege at one point ('Comments on sense and reference', 1892–5) makes a similar suggestion in connection with his well-known difficulties with talking about concepts.[72] He holds that 'the concept Φ' refers to an object, not a concept, because it is a complete expression, and lacks the predicative nature of a concept-word. He notes that 'the reference of the concept-word *A*' is in the same boat, and concludes (*NS*, p. 133/*PW*, p. 122):

> Indeed, we should really outlaw the expression 'the reference of the concept-word *A*', because the definite article before 'reference' points to an object and belies the predicative nature of a concept. It would be better to confine ourselves to saying 'what the concept word *A* refers to', for this at any rate is to be used predicatively: 'Jesus is, what the concept word "man" refers to' *in the sense of* [*in dem Sinne von*] "Jesus is a man".'

What is noteworthy for our purposes here is the suggestion that 'is a man' and 'is, what the concept word "man" refers to' are equivalent not only in reference but in *sense*.

[72] For more on this see Ricketts, chapter 6, this volume, §§5 and 6.

We can now return to our worries about Frege's opening discussion of identity in *S&B*. At the end of that discussion, Frege left the initial question whether identity is a relation between names or between objects unresolved. Moreover, he gave no detailed account of the specific cognitive value of identity statements, which played such a central role in his argument.

Our account of the so-called 'reference-relation' helps with the first of these worries. Frege had asked whether identity related objects or names, citing the *Bs* as holding the latter view. We saw that the formal system of *Bs* implicitly treats identity as relating objects, and that the project of *Gl* similarly requires that identity relate objects, not merely names. Frege's later explanations of the identity sign sometimes seem to confirm that this is his view. For example, in *Gg* he introduces his identity sign in this way: '"Γ = Δ" shall denote the True if Γ is the same as Δ; in all other cases it shall denote the False' (*Gg*, vol. I, §7, p. 11). In 'Comments on sense and reference', he states that 'the relation of equality, by which I understand complete coincidence, identity, can only be thought of as holding for objects, not concepts' (*NS*, pp. 130–1/*PW*, p. 120). Yet in his later writings Frege also makes statements that seem to fit better with the *Bs* view. Earlier in *Gg*, for example, he says that 'if I wrote "(2 + 3 = 5) = (2 = 2)" … I should only have designated the truth-value of *"2 + 3 = 5"s denoting the same as "2 = 2"*.' (*Gg*, vol. I, §5, p. 9). And in 'Comments on sense and reference' he treats 'The reference of the word "conic section" is the same as that of the concept-word "curve of the second degree"' and 'The concept *conic section* coincides with the concept *curve of the second degree*' as interchangeable (*NS*, p. 131/*PW*, p. 120).

One might be disposed to dismiss such statements as hand-waving pedagogy in the context of informal exposition. But this thought is undercut by a striking example in which Frege seems to assert the *Bs* view *in the midst* of discussing the importance of distinguishing use and mention. Frege opens 'Function and concept' (1891), the first published work in which he deploys the sense–reference distinction, by criticizing views which confuse 'form and content, sign and thing signified'. He sees such confusion as underlying the view that '2 + 5 and 3 + 4 are equal but not the same' – 'difference of sign' is transferred to 'difference of thing signified'. Against this, he says: 'What is expressed in the equation '$2 \cdot 2^3 + 2 = 18$' is that the right-hand complex of signs has the same reference as the left-hand one' ('Function and concept', p. 138).

In *S&B* the opening question as to whether identity is a relation between names or between objects is never actually answered. The apparent conflict is now easily resolved, however. The sentences '$a = b$', 'the reference of 'a' = the reference of "b"', and 'the reference of •a• = the reference of •b•' not only have the same truth-value, they express the same thought. Moreover, in spite of linguistic appearances, none of these sentences states a relation between anything but the objects a and b, just as none of the sentences 'p', '"p"is true' and 'the thought that p is true' expresses anything but p.

We can now return to the worry raised by Frege's initial introduction of sense as a response to the problem of identity. The problem was to explain how the two sentences '$a = a$' and '$a = b$' can differ in content, 'cognitive value', while expressing the same relation between the same things. The basic answer is that while the two names 'a' and 'b' have the same reference, they may differ in sense, by presenting the same object in different ways. Frege explains at the end of *S&B*: 'If we found '$a = a$' and '$a = b$' to have different cognitive values, the explanation is that for the cognitive value (*Erkenntniswert*) the sense of the sentence, viz., the thought expressed by it, is no less relevant than its reference, i.e. its truth-value [*Wahrheitswert*].'[73] When the sense of 'a' differs from that of 'b', 'the thought expressed in "$a = b$" differs from that of "$a = a$"… the two sentences do not have the same cognitive value. If we understand by "judgement" the advance from the thought to its truth-value… the judgements are different' (*S&B*, pp. 176–7).

To fully appreciate Frege's point here, we must recognize the intimate connections between *judgement*, *inference* and *knowledge* in his thought.[74] To judge is to recognize (*anerkennen*) the truth of a thought. The German term '*anerkennen*', like its English translation,

[73] *CP* translates the second occurrence of '*Erkenntniswert*' in this sentence as 'for the purpose of acquiring knowledge.' But I think it is important to see that Frege is both linking and contrasting truth-*value* (*Wahrheitswert*) and cognitive *value* (*Erkenntniswert*) here.

[74] Here again I follow Ricketts, especially 'Logic and truth'.

points to a link between judgement and knowledge.[75] Judgement is essentially oriented towards truth: the truth of the thought that is recognized to be true is internal to the act of judgement, as its *goal*. But truth is not the only goal of judgement. Judgement aims, finally, at *knowledge*, and for this truth is not enough. For the purposes of knowledge, 'the thought expressed … is no less relevant than [the] truth-value'. 'Cognitive value' (*Erkenntniswert*, value for knowledge) and 'truth-value' (*Wahrheitswert*) are both species of 'value' in the sense of a goal or end.[76]

Inference, for Frege, is a process whereby we make judgements on the basis of other judgements; there is no such thing as inference from a mere assumption or hypothesis. The role of inference is to justify some of our judgements on the basis of others. To say that one thought has another as a consequence, then, is to say that anyone who recognizes the first to be true would be in a position to justifiably recognize the second to be true, by an inference from the first. Thus inference, like judgement, is intrinsically ordered to truth and knowledge. This will imply, as a necessary condition, that if one thought has another a consequence, it is impossible for the first to be true and the second false; but it is implausible to attribute to Frege the idea that this is also a sufficient condition. For one thought to have another as its consequence, the drawing of the inference from the first to the second must *justify* the recognition of the truth of the second. But as we have seen, though for any two logical laws, it is impossible that the first be true and the second false, nonetheless some logical laws are *basic* and others not; such basic laws will not be *consequences* of non-basic laws on this account.

Now, if judgement is the advance from a thought to a truth-value, when we advance from different thoughts to the same truth-value, our acts of judgement are different, and so the knowledge acquired

[75] The tightness of the link can be disputed; there is a question whether one can 'recognize the truth' of a false thought – judge incorrectly. I assume here that this is possible. In this I disagree with Ricketts ('Logic and truth'). See my 'Judgment and truth in Frege'.

[76] Gottfried Gabriel and Hans Sluga have made clear the connection of 'truth-value' to neo-Kantian 'value theory'. See Sluga, 'Frege on meaning', in H. J. Glock (ed.), *The Rise of Analytic Philosophy* (Oxford: Blackwell, 1997), and Gabriel, 'Frege als Neukantianer', *Kantstudien*, 77 (1986), pp. 84–101.

is different. Thus we can see how the difference in sense between '*a*' and '*b*' can translate into a difference in 'cognitive value' between '*a* = *a*' and '*a* = *b*'. Moreover, we can begin to see how to understand the *specific* difference in cognitive value, the difference in *content*, in *consequences*, between '*a* = *a*' and '*a* = *b*.' The question is, fundamentally, why given '*a* = *b*' we can infer from '*Fa*' to '*Fb*,' whereas given '*a* = *a*' we cannot.

Suppose that one has recognized the truth of '*a* = *b*'. One has thereby grasped a thought which has as parts two senses, •*a*•, the sense of '*a*', and •*b*•, and recognized the truth of that thought. Now suppose further that one has judged that *Fa*. Here one has grasped another thought, which has as a part •*a*•, and recognized its truth. Further, suppose that one has also grasped the thought that *Fb*, which has as a part •*b*•. One can see here at least the beginnings of an explanation of how it is possible that all this should amount to a justification for the further recognition of the truth of the thought that *Fb*. For after all the two thoughts whose truth one has recognized are related in a familiar way to the third yet to be recognized as true, the first premise containing •*a*•, the second linking this to •*b*•, and the conclusion replacing •*a*• with •*b*• in the first. However, this purely formal explanation might seem insufficient; for after all the same could be said about any so-called inference of the form '*Fa*', '*aRb*', therefore '*Fb*'.

It is tempting to try to supplement this explanation in one of two ways. First, one might try to argue as follows. I have claimed that '*a* = *b*' and '•*a*• has the same reference as •*b*•' express the same thought. Consequently in judging that *a* = *b* we have judged that •*a*• and •*b*• have the same reference, and it is this specific link that enables us to replace the one sense with the other and infer from the judgement that *Fa* to the judgement that *Fb*. But clearly this 'explanation' is circular; it employs exactly the principle that it is intended to explain. Second, one might suggest that since •*a*• and •*b*• *give* us the very same object, when we have grasped both senses properly we will be enabled to infer from the judgement that *Fa* to the judgement that *Fb*. But this move eliminates the need for the premise '*a* = *b*' and ends up again reducing the content of '*a* = *b*' to that of '*a* = *a*'. For Frege it is crucial that the same object can be given to us in different ways without our *knowing* that this is so, in other words without our knowing that *a* = *b*.

The proper response to this situation is to give up the search for an *explanation* of the difference in inferential potential between '$a = a$' and '$a = b$'. The theory of sense and reference is misconceived if it is thought of as providing an explanatory account of the facts concerning inferences and identity. More generally, we should refuse to assign explanatory priority to any of the various concepts of truth, judgement, inference, thought, object, mode of presentation, naming and so on. This does not mean that the theory is left without a point or a purpose, however. For the theory displays in its categories – name, sentence, sense, reference, thought, truth-value – essential general structures of our acts of judging and asserting. These acts are taken as central – in accordance with the Context Principle. Thought, truth-value, sense, reference, are immanent, interdependent aspects of our cognitive activity. On this sort of view, the theory of sense and reference, in providing the formal demonstration above that the inference '$a = b$, Fa, therefore Fb' need not be a mere case of repetition of the premise in the conclusion, makes clear the logical place which identity judgements can occupy. It does not provide a demonstration that this place is filled. *We* fill that place by taking identity as a primitive sign whose sense and reference we grasp in employing it in accordance with the principle of substitution that governs it.

Thus we should not view the theory of sense and reference as explaining the phenomena governing the inferential interrelations of identity judgements. Nor should we, however, treat inferential relations among thought contents as primary, expecting this to secure for us a world of objects. Rather, we should simply take it that to speak of inferential relations involving identity statements, and to speak of the objects to which we refer and the senses through which we refer to them, are different ways of saying the same things. This is the ultimate resolution of the problem of identity. While Frege pointed the way to such a resolution, however, he never managed to follow it out systematically.

9 PROBLEMS AND PROSPECTS

Frege claims at times that the theory of sense and reference provides an *explanatory account* of important features of thought and language. In a letter to Jourdain, he writes that 'the possibility of our understanding sentences which we have never heard before rests evidently on this, that we construct the sense of a sentence out of parts that correspond to the words. … Without this, language in the proper sense would be impossible' (Frege to Jourdain, 1914, *WB*, p. 127/*PMC*, p. 79). Similarly, in correspondence with Peano, he seeks to '*explain* how it is possible that identity should have a higher cognitive value than a mere instance of the principle of identity', and adds that 'my distinction between sense and reference comes in in an illuminating way' (Frege to Peano, undated, *WB*, p. 196/*PMC*, p. 127, my emphasis).

Frege's explanation of the possibility of understanding new sentences is spelled out in his unpublished 'Logic in Mathematics' of 1914 (*NS*, 243/*PW*, 225):

> It is marvellous what language achieves. By means of a few sounds and combinations of sounds it is able to express a vast number of thoughts, including ones which have never before been grasped or expressed by a human being. What makes these achievements possible? The fact that thoughts are constructed out of building-blocks [*Gedankenbausteinen*]. And these building-blocks correspond to groups of sounds out of which the sentence which expresses the thought is built, so that the construction of the sentence out of its parts corresponds to the construction of the thought out of its parts. And as we take a thought to be the sense of a sentence, so we may call a part of a thought the sense of that part of the sentence which corresponds to it.[77]

Frege appeals here to his principle of compositionality for senses, made explicit in the last sentence. However, he puts a particular construction on that principle, shown in his metaphor of 'building blocks'. This metaphor suggests a conception of the parts out of which the thought is 'constructed', as independent of and prior to that thought. The Context Principle, however, requires a different conception of the relation of part to whole, according to which, as Aristotle puts it, 'a whole is necessarily prior to its parts' (*Politics*, 1253a20). Thus, the drive for an explanatory use of the sense–reference distinction leads away from the Context Principle here.

One possible motivation for the attempt to provide an explanatory account of the difference in inference potential between '$a = a$' and '$a = b$' is the worry that, if we fail to do so, we will be left with a picture of thoughts as individuated solely by their logical interconnections,

[77] The translation is Dummett's (*IFP*, p. 262).

and so with a coherentism which cannot explain the way in which truth is independent of sense. Purely on the basis of the inferential interconnection of thoughts we cannot fund a conception of these thoughts as being determinately *about* specific objects. By treating name-sense as the 'mode of presentation' of an *object*, we might hope to avoid such a coherentism and establish an independent link between language and the world. In this way our model of sense will capture the normative dimension of the responsibility of our judgements and assertions to the world, of *truth*. The *Bs* model dealt with this by building the objects about which we judge directly into the contents of judgement, but failed to respect the inferential proprieties governing the contents of our judgements. The new model retains an intimate link between the thought which we judge true or false and the objects about which we think, without threatening to disturb the facts about inference. In this scheme, however, the link to the world becomes explanatorily fundamental – name-sense, in presenting us with objects, is the underlying phenomenon which explains our capacity to judge in truth-evaluable ways; and truth-evaluability becomes the fundamental phenomenon in terms of which norms of inferential propriety are to be explained.[78]

[78] I have neglected a crucial element of Frege's account here, namely the *concepts* which we apply to the objects about which we think. The unpublished 'Comments on sense and reference' establish that Frege intended to apply the sense–reference distinction to functional expressions such as predicates, as well as names and sentences (*NS*, pp. 128–36/*PW*, pp. 118–25). The reference of a functional expression is a function from argument-referents to value-referents, so that in the case of a predicate, the reference is a concept, or function from objects to truth-value. Any discussion of Frege's account of the truth of thoughts must take into consideration not only the objects that the thoughts are about but also the concepts that those thoughts involve. There is a debate, however, over the nature of the *senses* of functional expressions. Dummett takes them to be modes of presentation of functions, while Geach takes them to be functions from argument-senses to value-senses. See Michael Dummett, *Frege: Philosophy of Language* (London: Duckworth, 1973) (hereafter *FPL*), and *IFP*; Peter Geach, 'Names and identity', in S. Guttenplan (ed.), *Mind and Language* (Oxford: Clarendon Press, 1975); for an illuminating discussion see Peter Sullivan, 'The functional model of sentential complexity', *Journal of Philosophical Logic*, 21 (1992), pp. 91–108.

There is much to say on this issue, but space does not permit an extended discussion. I will only offer a brief suggestion here: the two competing interpretations derive from the two sources of the notion of sense – the idea of thoughts as inferentially individuated sentential contents, and the idea of name-senses as modes of presentation of objects. Emphasis on the first idea leads to Geach's

Once again, however, the drive to put the sense–reference distinction to explanatory work leads to a certain backing away from the Context Principle. Names function in language by giving us objects, and although every object must be given in a particular way, it is not obvious that this involves any link to judgement or thought. As *Gl* would lead us to predict, this involves Frege in a residual form of psychologism. When we seek to found the notions of sense and thought on the idea of 'modes of presentation', taking this to be prior to and explanatory of the use of names in a system of inferentially interrelated sentences, these 'modes of presentation' become hard to distinguish from the psychological representations or ideas (*Vorstellungen*) accompanying the use of words. The worry then arises that an object is, after all, given to different individuals in different ways, even if those individuals use the same word for that object.[79]

Paradoxically, this psychologizing of sense arises in part from a move designed to secure the objectivity of thoughts. Frege tended to see the shareability and communicability of thoughts as intelligible only on a model in which thoughts are transcendent entities independent of all human activities of judging, asserting or thinking. In the 'Logic' of the 1880s this idea is already prefigured: 'What is true is true independently of the person who recognizes it to be true. What is true is *therefore* not a product of a mental process of inner act; for the product of one person's mind is not that of another's' (*NS*, p. 3/*PW*, p. 3, my emphasis). Frege's favoured metaphor for understanding, 'grasping' thoughts, also emphasizes this independence: 'What is grasped, taken hold of, is already there and all we do is take possession of it' ('Logic' (1897), *NS*, p.149/*PW*, p. 137). In 'Thoughts' (1918–19) he speaks of thoughts as occupying a 'third

interpretation while emphasis on the second supports Dummett's. The necessity of seeing these positions as competing and exclusive is questionable, however. The appearance of necessity derives from the seeming need to choose one of these ideas as explanatorily more basic. This choice does not have to be forced if we conceive of thoughts and senses as immanent in the activity of judging rather than transcendent Platonic entities.

79 The thought that Frege falls into psychologism in trying to make the notion of name-sense do explanatory work derives from John McDowell's 'On the sense and reference of a proper name', *Mind*, 86 (1977), pp. 159–85. My discussion is inspired by this thought of McDowell's as well as by his *Mind and World* (Cambridge Mass.: Harvard University Press, 1996).

realm' distinct from the external world of things and the internal world of ideas ('Thoughts', p. 363).[80]

Yet this conception of thoughts as existing in magnificent isolation from human activities of judging, speaking and acting in the world raises the worry that these same thoughts might have nothing to do with the world of objects. Hypostasized thoughts may stand in ideal logical relations to one another, but our capacity to recognize them as true or false is rendered mysterious by this picture. The idea that 'modes of presentation' of objects are *constituents* of thoughts can seem to provide the needed explanation and grounding of the truth-evaluability of thoughts. But modes of presentation, ways of being given, inevitably are modes of presentation *to us*, ways of being given *to us*. Taken as explanatorily basic they cannot be in turn elucidated in terms of inferential relations between the thoughts in which they occur, and our conception of them, if it is to have any content, drifts in a psychologistic direction. Thus we end up with a view of sense involving a curious mismatch between sentence-senses, substantial 'thoughts', objects occupying a 'third realm' of their own and name-senses, insubstantial 'modes of presentation', which are either nothing at all in their own right, or become reduced to occupants of the inner world of ideas. Frege is driven on the one hand to make thoughts, and so senses, into something too objective, and the other hand, senses, and so thoughts, into something too subjective.

Thus Frege's *psychologizing* of sense shows itself in his wavering over the *objectivity* of sense. On the official view, each sign in a language has a sense which is objectively determined by the structure of the language and which is shareable by all who have a sufficient mastery of the language. The sense of an expression is contrasted with the representation associated with it: the latter is private, subjective, and can vary from speaker to speaker, whereas the former 'may be the common property of many people, and so is not a part

[80] Beaney calls this view of thoughts 'semainomenalism'. He sees it as a departure from Frege's earlier view of contents, 'best characterized … as "states of affairs", constituents of the temporal first realm' (*FMS*, p. 217). This reading of the early notion of content reflects the way in which objects figure as parts of such contents. However, the inferential individuation of such contents points in another direction. The discussion of the objectivity of content in 1880s 'Logic' prefigures many of the features of the later explicit conception of a 'third realm'.

or a mode of the individual mind'. Frege argues that 'one can hardly deny that mankind has a common store of thoughts which is transmitted from one generation to another. In the light of this one need have no scruples in speaking simply of *the* sense' (*S&B*, p. 160). Yet he takes away with one hand that which he gives with the other. In a much-discussed footnote to the passage in which he holds that anyone who has mastered the language can grasp the sense of a word, he admits that 'opinions as to the sense may differ' (*S&B*, p. 158), and in the same paragraph in which he claims that we may speak of '*the* sense' he adds that 'It might perhaps be said: Just as one man connects this idea, and another that idea, with the same word, so also one man can associate this sense and another that sense.' His response is surprisingly weak: 'But there still remains a difference in the mode of connection. They are not prevented from grasping the same sense; but they cannot have the same idea.' Here he appeals to the objectivity of sense as an *entity*, but does not try to defend its objectivity as a *feature of language*. But both forms of objectivity are needed if the fact that a 'common store of thoughts' is transmitted from generation to generation is to be intelligible. The idea of a 'common store of thoughts' is threatened when the sense of our words is reduced to the interpretation placed on them by each individual speaker.

In his late essay, 'Thoughts' (*Der Gedanke*), both the subjectivizing, psychologizing aspect of Frege's thought about sense and the objectivizing, hypostasizing aspect are on display. Frege first explains a thought as 'something for which the question of truth can arise' ('Thoughts', pp. 353–4).[81] He then takes up the expression of thoughts by sentences ('Thoughts', pp. 355–7), and particularly the fact that the thought expressed by a given use of a sentence may depend on the circumstances in which the sentence is uttered, as in the case of sentences involving indexicals like 'today', 'yesterday', 'here' or 'there'. Frege mentions specifically the case of the first-person singular pronoun: 'The same utterance containing the word "I" in the mouths of different men will express different thoughts of which some may be true, others false' ('Thoughts', p. 358). He adds

[81] Given his account of judgement as 'the recognition of the truth of a thought' we can see here the ancestry of the notion of thought in *Bs*'s notion of 'judgeable content'.

that 'the occurrence of the word "I" in a sentence gives rise to further questions' ('Thoughts', p. 358).

He considers an example: 'Dr Gustav Lauben says, "I was wounded". Leo Peter hears this and remarks some days later, "Dr Gustav Lauben was wounded". Does this thought express the same thought as the one Dr Lauben uttered himself?' Frege moves from this to a discussion of proper names, remarking that 'knowledge of the language is a special thing when proper names are involved'. He suggests that different people will express different thoughts using the sentence 'Dr Lauben was wounded', insofar as they possess different identifying information about Dr Lauben and so associate different senses with the name 'Dr Lauben'. In such cases, he says, 'as far as the proper name "Dr Gustav Lauben" is concerned,' they 'do not speak the same language'. This is due to the fact that 'with a proper name, it is a matter of the way that the object so designated is presented. This may happen in different ways, and to every such way there corresponds a special sense of a sentence containing the proper name.' It is only through a *stipulation* on our part that we can assure 'that for every proper name there shall be just one associated manner of presentation of the object so designated' ('Thoughts', pp. 358–9). He remarks that there are occasions on which it is important that such a stipulation be fulfilled.

Yet it is hard to see how to guarantee the fulfilment of such a stipulation, given his preceding argument. That argument turned on the thought that speakers possessing differing knowledge of Dr Lauben would associate different modes of presentation with the name 'Dr Lauben'. Here the sense, as mode of presentation, already takes on a highly subjective cast, and it is hard to avoid the idea that each individual might associate his or her own sense with each word, so that the public language of communication would split into a multiplicity of idiolects. When Frege returns to 'I' this thought is reinforced ('Thoughts', p. 359):

> Now everyone is presented to himself in a special and primitive way, in which he is presented to no one else. So, when Dr Lauben has the thought that he is wounded, he will probably be basing it on this primitive way in which he is presented to himself. And only Dr Lauben himself can grasp thoughts specified in this way.

Here, Frege admits that there may be modes of presentation, and so thoughts, which are unshareable and incommunicable. With this, it is hard to see any difference remaining between 'mode of presentation' and subjective 'idea' or 'representation' (*Vorstellung*). The psychologizing of sense is complete. Even though Frege goes on to say that Dr Lauben *can* communicate a closely associated thought, using 'I' in the publicly accessible sense of 'the speaker', once an incommunicable sense has been let in, the fact (if it is a fact) that some senses *are* shareable becomes a mysterious and contingent feature of human psychology.

Frege recognizes that there is a problem here. Immediately after this discussion he raises 'a doubt': 'Is it at all the same thought which first that man expresses and then this one?' ('Thoughts', p. 360). In response to this, he argues for a distinction between thoughts and subjective representations (*Vorstellungen*). His ground for this is that 'representations are something we have', which 'need an owner', and have 'only one owner', whereas 'other people can assent to the thought that I express', so that 'I am not its owner'. This shareability and communicability of thoughts is equally required if there is to be any possibility of real disagreement of dispute ('Thoughts', pp. 360–3). At this point, Frege introduces the idea of a 'third realm,' in which thoughts are to reside. I have to 'acknowledge thoughts as independent of me' if I am to acknowledge that 'others can grasp them just as much as I' and so also to acknowledge 'a science in which many can be engaged in research' ('Thoughts', p. 368). Unlike a representation, 'we do not *have* a thought'; rather we grasp a thought in thinking, but the thinker 'is the owner of the thinking, not the thought' ('Thoughts', pp. 368–9). Frege's argument here emphasizes the objectivity of thoughts as necessary for the explanation of shared human intellectual activity. Yet this argument is too little, too late; it really bypasses the doubt it was supposed to address, namely how this is *possible*, if thoughts are *made up out of* apparently subjective modes of presentation of objects.

The conception of thoughts, the items which we recognize as true in judging, as transcendent entities occupying a third realm, attempts to ground the objectivity of the norms which govern our acts of judging in a set of mythical objects. Yet these thought-objects have to bear a double explanatory burden, since our judgements are appraised normatively both in terms of their relation to other judgements and in terms of their relation to the world. We may feel forced to choose a direction of explanatory priority – either we take sentences, thought, judgement, truth and inference as explanatorily prior, or we take naming, objects and the ways in which they are given as prior. On the first option we take heed of the Context Principle and focus on the third realm of logically interconnected thoughts. But we risk coherentism – we are unable to provide a satisfactory account of the independence of truth from inferential articulation. On the second option, we try to secure this independence through appeal to our being *given* a shared world of objects, of reference; but this ends up, through neglect of the context principle, plunging us into psychologism and so undercutting any satisfying account of logical norms. Frege's thought exhibits fundamental tensions which can be explained as aspects of an oscillation between these two options. I will conclude by examining two such tensions, one having to do with our grasp of thoughts and senses, the other with the relation of sense to reference.

Frege's conception of thoughts as *objects* is implicit in his attempt in *Bs* to state criteria of identity for judgeable contents, when taken together with *Gl*'s account of name-hood, and so object-hood, in terms of occurrence in identity contexts. Thus, in *Bs*, sentences can flank the identity sign, and judgeable contents can be 'determined in different ways'. This would suggest a conception of modes of presentation of *judgeable contents*. Dummett argues on this basis against the idea that *thoughts*, like the early 'judgeable contents', are to be individuated in terms of their consequences. Dummett calls this the 'map-reference view of language' (*IFP*, pp. 43–5), and complains that this view 'demands the introduction of a third feature of sentences, one which stands to the sense of a sentence as sense stands to reference'. Thus we would be forced 'to acknowledge a third feature of sentences, which we might call their *significance*: the significance of a sentence would consist in the particular manner in which it indicated a specific thought'. For Dummett this is fundamentally anti-Fregean:

> the whole point of Frege's notion of sense is that there is no place for such a conception … A proper name, for example, stands for an object; and the particular manner in which it does this *is* its sense … there is no room for a further notion of the particular way in which the sense of the name

is picked out, because everything that belongs to the manner in which the expression functions to determine a referent is part of its sense. And what goes for a complex proper name also goes for a sentence. The sentence stands for its truth-value; and the particular manner in which it determines one or other of the two truth-values *is* the thought it expresses.

Dummett conceives of the core notion of sense as that of mode of presentation, and concludes that, as a thought is the sense of a sentence and a truth-value is its reference, a thought must be a mode of presentation of a truth-value. However, Frege never says this. On the other hand, in the unpublished 'Logic' of 1897 he comes very close to the conception of significance which Dummett castigates as clearly 'not Frege's'.

Under the heading 'Separating a thought from its trappings' Frege discusses ways in which the same thought can be presented linguistically in different forms, using *Bs*'s example of active and passive voice, as well as *S&B*'s example of '*p*' and 'it is true that *p*'. He then goes on to say:

The distinction between what is part of the thought expressed in a sentence and what only gets attached to the thought is of the greatest importance for logic. The purity of the object of one's investigation is not of importance only to the chemist. How would the chemist be able to recognize, beyond any doubt, that he has arrived at the same results by different means, if the apparent difference of means could be traced back to impurities in the substances used? There is no doubt that the first and most important discoveries in a science are often a matter of recognizing something as the same again [*Wiederkennungen*]. However self-evident it may seem to us that it is the same sun which went down yesterday and rose today, and however insignificant this discovery may seem to us, it has certainly been one of the most important in astronomy and perhaps the one that really laid the foundations of the science. It was also important to recognize that the morning star is the same as the evening star, that three times five is the same as five times three.

Here he has segued from the discussion of thoughts expressed in different verbal forms to the discussion of objects given in different ways. His reference to '*Wiederekennungen*' (reidentifications) echoes *Gl*'s description of numbers as 'self-subsistent objects that can be recognized as the same again' (*selbständige, wiedererkennbare Gegenstände*) (*Gl*, §56, p. 68), and *S&B*'s example of the 'reidentification [*Wiedererkennung*] of a small planet or comet' (*S&B*, p. 155), and even his example of the rising sun is borrowed from *S&B*'s opening paragraph. He concludes that in logic 'the first and most important task is to set out clearly what the objects to be investigated are. Only if we do this shall we be able to recognize the same as the same: in logic too, such acts of recognition probably constitute the fundamental discoveries' (*NS*, pp. 152–4/*PW*, pp.141–3).[82] Frege doesn't speak explicitly of 'modes of presentation of thoughts', but talk of thoughts as 'objects' which have to be 'recognized as the same' clearly suggests the idea. Dummett is right to see a tension here, but it is a tension within Frege's thought, not between somebody's misunderstanding of Frege and Frege. This tension arises from a potential mismatch between the two notions which generate the level of sense: the notion of mode of presentation and the notion of inferentially articulated content. This mismatch becomes serious when we hypostasize senses as entities.

This potential mismatch is also evident in a related view of Frege's which Dummett sees as in need of revision: Frege's theory of indirect sense. As discussed above, Frege held that in indirect discourse contexts, words have as their 'indirect reference' their customary sense. He further held that in such contexts words have an 'indirect sense', presumably a mode of presentation of their customary sense. Here, though, we have exactly the notion of a way in which the sense of an expression is picked out, for which Dummett said there was no room in Frege's account. Dummett therefore proposed a revision of Frege's account: in indirect discourse contexts the sense of a word remains its customary sense; it is only its reference which is changed (*FPL*, pp. 266–9). That this is a real revision in Frege's view is shown by his appeal to the doctrine of indirect senses (and indeed a whole hierarchy of doubly indirect senses, triply indirect senses and so on) in dismissing a paradox of Russell about classes of propositions (Frege to Russell, 28 December 1902, *WB*, pp. 236–7/*PMC*, pp. 153–4). In this discussion Frege shows that his conception of thoughts as occupants of a 'third realm' is at the root of the idea of indirect sense – thoughts, as *objects*, can be given

[82] Similarly, in 'Concept and object', and in correspondence with Husserl, Frege speaks of the necessity of 'recognizing a thought as the same again' (*wiederzuerkennen*). ('Concept and object', p. 185, fn. 7; Frege to Husserl, 9 December 1906, *WB*, p. 105/*PMC*, p. 70).

in different ways just as much as numbers can. Dummett's worries about this notion have their source in the dubiousness of the idea that 'modes of presentation' can be constituents of *objects* which can be 'given in more than one way'.[83]

At the same time, Frege's conception of thoughts as objects, combined with the thesis that word-senses are the 'building blocks' out of which thoughts are constructed, results in a hypostasization of name-senses which opens up the question of the representational link between such senses and their references, a question which is out of place when name-senses are thought of as 'modes of being given' the reference. The intelligibility of this question entails that there can be sense without reference, and so thoughts without truth-value. Yet, as Dummett and Ricketts[84] have pointed out, Frege's commitment to the possibility of sense without reference brings with it an ineliminable use of the truth-predicate and undercuts the tenability of the claim that '*p*', '"*p*" is true' and 'the thought that *p* is true' are everywhere interchangeable. For if '*p*' is a sentence containing a name which has sense but no reference, '*p*' expresses a thought, but has no truth-value. In this case, therefore, '*p*' is neither true nor false, but '"*p*" is true' and 'the thought that *p* is true' are false. Here 'true' and 'false' appear to function as genuine predicates. This is not a merely inconvenient consequence of Frege's acceptance of sense without reference; it undercuts the very basis of his account of the 'peculiar and incomparable' act of judgement by reintroducing the notion of truth as a property of thoughts.

Ricketts sees this problem as arising from Frege's theory of generality ('Generality', pp. 184–5). He points out that Frege's account of the generality of a universally quantified sentence amounts to this: '$\forall x \Phi x$' implies all its instances, all sentences of the form 'Φn' where 'n' is a name. Hence, if we deny any instance 'Φn' we will be committed to denying the general statement '$\forall x \Phi x$'. However, this poses a serious problem when we admit descriptions of the form 'The Ψ' as proper names, as Frege does. For suppose, as is plausible,

[83] A similar set of concerns motivates David Bell's discussion in 'Thoughts', *Notre Dame Journal of Formal Logic*, 28 (1987), pp. 36–50.

[84] Michael Dummett, *FLP*, *ILP*, also: 'Truth', in *Truth and Other Enigmas* (London: Duckworth, 1978). Thomas Ricketts, 'Generality, meaning and sense in Frege', *Pacific Philosophical Quarterly*, 67 (1986), pp. 172–95 (hereafter 'Generality').

that if 'Φ(The Ψ)' is true, then there is exactly one Ψ. Then whenever this condition is not met, we cannot affirm 'Φ(The Ψ)' – so it would seem that we should deny it. But this means that we will have to deny the generalization '$\forall x \Phi x$' as well, even if this is a logical truth like '$\forall x(x=x)$'.

Frege's solution to this problem is to deny that 'Φ(The Ψ)' *implies* that there is exactly one Ψ; this is not a consequence of the first claim, but a *presupposition* of our use of the name 'The Ψ'. When this presupposition fails, the name 'The Ψ' has no reference, and the entire sentence 'Φ(The Ψ)' lacks a truth-value. We do not *assert* that Φ(The Ψ), but we do not *deny* it either, and so the account of generality is preserved; but in saving the account of generality, we are forced to admit that there can be sense without reference, thought without truth-value.

Ricketts's diagnosis of the situation is persuasive but incomplete.[85] Frege applies his doctrine of presupposition not only to definite descriptions but to genuine proper names: 'that the name "Kepler" designates something is just as much a presupposition for the assertion

Kepler died in misery

as for the contrary assertion.' (*S&B*, p. 168). Ricketts's argument will not provide a clean explanation of this case unless we suppose that Frege held a description theory of the sense of all proper names. If we try to run Ricketts's explanation directly, we will have to put, in place of the claim that there is exactly one Ψ, the explicitly metalinguistic claim which Frege introduces:

'Kepler' designates something, i.e. $\exists x$('Kepler' designates x).

Here 'designates' appears as a relation between 'Kepler' and an object. However, we have to ask whether we are speaking of 'Kepler' as a *figure*, or as *a sign*, a figure in *use*. Is the *presupposition* of the assertion that Kepler died in misery, really (1) $\exists x$('Kepler' designates x), or (2) $\exists x$(*Kepler* designates x)? In the first case, we have employed a 'designation relation' between a sign-in-use and its reference. But, as I argue above, the doctrine that truth is not a property

[85] The argument that follows is inspired by McDowell's discussion in 'Truth-value gaps'.

of thoughts implies that there is no such relation. To admit such a relation is precisely to admit as well a real property of truth for *sentential signs*, and so also for the thoughts which those signs essentially express. If we do not admit such a relation, I have argued, then to assert that 'n' designates x is simply to assert that $n = x$. In that case, the supposed presupposition would reduce to the bare logical truth '$\exists x(\text{Kepler} = x)$'.

On the other hand, in the second case, we have only a designation relation between *figures* and objects. But, even if there is no designation relation between *signs* and references, or between name-senses and references, there is surely no reason for Frege to deny that there is such a relation between *figures* and references. For it is after all an empirical fact that we use the figure ⋆Kepler⋆ to designate Kepler; this point was central to the opening argument of *S&B*. As in the case of the truth-predicate for figures discussed above, though, the proper analysis of the designation relation for figures is:

$$\star a^{*} \text{ designates } b = \star a\star \text{ expresses } \bullet a\bullet \text{ and } \bullet a\bullet \text{ designates } b.$$

Here given that senses essentially designate their references, the second clause adds nothing, and all the interesting work goes into the expression relation between figures and name-senses. But what is important for our purposes is that the admission of such a relation does not bring with it a need for sense without reference. To deny that ⋆Odysseus⋆ designates something in *this* sense, or to claim that ⋆Odysseus was set ashore at Ithaca while sound asleep⋆ is neither true nor false in *this* sense, need not imply that we have sense without reference, or a truth-valueless thought. For after all no such conclusion follows from the claim that ⋆Suessydo⋆ fails to designate something, or that ⋆Suessydo saw tes erohsa ta Acahti elihw dnuos peelsa⋆ lacks a truth-value.

Thus, if we understand Frege's doctrine of presupposition along the lines of (2), we can defend his account of generality without having to accept truth as a real property of thoughts or designation as a real relation between name-senses and objects. Frege, however would object to this solution; for it would require assimilating the case of ⋆Odysseus⋆ to that of ⋆Suessydo⋆ – both would have to lack not only reference, but also sense. And yet, isn't it *obvious* that the two cases differ in that we *understand* the sentence 'Odysseus was set ashore at Ithaca sound asleep'? Don't we grasp in this a thought, which will 'remain the same whether "Odysseus" refers to something or not' (*S&B*, p. 163)?

To this, the proper answer is that, *if* by 'grasping a thought' we mean an act which is preparatory to *judging*, recognizing as true, then if *Odysseus* lacks a reference, we have *not* here 'grasped a thought'. Our inner experiences may be indistinguishable in this case from those which we have when grasping a thought, but to suppose that *this* guarantees that we *have* grasped a thought is to slide once again into a form of psychologism. In thus taking sense to be transparent to the 'grasping' mind, we threaten to reduce it to the level of a mere mental 'representation'. We must accept that we can be under the *illusion* of having grasped a thought, and so of being in a position to ask a question which we can proceed to answer.

Jim Conant, in 'The search for logically alien thought',[86] has pointed out that Wittgenstein, in the *Tractatus*, saw clearly the need to repudiate this residual element of psychologism in Frege's thought.

> Frege says: Every legitimately constructed proposition must have a sense; and I say: Every possible proposition is legitimately constructed, and if it has no sense this can only be because we have given no meaning to some of its constituent parts.
>
> (Even if we believe that we have done so).[87]

It is a tribute to the 'great works of Frege'[88] that even if they did not reach fully to this *Tractarian* insight, they were capable of inspiring it.

[86] James Conant, 'The search for logically alien thought: Descartes, Kant, Frege, and the *Tractatus*', *Philosophical Topics*, 20 (1991), pp. 115–80.

[87] Ludwig Wittgenstein, *Tractatus Logico-Philosophicus* (London: Routledge and Kegan Paul, 1922), 5.4733.

[88] *Ibid.*, Preface.

WILLIAM TASCHEK

8 On sense and reference: a critical reception

As influential as Frege's distinction between sense and reference has been in shaping nearly all contemporary work in the philosophy of language – as well as considerable portions of the philosophy of mind – many of its most prominent critics and proponents alike have, it seems to me, failed adequately to understand it. In consequence, they have failed adequately to assess its originality and philosophical importance. While often important and insightful in their own right, their interpretations are too often structured by commitments and concerns different from, or even alien to, those that originally motivated Frege. There is, in particular, a widespread failure to appreciate the central and controlling role that Frege's concern with and distinctive understanding of logic played in motivating and shaping the distinction. Epistemological considerations are over-emphasized at the expense of logical ones, thus preventing us from fully understanding the issues that led Frege to draw the distinction in the first place and, so, from adequately assessing its significance.

Once we accord Frege's concern with logic its proper place, we see that the central issues raised by the phenomena that led him to distinguish sense from reference are best understood as primarily logical and only secondarily as epistemological. For Frege, we are obliged to distinguish sense from reference to do justice to differences between sentences that cannot sensibly be accommodated by a theory of reference alone – differences, nevertheless, that *logic* obliges us to recognize. The considerations that govern the logical appraisal of our assertions or judgements require an appeal to something over and above their referential truth conditions. The challenge is to provide an account of the contents of our assertions and judgements – and, in turn, of the semantic properties of our sentences – that does adequate justice to this fact.

In what follows, after briefly elaborating and defending this perspective on the origin and role of Frege's distinction, I shall consider the critical reception the distinction has received in twentieth-century philosophy of language. Necessarily, my approach will be selective, for the critical literature generated by Frege's distinction is too massive and theoretically diverse to permit an adequate, unified and comprehensive treatment.[1] My focus will be on a somewhat artificially circumscribed debate the central concern of which is whether an adequate theory of meaning obliges or even permits appeal to Frege's notion of sense.

Two broad traditions have dominated this debate, a critical one concerned to demonstrate the semantic irrelevance (if not incoherence) of the notion, and a defensive neo-Fregean backlash. The critical tradition has its origins in the work of Bertrand Russell and finds its latest expression in certain of the so-called direct reference theorists. For better or worse, the literature here has concentrated almost exclusively on the question of sense for singular terms – in particular, proper names (though, more recently, also for indexicals).[2] Perforce, my own discussion will follow this lead.[3]

[1] The more significant and regrettable lacunae include discussion of (i) the impact of Frege's distinction on the work of Carnap and Church, and how their interpretations of the distinction have played out in the development of intensional logic and formal semantics; (ii) the relevance of Quine's various animadversions against meaning and related notions to the theoretical usefulness and/or coherence of Frege's notion of sense; (iii) the distinctive issues raised for a theory of sense by indexical expressions; and, not unrelatedly, (iv) the motivations for, and viability of, that variety of neo-Fregeanism that insists on the object-dependent nature of the sense of most proper names and all indexical expressions.

[2] For an overview of the central issues raised by indexicals for Frege's theory of sense see John Perry, 'Frege on demonstratives', *Philosophical Review*, 86 (1977), pp. 474–97; Gareth Evans, 'Understanding demonstratives', in *Collected Papers* (Oxford: Oxford University Press, 1985), pp. 291–321 (originally in H. Parret and J, Bouveresse (eds.), *Meaning and Understanding* (Berlin: Walter de Gruyter, 1981)); David Kaplan, 'Demonstratives', in J. Almog, J. Perry and H. Wettstein (eds.), *Themes from Kaplan* (New York: Oxford University Press, 1989), pp. 481–614.

[3] Most recently, especially in the United States, the relevant issues tend to get raised in the context of a broader concern with the semantics of propositional attitude ascriptions and theories of mental content generally. Given the role that senses play in Frege's own discussion of attitude reports, the reasons for this are, perhaps, obvious. We cannot, however, do justice to the full range of concerns raised by this

According to the neo-Fregean, once we correctly understand the considerations that led Frege to introduce his distinction, we shall see that the conception of sense attacked by the Russellians – an essentially description-theoretic conception – is not one to which Frege was committed. Moreover, these considerations, when properly understood, succeed in showing that an adequate theory of meaning cannot get by solely on the resources provided by a theory of reference and, so, will inevitably find itself obliged to appeal to some notion of sense. Michael Dummett has been the most influential expositor and defender of Frege's views on sense in the twentieth century, and though there are prominent neo-Fregeans who differ with him on important points, my discussion will focus on the neo-Fregean position as Dummett develops it.

I THE LOGICAL BASIS OF SENSE

Frege's approach to logic was shaped in fundamental ways by his understanding of judgement, assertion and inference, and by the way he took these practices to be constitutively subject to logical appraisal. Assertion and inference are, by their nature, subject to assessment as correct or incorrect. The notion of correctness constitutive of assertion is inextricably bound up with the notion of *truth*: an assertion is correct, in the first instance, just in case what is asserted is true. But, for Frege, to be subject to assessment as true or false is also and immediately to be subject to logical assessment. To make an assertion is to aim at saying something true and to commit oneself to the truth of what one has said in such a way as to be subject to a determinate range of logical evaluations. To understand an assertion is to appreciate what, in making the assertion, the speaker has thereby committed herself to.

For Frege, then, the principles of logic determine standards that constitutively govern those of our cognitive and linguistic practices that concern themselves essentially with truth, standards we cannot opt out of and still take ourselves to be making assertions or drawing inferences.[4] To articulate these principles, the logician

literature. Nevertheless, there is substantial overlap, and much of what we discuss below will be relevant to those concerns.

[4] As Frege puts the matter in the Preface to *Gg*, the laws of logic 'are the most general laws, which prescribe universally how one should think if one is to think at all'.

must provide a systematic account of all, but only those features of our assertions in virtue of which they are subject to appraisal in whatever logically relevant ways they are. This will require first that she distinguish between assertions proper and *what* is asserted – their contents. Two assertions will differ in their logically relevant contents just in case they differ in their potential for logical appraisal. An assertion's potential for logical appraisal is fixed by the *way* in which its content articulates the conditions under which that assertion is to be counted true. For Frege, then, it is what one asserts – the content of one's assertion – that is in the first instance true or false and *simultaneously* the locus of inferential potential.

The next step for the logician is to provide an account of the logical complexity of these contents – a systematic account of the logically relevant ways in which assertoric contents may differ. Frege's conception of logical complexity was fundamentally compositional. Substituting, within a sentence, one expression for another with the same logically relevant content must leave unchanged the logically relevant content of the sentence as a whole. In the *Begriffsschrift*, before drawing his distinction between sense and reference, Frege identified the logically relevant contents of sentences with what he there called 'judgeable contents' (*beurteilbare Inhalte*). He identified these in turn with just those objective circumstances the obtaining or not of which determines the truth or falsity of the relevant assertion. He called the logically relevant content of (monadic) predicate expressions 'concepts' and proceeded, in effect, to identify concepts with functions from appropriate arguments to judgeable contents.[5]

Frege was aware that this way of viewing matters faces a *prima facie* difficulty when it comes to expressing the identity of contents. If two expressions *a* and *b* possess the same content, and if we take the content of '=' to be an ordinary relation, then, by compositionality, the sentence ⌜$a = b$⌝ must have the same judgeable content as the sentence ⌜$a = a$⌝. Frege took it as obvious, however, that ⌜$a = b$⌝ and ⌜$a = a$⌝ express different judgeable contents. Given his insistence that judgeable content concerns only what is logically relevant, to suppose that such sentences differ in judgeable content is just to suppose that they differ in some logically relevant way – which, of course, they do. An assertion of ⌜$a = a$⌝ cannot, for example, always

[5] See *Bs*, especially the Preface and §§1–11.

be substituted *salva validitate* in any inference in which an assertion of ⌜$a = b$⌝ occurs as premise. To maintain his identification of judgeable contents with objective circumstances, Frege found himself obliged to distinguish the circumstances the obtaining of which would make each of these sentences true. Frege suggested that we understand the content of '=' in such a way that sentences like ⌜$a = b$⌝ do not, in fact, express a relation between the contents of *a* and *b* but, rather, between the expressions themselves. Since the metalinguistic circumstance that *a* has the same content as *b* differs from the metalinguistic circumstance that *a* has the same content as *a*, Frege has, it seems, successfully avoided the problem.

By 1891, Frege became dissatisfied with his *Begriffsschrift* views and with the treatment of identity sentences to which they seemed to commit him. It is noteworthy, however, that in 'Function and concept' – in which he concludes for the first time in print that we must distinguish the sense of a sentence from its reference – Frege takes as his starting point not the difficulty with identity-statements that led to his *Begriffsschrift* proposal but, rather, a generalization of it.[6] Consider, for example, the concept that would, for Frege, constitute the content of the predicate expression '( ) is the second planet from the Sun'. Now consider the names 'Hesperus' and 'Phosphorus', which both have the planet Venus as their content. Since no function, and so no concept, can have different values for the same argument, the value of our concept for the argument Hesperus must be the same as its value for the argument Phosphorus. In the context of the *Begriffsschrift*, this means that the circumstance that is required to obtain in order for the sentence

(1) Hesperus is the second planet from the Sun

to be true is the same as the circumstance that is required to obtain in order for

(2) Phosphorus is the second planet from the Sun

to be true. Insofar as Frege's notion of circumstance is clear at all, this seems just as it should be: (1) and (2) should be assigned the same referential truth conditions. For Frege, though, it is also clear that logic obliges us to distinguish the judgeable content of an assertion of (1) from that of (2): an assertion of (1) cannot be substituted

[6] 'Function and concept', *CP*, pp. 137–56, see esp. p. 145.

salva validitate in any inference in which an assertion of (2) occurs, nor *vice versa*. Consequently, judgeable contents cannot be circumstances: judgeable contents, individuated by appeal to logical considerations, cannot be the values of concepts for appropriate arguments.[7]

Frege came to appreciate the inadequacy of his *Begriffsschrift* solution in part, no doubt, by recognizing how easy it is to generalize the difficulty with identity sentences to other – perhaps, even, to all other – sentence forms. His *Begriffsschrift* proposal for resolving this difficulty is plausible only if the particular metalinguistic circumstances hypothesized to be the real contents of identity sentences genuinely reflect what logic obliges us to take such sentences to be about. Clearly, though, it would be quite implausible to generalize the *Begriffsschrift*'s metalinguistic strategy across the board. Doing so would deprive nearly every sentence of its ordinary significance, of its capacity to express, as Frege might say, 'proper knowledge'.[8] By reflecting on the general case, Frege came to see that, if we are to do justice to the truth conditions of identity sentences and to the role they play in our inferences, we have no serious choice but to suppose that the constituent names contribute just those objects that we ordinarily take them to designate to the truth conditions of the sentences in which they occur. What Frege realized, in effect, is that two sentences, despite having the same referential truth conditions, may still differ in ways that *logic* obliges us to acknowledge.

As Frege now saw matters, an adequate account of the differences and similarities between assertoric contents that logic obliges us to acknowledge will require an appeal not only to features that it is the business of a theory of reference to articulate, but also to features that require the resources of a theory of sense.[9] While it is up

[7] Eventually, of course, for reasons he outlines in detail in 'On sense and reference', *CP*, pp. 157–77, Frege identified the values of concepts for appropriate arguments with truth-values. For useful discussion of these issues, see Tyler Burge, 'Frege on truth', in L. Haaparanta and J. Hintikka (eds.), *Frege Synthesized* (Dordrecht: Reidel, 1986), pp. 97–154.

[8] See Frege's discussion in the first paragraph of 'On sense and reference', *CP*, pp. 151–2, where, in criticizing the *Begriffsschrift* solution to the identity problem, he writes, 'In that case the sentence a = b would no longer refer to the subject matter, but only to its mode of designation; we would express no proper knowledge by its means. But in many cases this is just what we want to do.'

[9] As he remarked in a letter to Husserl (*PMC*, pp. 61–4), his *Begriffsschrift* notion of a judgeable content involved a confused conflation of what he now held must be distinguished in terms of his notions of sense and reference.

to the theory of reference to articulate the conditions that must be satisfied in order for a sentence to be true, it belongs to the theory of sense to articulate what a speaker grasps when, by understanding the sentence, she grasps these conditions.[10] For it is in virtue of this understanding that a speaker incurs the inferential commitments she does by asserting the sentence. When asserting a sentence, a speaker's commitment to its truth, to the satisfaction of its referential truth conditions, cannot be divorced from her inferential commitments.

The need to distinguish between sense and reference manifests itself first, for Frege, in relation to sentences. The need to extend the distinction to the level of sub-sentential expressions is a consequence of his continued commitment to compositionality. Accordingly, two coreferential sub-sentential expressions will differ in sense if the sentence that results from substituting one for the other has different logical properties from the original, despite having the same truth-value. They differ in sense if they present their shared reference in different ways such that it is by grasping this difference that a speaker appreciates the different inferential properties of sentences containing them. In general, then, Frege views the sense of a sub-sentential expression as *a way of thinking about its* reference – or, if one prefers, a way in which the expression presents the reference to the speaker. Nothing is relevant to individuating these ways of thinking (or 'modes of presentation') except what is distinctive about the contribution the expressions make to the logical properties of the sentences in which they occur. To grasp the sense of an expression, then, is to be able to think of its reference in the relevant sort of way – a way that enables one to appreciate the logical properties of the sentence containing it.

If the critical tradition has not always adequately appreciated that Frege was led to introduce his notion of sense on the basis of strictly logical considerations, at least some of the blame for this must lie with Frege himself. For nowhere does he explicitly and systematically explain why he distinguishes the thoughts expressed by sentence pairs of the sort that puzzled him. Instead, he assumes that his readers will have little trouble appreciating why he distinguishes thoughts (or, earlier, judgeable contents) where he does. And even

[10] See *Gg*, §32, pp. 221–2.

in those passages where he is more expansive, he tends to appeal to criteria that advert to very general features of our epistemic and linguistic practice, while relying on his readers to appreciate the underlying *logical* relevance of those features.

Moreover, in the scattered passages in which he takes up these matters, Frege tends to rely on two prima facie different criteria, thereby further complicating the interpreter's task. On the one hand, there are a variety of passages that make it evident that Frege holds that

> (A) Two sentences express different thoughts whenever it is possible for a speaker competent in the use of both sentences to extend his knowledge by coming correctly to believe what would be expressed by the (literal) utterance of one, despite his already knowing the truth of what would be expressed by the (literal) utterance of the other.[11]

But there are at least an equal number of passages that support his holding the following criterion:

> (B) Two sentences express different thoughts whenever it is possible for a speaker competent in the use of both sentences to believe what would be expressed by the (literal) utterance of one while, without changing her mind, failing to believe – either disbelieving or suspending judgement on – what would be expressed by the (literal) utterance of the other.[12]

While the possibilities adverted to in (A) and (B) are indeed different, a careful reading and comparison of the passages in which he deploys these criteria (many of which overlap) make it clear that Frege intended them to be understood as issuing from the same, more fundamental, underlying considerations. The ultimate source of both possibilities lies in the fact that the sentences at issue differ in their logical potential.

[11] See, for example, the passage from 'Function and concept' mentioned earlier; the famous opening passage from 'On sense and reference', and the 'Dr Lauben' discussion in 'Thoughts', *CP*, pp. 358–60.

[12] See, for example, the important and illuminating 'Aphla' and 'Ateb' passage from an undated letter from Frege to Philip Jourdain (*PMC*, pp. 78–80) or, again, the opening passages from 'On sense and reference'.

This is easiest to see with respect to (B), at least once the relevant notion of *possibility* appealed to there is clarified. Given Frege's well-known anti-psychologism, this cannot be a matter of mere psychological possibility. The notion at issue here is best understood *normatively*. Frege would not, however, have concerned himself with merely pragmatic permissions and restrictions. The notion is best understood in a sense appropriate to purely logical concerns. Properly understood, then, (B) says that two sentences will differ in the thoughts they express whenever a speaker *may, without violating any logical norms*, believe what is expressed by the utterance of one while failing to believe – either suspending judgement on or believing the negation of – what is expressed by the utterance of the other. This amounts to no more than saying that two sentences differ in the thoughts they express if they differ in their logical potential.

Many interpreters, however, focus on (A), taking it or some variant of it to capture best the issues of central concern to Frege. Viewing matters this way, the possibility adverted to in (B), if considered at all, is taken to be secondary to or consequent upon that adverted to in (A). In any case, the explicitly epistemic orientation of (A) is doubtless a major reason why for so many interpreters it is obvious that Frege introduced the notion of sense primarily in the service of epistemic considerations.

I do not mean to deny that differences in sense are epistemically significant. There is certainly no denying that Frege appeals to epistemic notions in the various passages from which (A) is drawn. Nevertheless, we should not allow this to mislead us into overlooking the more fundamental role played by logical concerns of the sort already discussed. If we say that sentences differ in their *cognitive significance* whenever they admit of the epistemic possibilities adverted to in (A), then the only differences in cognitive significance of interest to Frege are those that arise *because* the sentences differ in some straightforward logical respect.

Consider, for example, the famous opening passage of 'On sense and reference' (*CP*, p. 157) where Frege asserts that true sentences of the form $a=b$ differ in 'cognitive value' (*Erkenntniswert*) from counterpart statements of the form $a=a$, and that the former 'often contain valuable extensions of our knowledge'. To see how we can account for this difference in 'cognitive value' or such 'valuable extensions of one's knowledge' by appealing only to logical differences between such sentences, consider a case in which we would ordinarily say that a subject, while already knowing that $a = a$, comes to learn that $a = b$. Suppose that prior to acquiring this new knowledge, our subject also already knows that *Fa*, that *Gb*, and that $(\forall x)(Fx \,\&\, Gx) \supset Mx)$. Also assume that she disbelieves that $\neg Fb$ (i.e., believes that $\neg Fb$), while suspending judgement on whether *Ga*, or whether *Ma*, or, for that matter, on whether *Mb*. Prior to learning that $a = b$, our subject's system of beliefs is cognitively deficient in a number of obvious respects. Moreover, it is evident that no amount of purely logical reflection will enable her to appreciate this deficiency: though some of her beliefs are false, she cannot be accused of having logically inconsistent beliefs, nor can she be accused of having overlooked any logical implications.[13] It is easy to see, however, that by coming to learn that $a = b$, our subject will immediately be in a position to overcome these deficiencies. Not only is she poised to acquire new knowledge – the knowledge that *Fb*, that *Ga*, and that *Ma* – but, with respect to *Fb*, she is in a position to free herself from a false belief. Indeed, were she not to change her mind about *Fb*, she would be subject to serious logical censure – she would justly be accused of having inconsistent beliefs.[14] And, depending on the circumstances, were she to fail to come to believe that *Fb*, that *Ga*, and hence that *Ma*, she might also be subject to logical censure – though the censure here would be less serious (failing to draw an obvious inference).

My suggestion, then, is that when Frege speaks of different sentences having differing 'cognitive value' (despite having the same

[13] Given that $a = b$, I am willing to grant that our subject's belief that *Fa* and her belief that $\neg Fb$ are metaphysically incompatible – they cannot both be true. But acknowledging this in no way suffices to ground an accusation of inconsistency in the sense relevant to the logical appraisal of beliefs, assertions, etc. While metaphysically incompatible, her beliefs are not logically inconsistent.

[14] Strictly speaking, of course, logic obliges our subject only to correct the inconsistency in her beliefs. So she must either give up her false belief that $\neg Fb$ or her correct belief that *Fa*. If we assume, for the sake of the example, that our subject is more strongly committed to the latter belief than to the former, the cognitive value of coming to believe that $a = b$ cannot be denied. But even if we refrain from assuming any specific preferential commitments in her prior beliefs, coming to learn that $a = b$ (despite already believing that $a = a$) can be counted a valuable cognitive achievement just in virtue of the opportunity it affords one for correcting cognitive deficiencies of the sort exemplified in the example – whether or not any particular subject takes advantage of this opportunity.

referential truth conditions), or of how coming to learn the truth of what is expressed by one such sentence may issue in a 'valuable extension of our knowledge', despite one's already knowing the truth expressed by the other, we do best to understand him as having in mind the possibility of significant cognitive improvements of just these sorts. And to make sense of such cognitive improvements, we need only recognize that different sentences, despite having the same referential truth conditions, may have different logical properties.

2 THE TRADITIONAL RUSSELLIAN

Few today will deny that the most enduring and influential tradition critical of Frege's doctrine of sense was inspired by the work of Bertrand Russell. Indeed, for better or worse, Russell's understanding of, and reaction to, Frege's distinction has structured in a number of crucial respects the discussion that was to follow, both critical and defensive. While Russell himself denied that any coherent theoretical role could be found for the notion of sense, this did not, as we shall see, prevent those more sympathetic to the notion from trying to appropriate what they took to be Russell's insights while retaining a full-fledged sense–reference distinction.

The guiding idea of the Russellian tradition is the view that the semantic content expressed by the assertoric use of an indicative sentence is to be thought of as a piece of (truth evaluable) *information* – in a sense to be examined below.[15] Thus, for the Russellian, the possibilities that led Frege to distinguish the thoughts expressed by puzzling sentence pairs will be of genuine semantic relevance only if they expose the need to distinguish between the information content semantically encoded by each sentence. The *traditional* Russellian, as I shall call him – exemplified most clearly by Russell himself – allows that these possibilities *do* in fact show that the relevant sentences encode different information and, so, agrees with Frege that the sentences differ in a semantically relevant way. He

[15] For a clear expression of this point of view, see the opening pages of Nathan Salmon, 'Reference and information content: Names and descriptions', in D. Gabbay and R. Guenthner (eds.), *Handbook of Philosophical Logic* (Dordrecht: Reidel, 1990), pp. 409–61.

denies, however, that accommodating this fact requires anything like a theory of sense.

The *neo*-Russellian, on the other hand – as exemplified by certain of the so-called direct-reference theorists (the most prominent of which are, perhaps, Nathan Salmon and Scott Soames) – *denies* that differences in cognitive significance of the sort that concerned Frege suffice to show that distinct information is semantically encoded by the relevant sentences. Since the relevant sentences are not taken to encode different information, there is no *semantic* difference for a theory of sense to explain.

The Russellian's focus on and understanding of the notion of information content is guided by something like the following line of reasoning. His interest, from the outset, is on the sorts of possibility adverted to in (A) above. What would account for the possibility that someone might extend his or her knowledge simply by coming correctly to believe what is expressed by one sentence, despite already knowing the truth expressed by some distinct sentence, except for the fact that the two sentences semantically encode different information about the world? And how can two sentences encode different information about the world unless they require, in order to be true, that different conditions be satisfied by the world? Accordingly, sentences will differ in their information content just in case they differ in their referential truth conditions – either because they are about different features of the world or because they represent the same features as involved with each other in different ways.[16]

It is now obvious why, for the traditional Russellian, it makes no sense to appeal to a level of semantic description distinct from the referential when attempting to explain the semantic differences between sentences of the sort that puzzled Frege. For the traditional Russellian, the challenge that Frege took himself to be faced with – the challenge of explaining how two sentences with the same referential truth conditions can nevertheless differ in some semantically

[16] This way of understanding information content is closely tied to both the traditional and neo-Russellian's view that the semantic contents of sentences are *structured propositions* the constituents of which include the very objects, properties, propositional functions or other 'features' of the word whose 'involvement' with each other is relevant to the truth of the relevant sentences.

relevant way – will seem deeply confused. He will simply not admit that two sentences can differ in the information they semantically encode and yet fail to differ in their referential truth conditions. So there is nothing for a theory of sense to explain.

Traditional Russellians must suppose that Frege failed to identify the correct referential truth conditions of the sentence pairs he found puzzling. Since the relevant sentences are assumed to differ in the information they convey, what is required is an alternative referential analysis – one that captures, without residue, all of the semantically relevant differences. In effect, the traditional Russellian is committed to a strategy similar to the one Frege himself deployed in the *Begriffsschrift*.[17]

Russell's theory of definite descriptions is, in large measure, an alternative implementation of just such a strategy. A central aim of that theory is to show how one can capture the semantic differences between pairs of sentences of the sort that concerned Frege directly in terms of a difference at the level of reference – or its Russellian counterpart, *meaning* – and thus, without any need to appeal to a notion like sense.[18] Russell, for example, agreed with Frege that someone could extend her knowledge by coming to believe what was expressed by a sentence like

(3) The author of *Waverly* was Scottish

despite already correctly believing what is expressed by

(4) The author of *Ivanhoe* was Scottish.

[17] The traditional Russellian is not, of course, obliged to adopt the particular, essentially metalinguistic analysis offered there.

[18] At least as important for Russell, of course, was his view that it was obviously possible for sentences containing non-denoting descriptions to be meaningful. Given the framework within which he was operating, this would not be possible if we assume that the semantic role of a definite description is simply to refer to some particular object – that is, if we assume that definite descriptions are *logically* proper names. So either it is a mistake to suppose that such sentences are meaningful after all – an option Russell was unwilling to consider – or we must analyse sentences containing definite descriptions in a way that shows how they can indeed be meaningful despite the fact that there is no one object that a contained description picks out. The theory of descriptions was in part designed to do precisely this. See Bertrand Russell, *Logic and Knowledge: Essays 1901–1950*, ed. R. C. Marsh (London: Allen and Unwin, 1956), and *Introduction to Mathematical Philosophy* (London: Allen and Unwin, 1919), pp. 167–80.

But as is shown by the following (simplified) representations of their Russellian truth conditions, Russell took these sentences to be *about* different features of the world:

(3) $(\exists x)(Wx \,\&\, (\forall y)(Wy \supset x = y)) \,\&\, Sx)$

(4) $(\exists x)(Ix \,\&\, (\forall y)(Iy \supset x = y)) \,\&\, Sx)$.

According to Russell, these two sentences say of *distinct* properties – the property of being the author of *Waverly* and the property of being the author of *Ivanhoe*, respectively – that each is uniquely instantiated by an object that also instantiates the property of being Scottish.[19]

Russell's general strategy was motivated by his view that we can make no sense of semantically relevant differences in cognitive significance except in terms of differences in information content (and, so, differences in referential truth conditions). The detailed implementation of the strategy, however, was determined in large measure by Russell's specific epistemic commitments. For a speaker to talk about – indeed, even to think about – a specific object, she must, according to Russell, *know* what object she has in mind. This means she must be able epistemically to isolate that object from all others; her knowledge must uniquely determine the intended object.[20] Any difference in the cognitive significance of two sentences that, according to Frege, involve predicating the same thing of the same object is best accounted for, according to Russell, by a difference in the knowledge by means of which the speaker is

[19] Notice that neither sentence concerns itself directly or essentially with Sir Walter Scott. No individual is mentioned. Scott does not, as Russell himself would put it, occur as a constituent in either of the propositions expressed by these sentences.

[20] See Bertrand Russell, 'Knowledge by acquaintance and knowledge by description', in *Mysticism and Logic* (Totowa, N.J.: Barnes and Noble, 1917), pp. 152–67, esp. p. 159, and *The Problems of Philosophy* (Oxford: Oxford University Press, 1959), ch. 5, pp. 46–59. The claim that in order to talk or think of a particular object – that is, in order for a particular object to be the immediate object of a propositional attitude – one must *know which* object one is talking or thinking about may seem innocent enough. Indeed, on some readings it is surely truistic. Troubles arise, as we shall see, only with particular accounts of what is involved in, or required to count as, 'knowing which object'. One will be tempted by what I believe are the problematic accounts as soon as one views the requisite knowledge as something one can possess independently of, and prior to, having any particular propositional thoughts concerning that object.

enabled, in each case, to think about that object. But how are we to think of this knowledge in order to understand both how it enables speakers epistemically to isolate the relevant object, and how differences in such knowledge can differentially affect the information content of the sentences in which they appear? How indeed, except in terms of cognitively transparent, uniquely individuating descriptive conditions that are directly incorporated into the referential truth conditions of the relevant sentences.

In the case of sentences containing definite descriptions, Russell's theory brilliantly accommodates these various demands. But the adequacy of his strategy ultimately depends on whether, and how satisfactorily, it can be generalized to accommodate puzzling sentence pairs other than those containing co-denoting definite descriptions. Two sentences that differ at most in containing distinct, co-designating *proper names* can also differ in cognitive significance. For the traditional Russellian's strategy to work for such sentences, he will also have to deny that they have the same referential truth conditions, and he will be obliged to offer plausible alternative referential analyses.

Given the theory of descriptions and its motivations, it comes as no surprise that Russell proposed that we view ordinary proper names as, in effect, convenient abbreviations for underlying definite descriptions. Names that differentially affect the cognitive significance of sentences in which they appear must abbreviate different descriptions. And so the substitution of one such name for another in a sentence will result in a sentence with different referential truth conditions from the original.

Of course, in order to avoid reintroducing, at a different level, the very problem that Russell's alternative referential analysis was intended to solve, the descriptors of these underlying descriptions must themselves contain no expressions which – though intuitively picking out the same feature of the world as some other possible expression – might, nevertheless, differentially affect the cognitive significance (and, so, information content) of the sentences in which they appear. In effect, this is to require that these descriptors contain only *logically proper* expressions, the references of which are always only features of the world with which we are immediately *acquainted*. Indeed, it turns out, on Russell's view, that the information semantically encoded by any sentence we understand can directly concern *only* items with which we are, in his sense, acquainted.[21] And one will count as being acquainted with some feature of the world *X*, in the sense required, only if *X* exists and, for any feature of the world *Y*, if one is also acquainted with *Y*, and *X* is the same feature as (a different feature from) *Y*, then it is impossible for one to fail to recognize that they are the same (different). Reflection, however, suggests that we can be acquainted in this sense only with such allegedly epistemically intimate items as sense data, universals, and the like. Since, on Russell's view, our talk and thought never directly concerns objects with which we cannot be acquainted, it will never directly concern tables and chairs, our friends, lovers, or family members, not even our own hands and feet!

The difficulties facing any such proposal are familiar. The ontology of sense data and the variety of foundationalist epistemology that goes with it are, at best, highly tendentious and most likely untenable. Moreover, the notion of acquaintance that Russell needs if his generalized proposal is to have any chance at plausibility is notoriously difficult, if not impossible, to spell out and defend coherently. Finally, like Frege's own *Begriffsschrift* proposal, the alternative referential analyses to which Russell would be committed, even assuming they can be specified coherently, manifestly fail to do justice to what we ordinarily take such sentences to be essentially about – at least in the case of ordinary proper names and indexical expressions.

The extent of this last failure was forcefully revealed by a set of detailed and devastating considerations advanced in the early 1970s by a number of different philosophers. Especially influential here was Kripke's *Naming and Necessity*, in which he demonstrates just how poorly any essentially description-theoretic approach to the semantics of proper names does justice to various fundamental semantic, metaphysical and epistemological intuitions associated with our use of proper names.[22] It is important to realize, however, that Russell's view that ordinary proper names are disguised or

[21] In 'Knowledge by acquaintance', p. 159, Russell puts the matter this way: '*Every proposition which we can understand must be composed wholly of constituents with which we are acquainted.*' (Russell's emphasis).

[22] Saul Kripke, *Naming and Necessity* (Cambridge, Mass.: Harvard University Press, 1980). For a useful overview of the separate epistemic, metaphysical and semantic

abbreviated definite descriptions was not the only, nor perhaps even the primary, object of criticism in *Naming and Necessity*. Kripke's arguments were also directed at Frege's view that ordinary proper names possess a sense. The view that arguments like Kripke's succeeded in undermining not only the traditional Russellian's position but also Frege's played an important role in motivating the approach that I earlier called neo-Russellian.

3 CLEARING THE WAY FOR THE NEO-RUSSELLIAN

That arguments marshalled against Russell's proposal should have been taken in turn to prove fatal to Frege's doctrine of sense is not that surprising. For the conception of sense under attack was the familiar – and at the time widely accepted – description-theoretic conception. Like Russell, description theorists with respect to sense saw little chance of accounting for the differential contribution that distinct co-referring names can make to the cognitive significance of sentences in which they occur except in terms of differences in the cognitively transparent descriptive conditions that speakers competent in the use of the names associate with them. They shared with Russell not only the view that for a speaker to talk about a particular object she must be able epistemically to isolate it from all other objects, but also the view that to epistemically isolate an object (at least one with which she is not acquainted), a speaker must be able to describe it uniquely. Accordingly, an object will count as the reference of a given name in virtue of its satisfying the associated uniquely individuating descriptive conditions.[23]

Despite its overwhelming popularity at the time, proponents of the description theory of sense rarely bothered to spell it out in much more detail than I have done here. A central difficulty concerns precisely how we are to understand the 'association' between name

arguments to be found in *Naming and Necessity*, see Nathan Salmon, *Reference and Essence* (Princeton, N.J.: Princeton University Press, 1981), pp. 23–32, and more recently Scott Soames, *Beyond Rigidity: The Unfinished Semantic Agenda of Naming and Necessity* (New York: Oxford University Press, 2002), esp. pp. 18–54.

[23] If no object satisfies the conditions, then according to this view the name has no reference – though, consistently with Frege's repeated claims, the name may still be said to have a sense. Sentences containing such a name may still count as meaningful, as expressing thoughts. See note 13 above.

and descriptive conditions to which it appeals. The most straightforward interpretation – the suggestion that the sense of a name is always *equivalent* to the sense of some definite description – is also the most problematic. First of all, it is not clear how to understand this suggestion without its collapsing into the traditional Russellian proposal. In that case, if anything like Russell's theory of descriptions provides a correct account of the semantic role of definite descriptions, any differences in cognitive significance between co-designating names will be accounted for by differences in the referential truth conditions of the sentences in which they appear. The presumed need for a theory of sense would evaporate.

Secondly, if we are obliged to take some specific uniquely identifying description, say ⌜the Φ⌝, to be equivalent in sense to some name n, then we shall, it seems, be obliged to view the sentence ⌜n is the Φ⌝ as equivalent in sense to ⌜The Φ is the Φ⌝ and so, presumably, as necessarily true and knowable a priori. But virtually any uniquely identifying description that an ordinary speaker is likely to associate with a name at her disposal is such that she would ordinarily regard the relevant instances of ⌜n is the Φ⌝ as only contingently true and knowable only a posteriori. It would seem, then, that none of the uniquely identifying descriptions that an ordinary speaker might associate with the names at her disposal will be legitimate candidates for giving the sense of those names.

To avoid the first problem, the description theorist must explain how to view the semantic role played by the descriptive conditions associated with a name in such a way that any difference in descriptive conditions that are satisfied by the same object are not reflected in a difference in the referential truth conditions of sentences containing the relevant names. At the same time, though, his account must allow that the difference in descriptive conditions *is* directly reflected in differences in the thoughts expressed by sentences containing the names. He must do this, however, in a way that avoids the second problem.[24]

There are various ways in which advocates of a description-theoretic approach to sense have attempted to avoid these problems.

[24] It is not implausible to think that David Kaplan's distinction between character and content was in part introduced to accomplish precisely this. See Kaplan, 'Demonstratives'.

Perhaps the most common – at least with respect to the second issue – was to adopt some variety of what has come to be called the *cluster* version of the description theory. According to this view, roughly, an object counts as the reference of a name not in virtue of its uniquely satisfying some one specific description but, rather, by its uniquely satisfying some sufficient number of descriptions that the speaker has associated with the name.[25] But even if we suppose that this idea can be developed in a way that avoids the two problems just mentioned, difficulties remain that raise serious doubts about the ultimate tenability of *any* description-theoretic conception of sense for proper names. By reflecting on our ordinary practice in using names, Kripke, in effect, shows that to be able to use a proper name to refer successfully to a particular object, it is neither necessary nor sufficient that the speaker associate *any* (non-circular) uniquely identifying description with the name.[26]

What Kripke's discussion helps us to see is that the assumption that in order to refer to a particular object one must be able epistemically to isolate it from other objects – on at least the description-theoretic understanding of what this involves – cannot be supported by a close examination of our actual name using practice. On the one hand, names are often successfully used by speakers who have very little descriptive knowledge of their references and, in many cases, certainly not enough to form a correct uniquely identifying description. On the other hand, even when a speaker does associate what she takes to be a uniquely identifying description of the reference of a name at her disposal, the object satisfying the description (or a sufficient number of descriptions in the associated cluster) may not be the object to which she succeeds in referring when using the name. Indeed, a speaker may successfully refer to an object by her use of a name even though *no* object uniquely satisfies the associated description(s).

[25] The cluster theory is often (I think mistakenly) credited to Wittgenstein on the basis of the 'Moses' discussion in remark 79 of Ludwig Wittgenstein, *Philosophical Investigations* (New York: Macmillan, 1953). The view was first deployed in connection with the problems raised in the text by John Searle, 'Proper Names', *Mind*, 67 (1958), pp. 166–73.

[26] The proscription against circular descriptions is explained and defended in Kripke, *Naming and necessity*, pp. 68–70.

A speaker's ability to use a name (one and the same name) to refer to a particular object is typically something that will remain stable across considerable variations in both the number and accuracy of her beliefs about that object – in any case, considerably more variation than description theorists can permit. Moreover, different individuals can be competent in the use of the *same* name – in the sense that they could use that name to communicate successfully with each other about its reference – despite significant differences in their knowledge and beliefs about the reference – again, more differences than description theorists can allow. But if this is right, then it is not plausible to suppose that the intra- and inter-personally stable compositional contribution that a proper name makes to the propositional content (whether information content or thought) semantically expressed by sentences containing it must consist in or be fixed by the descriptive information that speakers associate with the name.

Critics of description-theoretic approaches came to realize that it was implausible to assume that the circumstances that enable a speaker to count as a competent user of a name – specifically, the circumstances responsible for making one object rather than another the name's reference – are exhausted by facts about her individual, cognitively transparent, descriptive resources. Rather, whether a speaker counts as competent in the use of a name will often essentially be a matter of her being situated in a context (social and/or physical) that relates her in the right sort of way to the name's reference. How best to understand this contextual involvement is a matter of considerable dispute.[27] But what is important – indeed, revolutionary – about this way of thinking is the recognition that factors external to and, in an important sense, independent of a speaker's individual cognitive resources may play an essential

[27] If it is not in virtue of their possessing the same, or sufficiently similar, uniquely individuating descriptive knowledge of its (purported) reference that speakers count as competent in the use of a name, then in virtue of what do they so count? It was in connection with this question that the so-called 'causal *theory* of reference' was introduced. In offering the considerations he does, Kripke cautiously says that he does not intend to be providing a causal *theory* of reference but, rather, a 'better picture than the picture presented by the received views. See Kripke, *Naming and necessity*, pp, 93–6. Michael Devitt, however, is not so modest and proposes in *Designation* (New York: Columbia University Press, 1981) to articulate and defend just such a theory. This is not the place to attempt to survey, yet alone to assess, the various ways in which this 'theory' gets interpreted and deployed.

enabling role vis-à-vis her ability to refer to a particular object using a name. Once we acknowledge that a speaker's contextual situation might play such an enabling role, we are free to unburden the competent name user of the particular epistemic responsibilities with which description theorists have saddled her.

4 THE NEO-RUSSELLIAN ALTERNATIVE[28]

On the basis of considerations such as these, the neo-Russellian concedes that Russell was wrong to hold that ordinary proper names are best viewed as disguised or abbreviated definite descriptions. Nevertheless, he takes these considerations to provide powerful support for the more fundamental Russellian idea that, in providing an account of the content semantically conveyed by a sentence, there is no need or room for a theory of sense over and above a theory of reference.

According to the neo-Russellian, arguments like Kripke's help us to see that it is only the reference of a name that remains stable across the intra- and inter-personal diversity of descriptive beliefs held by competent users of that name; and so it is only their references that names can plausibly be taken to contribute compositionally to the information content semantically expressed by the sentences in which they occur. For the neo-Russellian, then, the substitution of distinct but co-referring names within a sentence will, with a few exceptions, result in a sentence that semantically expresses the same propositional content – which is to say, the same information content – as the original sentence. But what, then, of the considerations that led both Frege and the traditional Russellian to suppose that these sentences express different contents?

The neo-Russellian admits that a speaker can acquire new information when she comes to believe what is conveyed by the utterance of a sentence with the same referential truth conditions as a distinct sentence whose content she already believes. He will insist, however, that this does not show that the relevant sentences *semantically encode* different information. Rather, he points out, the

[28] Much of the material in this section derives from a more expanded discussion in William Taschek, 'Frege's puzzle, sense, and information content', *Mind*, 101 (1992), pp. 767–91.

acquisition of new information may well be the result of semantically irrelevant psychological or pragmatic factors. And once it is shown that this is in fact the case, there will be no call for an independent theory of sense. Thus, for example, one of the leading neo-Russellians, Nathan Salmon, proposes:

> To be sure ⌜$a = b$⌝ *sounds* informative, whereas ⌜$a = a$⌝ does not. Indeed, an utterance of ⌜$a = b$⌝ genuinely imparts information that is more valuable than that imparted by an utterance of ⌜$a = a$⌝. For example, it imparts non-trivial linguistic information about the sentence ⌜$a = b$⌝ that it is true, and hence that the names *a* and *b* are co-referential. But this is pragmatically imparted information … If Frege's strategy is ultimately to succeed, a further argument must be made to show that the information imparted by ⌜$a = b$⌝ that makes it sound informative is, in fact, semantically encoded.[29]

As the neo-Russellian views matters, then, to account for the characteristic appearance of a difference in information content, with respect to sentence pairs of the sort that puzzled Frege, one need only suppose that the utterance of one of the sentences *pragmatically imparts* information that is not imparted by an utterance of the other; it is not required that the two sentences *semantically encode* different information.

This kind of proposal, however, immediately faces two general – and ultimately related – difficulties. The first arises as follows. To remain true to his Russellian starting point, the neo-Russellian cannot allow the content semantically expressed by the assertoric utterance of a sentence to be individuated by criteria any more fine-grained than those that individuate pieces of information. Consequently, he is committed to claiming that *what* is asserted when a competent and rational speaker assertively utters, say,

(7) Superman is Superman

or

(6) Superman can fly

(with the primary intention of asserting the proposition semanticly encoded in these sentences) will be precisely the same as what she would have asserted were she, respectively, to have assertively uttered

[29] Nathan Salmon, *Frege's Puzzle* (Cambridge, Mass.: MIT Press, 1986), pp. 78–9.

(8) Clark Kent is Superman

or

(9) Clark Kent can fly.

But most competent speakers of English share, and feel deeply committed to, the intuition that someone can *without logical inconsistency* assert what is semantically expressed by (7) or (6) while asserting the negations of (8) or (9) respectively, *viz.*,

(10) Clark Kent is not Superman

(11) Clark Kent cannot fly.

It is difficult to see, though, how the neo-Russellian can avoid crediting such a person with making logically inconsistent claims. After all, for the neo-Russellian, asserting what is expressed by (10) or (11) will involve asserting the negation of precisely what one would have asserted were one to have asserted, respectively, what is expressed by (7) and (6).

Moreover, on the neo-Russellian view, we are precluded from assuming that someone can assert what is (semantically) expressed by (7) or (6) without (indeed, without *thereby*) asserting what is expressed by (8) or (9) – again, contrary to the deeply felt intuitions of most competent speakers of English. According to their view, then, if we say of someone that she both asserts what is expressed by (7) or (6) but does not assert what is expressed by (8) or (9), *we*, it seems, must be credited with inconsistency! There is, however, simply no precedent in our ordinary practice of attitude ascription for crediting such ascriptions with inconsistency.

Neo-Russellians are aware that their views have these counter-intuitive consequences, and they have, in a variety of ways, made serious efforts to explain them away. The usual strategy is to argue that our offended intuitions are themselves the inevitable result of confusions bred of various sorts of pragmatic exigency and, thus, ought to possess no authority vis-à-vis our semantic theorizing.[30] It would take us too far afield to attempt to outline, let alone assess, all these efforts here. For present purposes, it is enough to appreciate the

[30] For an especially rigorous articulation of this strategy, see Soames, *Beyond Rigidity*, ch. 6, p. 131–46. Unfortunately Soames's important book appeared well after this paper was originally sent to the editors of this *Companion*. For various

source and nature of these counter-intuitive consequences – specifically the fact that they concern *logical* relations between assertoric contents – and to recognize that if their position is to be sustained, the neo-Russellian must successfully explain these intuitions away. We shall be returning to these issues again shortly.

The second sort of difficulty that neo-Russellian's face concerns whether or not it is plausible to suppose that we can adequately explain the sorts of cognitive achievements that concerned Frege by appealing only to the acquisition of information that is *pragmatically imparted*. Reflection, I submit, suggests that it is not. In the first place, coming to have metalinguistic beliefs of the sort gestured at by neo-Russellian proposals like Salmon's does not appear to be necessary for significant cognitive achievements of the kind that concerned Frege.[31] Even if we were to assume that possessing the kind of metalinguistic concepts and elementary semantic knowledge presupposed by their proposal is, as a general fact, necessary to count as a competent speaker, the neo-Russellian has offered no compelling reasons to believe that whenever we credit a speaker with cognitive achievements of the sort at issue we are obliged to assume that she has *deployed* those concepts and knowledge by coming to have the relevant metalinguistic beliefs. If not, then the acquisition of new *information* (understood in the Russellian way) – whether pragmatically imparted or semantically encoded – is not

reasons, it proved unfeasible to try to incorporate a critical discussion of the novel arguments and proposals that Soames there offers in defence of what I'm here calling neo-Russellianism. Some of the concerns raised here apply equally well and straightforwardly to Soames's proposals. But an adequate critical appraisal of other aspects of Soames's discussion would require additional discussion, for which unfortunately there was no room.

[31] When attempting to explain differences in cognitive significance of the sort that led Frege to distinguish the thoughts expressed by distinct referentially isomorphic sentences, it is, of course, open to neo-Russellians to appeal to any type of information that may be, more or less systematically, pragmatically imparted by such sentences. In particular, they need not – though they often do – appeal to *metalinguistic* information. Soames, for example, in *Beyond Rigidity* does not propose that the information that is pragmatically imparted, and that is standardly confused with the information semantically encoded, is best thought of as metalinguistic. Nevertheless, in my critical discussion I focus my attention on the sort of *metalinguistic* information that Salmon appeals to in the passage quoted above. It should be clear, though, how analogous arguments to those I give can be deployed vis-à-vis other proposals about what the relevant pragmatically imparted information might be.

what is relevant to the differences in cognitive potential of the sort that puzzled Frege.

Given our discussion in §1, none of this should come as a surprise. For there, recall, we saw that differences in the cognitive significance of distinct but referentially isomorphic sentences interested Frege only insofar as he took them to manifest underlying *logical* differences in the contents being expressed. If we grant that the contents semantically expressed by the relevant sentences do indeed differ in some logically relevant way, then what our last reflections show is the very unsurprising fact that a speaker's understanding of a sentence – in particular, her appreciation of the logical properties of its content – does not require, yet alone consist in, the speaker's 'taking in' any information (metalinguistic or otherwise) that the sentence makes merely *pragmatically* available. Failure to attend to or take in the relevant pragmatically available information need not preclude grasp of the relevant logical properties. But grasp of the relevant logical properties is all that is required to explain the relevant differences in cognitive potential.

Moreover, even in those cases where we can assume that the speaker has taken in the relevant metalinguistic information, it is hard to see how the acquisition of such information *by itself* is ever sufficient to account adequately for the sorts of cognitive achievement of interest to Frege. Suppose, for example, that Lois already correctly believes what is expressed by the following two sentences:

(7) Superman is Superman

(6) Superman can fly.

Ordinarily, we would say that were she now to come to believe what is expressed respectively by the sentences

(8) Clark Kent is Superman

and

(9) Clark Kent can fly

this would involve a significant cognitive achievement. According to a neo-Russellian, this might consist in Lois's coming to acquire some new, pragmatically imparted, metalinguistic information: presumably, in the case of (8), the information that 'Superman' and 'Clark Kent' refer to the same thing; and in the case of (9), something like the information that the reference of 'Clark Kent' can fly. But now why should acquiring such information be deemed valuable?

One might, of course, be interested in metalinguistic matters for their own sake; but any value accorded the acquisition of such metalinguistic information based on an interest of this sort cannot do justice to the importance we ordinarily attach to the cognitive achievements at issue here.[32] Presumably, then, it cannot be for its own sake that the acquisition of such information will be deemed of cognitive value in a sense relevant to Frege. It would seem, then, that acquiring metalinguistic information of the sort at issue can be deemed valuable, on whatever occasions it is, only because possessing such information, together with a rudimentary knowledge of semantics, will enable one to come to believe something *else* (typically, something non-metalinguistic) with a prior and independently acknowledged cognitive value. But now the problem should be evident. For what might this *something else* be except, in the first case, the belief that *Clark Kent is Superman* and, in the second case, the belief that *Clark Kent can fly*? Isn't this precisely what we would expect someone with the relevant metalinguistic beliefs plus the relevant rudimentary knowledge of semantics to conclude? But if this is right, then we shall be obliged to credit (8) and (9) with the cognitive values that they possess neither in virtue of their intrinsic information content – for that is the same as the information expressed, respectively, by (7) and (6) – nor in virtue of some pragmatically grounded accompanying information – for, in the cases at hand, that can play at best an instrumental role in enabling Lois to acquire these *new* beliefs. But, of course, once this is granted, no good reason remains for distinguishing the contents expressed by (8) and (9) from the contents of these new beliefs.

What these last reflections help us to see is that the neo-Russellians have not, in fact, successfully addressed the real challenge that led Frege to introduce his notion of sense. Their strategy is premised on the assumption that Frege, like themselves, always took differences in cognitive significance to manifest differences in information content, and that his error was to mistake pragmatically

[32] The reader should be reminded of the misgivings Frege voices in 'On sense and reference' about his earlier metalinguistic proposal for dealing with the problem of identity statements in the *Begriffsschrift* – especially his remark that, if the content of an identity sentence were construed in the metalinguistic way he had earlier proposed, we would express no *proper knowledge* by its means.

imparted information for semantically encoded information. But if there is anything in Frege's theoretical apparatus that can plausibly be supposed to manifest a concern with information content – as that notion is understood by the Russellian – it is not his theory of sense, but his theory of reference. For Frege, differences in Russellian information are best viewed as differences in referential truth conditions.[33] What Frege took himself to notice, though, was that despite having the same referential truth conditions – that is, despite semantically encoding the same information – two sentences might nevertheless differ in *logically* relevant ways. The challenge to which Frege's notion of sense was a response was to account for *that* possibility. If the neo-Russellian wants genuinely to engage with Frege, he will have to concern himself more directly with this challenge.

To do this, the neo-Russellian must say more about the relationship between information content and logic. This, I suggest, is the real lesson to be learned from the first of the two difficulties just examined. The neo-Russellian has, so far as I can tell, two options, neither of which seems very attractive. He might, for example, try to deny that there can be any difference in the logical properties of two contents that are not a direct reflection of differences in information.[34] But if the neo-Russellian takes this route, then – since he is unwilling to admit that the relevant sentence pairs encode different information – he is forced to claim that our ordinary practice of logically appraising assertion and belief is systematically muddled. If there are no logical differences between what is semantically expressed by the utterance of two sentences that encode the same Russellian information content, then when a person competently asserts both that Superman can fly and that Clark Kent cannot fly, we must be literally wrong to suppose – as we standardly do – that the claims she makes are logically consistent.[35] Consequently, accepting the present proposal would require a radical reconception of the

[33] I am, of course, ignoring here important differences between the ontology of Frege's theory of reference and the ontology of the Russellian's theory of information content. These differences do not, however, affect the present point.

[34] This is doubtless what the traditional Russellian should want to say at this point. And to the extent that he does, he will at least have managed to engage directly with Frege's challenge. Unfortunately, this does nothing to mitigate the other problems the traditional Russellian faces.

[35] Any such argument will be plausible, of course, only to the extent that one finds it plausible to suppose that merely pragmatic exigency can support a systematic

nature and epistemology of logic and, so, of the sort of normative control that we ordinarily take logic to have over our cognitive and linguistic practices. Articulating and defending such a reconception would certainly involve a direct engagement with Frege's challenge, but I am not aware of any neo-Russellian who has attempted to do so with any plausibility.

The neo-Russellian has an alternative – one that will not oblige him to call into question the standards by means of which we ordinarily assess the consistency or inconsistency of our assertions and beliefs. He can claim that it is not (or not solely) in virtue of their *contents* – of *what* we assert or of *what* we believe – that our assertions and beliefs are subject to logical appraisal in the ways that we ordinarily take them to be. In other words, to preserve our ordinary assessments of consistency and inconsistency, the neo-Russellian must radically sever the connection between the considerations he takes to be relevant to the individuation of contents and those he takes to be relevant to the logical appraisal of our assertions and beliefs. But if this is the route the neo-Russellian proposes to take, we have the right to demand of him, first, an account of the considerations he thinks *are* relevant to individuating content – an account that explains why it should *not* be in virtue of their contents that our assertions and beliefs are subject to logical appraisal in the ways that they are – and, second, an account of that in virtue of which our assertions and beliefs *are* subject to logical appraisal – an account that explains why these should not be relevant to the individuation of content. Without addressing these issues, the neo-Russellian has offered no serious challenge to Frege. For if Frege is right, it is precisely in virtue of the logical features of our assertions and beliefs, as these are manifest in our standard practice of (inter- as well as intra-subjective) logical appraisal, that it makes whatever sense it does to distinguish *what-is-asserted* from the asserting – that is, to credit our assertions with having objective contents at all.

practice of inconsistency of the sort to which the neo-Russellian appears to be committed. As a general methodological point, however, it would seem that any semantic theory that commits one to supposing that a certain very common linguistic practice has speakers systematically contradicting themselves, contrary to virtually all native speakers' pre-theoretical intuitions, ought to be adopted only as a last resort.

Whatever one thinks of the prospects for adequately responding to these demands, I hope it is clear by now that the real issue between Frege and the neo-Russellian has little to do with questions of sameness or difference of information content – at least not when it is individuated in the way Russellians individuate it. Rather, it has to do with much more profound issues about the nature of logic and, in particular, with how best to understand its normative status vis-à-vis our cognitive and linguistic practice. What Frege realized was that if we are to make sense of the normative demands that we ordinarily take logic to place on this practice, we must acknowledge that two assertions, or two beliefs, may differ in logically relevant ways despite the fact that they have the same referential truth conditions. Adequately accounting for this is exactly the challenge to which Frege took himself to be responding when he introduced his theory of sense. If I am right, the neo-Russellian has not yet so much as engaged with this challenge let alone met it.

5 A NEO-FREGEAN ALTERNATIVE

While the positive side of the neo-Russellian program faces daunting challenges, the critical side – directed against a specifically description-theoretic understanding of the sense of proper names – remains unaffected. Not surprisingly, a central feature of the neo-Fregean reaction to this critique is their insistence that nothing in the considerations that led Frege to introduce sense, nor anything about the work to which Frege puts the notion, commits him to a description-theoretic understanding of sense as it applies to proper names. Indeed, our discussion in §1 should already have made this abundantly clear. But even though neo-Fregeans reject the specifically description-theoretic commitments of their predecessors, most nevertheless retain the view that the notion of sense is to be understood primarily in epistemic terms. In particular, they continue to assume that for a speaker to be competent in the use of a name she must possess *some* determinate procedure or capacity for epistemically isolating an object as the reference of the name. According to these neo-Fregeans, the description theorists' approach failed primarily because they did not recognize the possibility of non-descriptive modes of identification or discrimination. As we shall see, however, by continuing to view the role of sense in these essentially epistemic terms, neo-Fregeans themselves end up imposing conditions on the individuation of sense that are neither required nor warranted by Frege's central (logical) concerns – conditions, moreover, that face difficulties strikingly similar to those faced by their description-theoretic predecessors.

To see what I have in mind, consider Michael Dummett's reconstruction of Frege's arguments for introducing the notion of sense.[36] As Dummett views matters, a theory of sense aims to provide, for each expression in the language, an adequate account of what a speaker knows just insofar as she understands that expression. And what a speaker knows when she understands an expression is, according to Dummett's reading of Frege, its reference.[37] What Frege argued, then, according to Dummett, is that it makes no sense to suppose that there could ever be *bare* knowledge of the reference of an expression. In the case of proper names, this amounts to saying that it is unintelligible to suppose that we can provide a *complete* account of a speaker's knowledge of the reference of some name *n* by means of an ascription of what Dummett calls 'predicative knowledge' – that is, by means of an ascription of the form ⌜S knows, of *a*, that *n* refers to it⌝, where *a* occurs transparently. Rather, the availability of any such ascription presupposes that the speaker possesses knowledge the proper ascription of which would take the form, ⌜S knows that *a* is *Φ*⌝, where

[36] My discussion of Dummett draws primarily from the following: 'Frege's distinction between sense and reference', in *Truth and Other Enigmas* (London: Duckworth, 1978), pp. 116–44; *Frege: Philosophy of Language* (London: Duckworth, 1973), esp. pp. 81–203; 'What is a theory of meaning?' parts I and II, in *The Seas of Language* (Oxford: Oxford University Press, 1993), pp. 1–93, though see esp. the Appendix to part I; and *The Logical Basis of Metaphysics* (Cambridge, Mass.: Harvard University Press, 1991), ch. 5, pp. 105–40.

[37] This way of putting the matter can be misleading. For familiar reasons, Dummett typically resists saying that understanding a *sentence* consists in knowing its reference, that is, knowing its truth-value, preferring instead to say that understanding a sentence – grasping its sense – consists in possessing (roughly) a *way of knowing* its truth-value. As grasp of the sense of a sentence is, on Dummett's view, derived from a grasp of the senses of its constituent expressions, we must understand grasping the sense of sub-sentential expressions accordingly. A more accurate representation of his view here, then, will be that understanding a name *a* consists in possessing a *way* of knowing its reference. Understanding a predicate *F* will consist in possessing a *way* of knowing whether or not something satisfies it. When combined, they provide us with a *way* of knowing the truth-value of the sentence *Fa*.

a occurs obliquely – what Dummett calls 'propositional knowledge'. According to Dummett, to count as competent in the use of a name, the sorts of propositional knowledge that a speaker needs to possess must provide her with a method or procedure for recognizing an object as the reference of the name.[38] But it follows from none of this, Dummett insists, that the relevant propositional knowledge must always be such as to provide a speaker with the sort of uniquely individuating *descriptive* knowledge presupposed by description theorists.[39]

To the extent that it is convincing, the argument so far establishes that for a speaker to associate a reference with a name, she must attach some sense or other to it – which, if Dummett is right, means that she must have some way of, or procedure for, recognizing an object as the reference of the name. It does not, however, provide any reason to suppose that different speakers must attach the same sense, the same way of recognizing an object, to any one name. Nor, in fact, does it even show that a single speaker must, when she uses the same name on different occasions, attach the same way of recognizing its reference. But why then should we take the sense of a name – conceived so as to tolerate such widespread inter- and intra-subjective variability – to be any part of what is *semantically* conveyed by the assertoric utterance of a sentence in which the name occurs rather than, say, merely part of a subjective psychological mechanism by which speaker's attach a semantic value to the name? In short, it would be implausible to view sense, so conceived, as possessing the sort of objectivity that was essential to Frege's conception of it.

Dummett is fully aware of these limitations and sees Frege as offering a second argument that complements the first and fills the gaps. This second argument concerns how we must understand what is involved in knowing the reference of a name if we are to make sense of the use of language in communicating knowledge.

[38] Of course, to support the idea that different names of the same object can have different senses, we need also to assume that for any object there can be distinct backings of propositional knowledge sufficient to provide a speaker with distinct procedures for recognizing that object, and that it is possible for a speaker with two such procedures not to recognize that they target the same object.

[39] See Dummett, *Frege: Philosophy of Language*, pp. 97–8, 110–11, and 'What is a theory of meaning?', p. 24.

To avoid the shortcomings of the first argument, we need reasons to suppose that for different speakers to count as competent in the use of the same name they must attach the *same* sense to it. For only then is it plausible to suppose that the sense of a name is something *semantically* conveyed by the utterance of a sentence containing it, rather than a merely subjective accompaniment. As Dummett views matters, these reasons are provided as soon as we acknowledge that the semantic content of a sentence is to be identified with what *any* speaker competent in the use of that sentence would come to know in recognizing the truth of what she understands when she understands a competent utterance of the sentence. And this is something we are obliged to acknowledge if we are to understand the use of language in communication. For, it is evident that it will not suffice for a hearer to count as having understood an assertoric utterance of a sentence of the form $\ulcorner a = b \urcorner$ if the content she ascribes to it is the same as the content of the knowledge she would possess simply in virtue of knowing the truth of what is expressed by a corresponding sentence of the form $\ulcorner a = a \urcorner$, or *vice versa*. More generally, to have communicated successfully, the content assigned to a sentence upon hearing it must be the same as the content expressed by the competent assertion of that sentence; and, for this, it is not enough that the contents assigned by speaker and hearer merely possess the same referential truth conditions.[40]

Since the extra-linguistic knowledge that any competent speaker will possess simply in virtue of knowing the truth of what is expressed by a sentence of the form $\ulcorner a = a \urcorner$ is distinct from the extra-linguistic knowledge that she will possess simply in virtue of knowing the truth of what is expressed by a corresponding

[40] Plainly, it is a necessary condition on the acquisition of knowledge by testimony that one correctly *understand* the assertion of the speaker. If what one takes some speaker to have asserted has the same referential truth conditions as what he in fact did assert, this will not, other things being equal, suffice for the transmission of knowledge, for this is insufficient for understanding. So, for example, if Jimmy comes to believe that Clark Kent can fly directly on the basis of Lois's having asserted 'Superman can fly' – the former being what he took her to have asserted – Jimmy's new true belief would not count as *knowledge*. For an illuminating discussion of the important role considerations about knowledge transmission play in motivating the semantic importance of a notion of sense, see Richard Heck, 'The sense of communication', *Mind*, 104 (1995), pp. 79–106.

sentence of the form ⌜$a = b$⌝, we must suppose that these sentences differ in the content they semanticly express. Add compositionality, and we can now be assured that the contribution *b* makes to the semantic content of the sentences in which it occurs must not only be distinct from that made by *a* – despite the fact that they have the same reference – but must also be something the grasp of which is common to speakers who share a competence in the use of these names. If, as Dummett suggests, grasping the sense of a name consists in possessing a procedure for recognizing an object as its reference, then it would seem that sharing a competence in the same name will have to consist in possessing the same recognitional procedure.[41]

Though Dummett has more to say about these matters, I can now state my principal misgiving: to the extent to which they are persuasive, it follows from neither of Dummett's arguments, either taken by themselves or together, that to be competent in the use of a name a speaker must possess a procedure for correctly recognizing an object as the reference of the name – let alone the same procedure as other speakers competent in the use of the same name. There are, moreover, independently compelling reasons – reasons analogous to those raised against the description theorist – for supposing that possession of any such a procedure is not required.

Consider again the first argument Dummett offers on Frege's behalf. What is there in the idea that predicative knowledge presupposes propositional knowledge – insofar as the idea is at all plausible – that by itself obliges us to assume that propositional knowledge of the sort necessary to underwrite an ascription of the form ⌜S knows, of *a*, that *n* refers to it⌝ must provide the speaker with such a procedure? For many, the only sort of propositional knowledge that is always relevant to underwriting the legitimacy of ascriptions such as

(i) ⌜S knows, of *a*, that *n* refers to it⌝

is plausibly the sort of propositional knowledge we attribute by means of some ascription of the form

(ii) ⌜S knows that *n* refers to *b*⌝

[41] See Michael Dummett, *The Interpretation of Frege's Philosophy* (London: Duckworth, 1981), pp. 186–7.

where *b* occurs obliquely and refers to the same object that *a* does in (i).[42] Such a proposal avoids the difficulties facing any assumption that understanding a name might consist in a *bare* knowledge of its reference – for a speaker competent in the use of the distinct co-referring names *n* and *m*, may know that *n* refers to *b* even though she does not know that *m* refers to *b*. It also avoids the shortcoming that Dummett's first argument faced – for knowledge of the sort ascribed by a particular instance of the form (ii) is, plausibly enough, just the sort of propositional knowledge (if any) that we should expect *any* competent user of the name *n* to possess.

Though Dummett seems inclined to accept much of this, he wants to insist that in order correctly to ascribe knowledge of the sort represented by (ii) – where the truth of such an ascription suffices for competence in the use of a name – we are obliged to suppose that the subject of such ascriptions possesses a procedure for recognizing an object as the reference of the name. But why? As far as I can tell, Dummett believes that we are so obliged, for otherwise we shall not be able adequately to distinguish knowledge of the sort ascribed by instances of (ii) from knowledge of the sort that would be ascribed by corresponding instances of

(iii) S knows that ⌜*n* refers to *b*⌝ is true.

And while Dummett is surely right that knowledge of the latter sort is insufficient for knowledge of the first sort – and, so, is insufficient for understanding a name – this fact alone does not establish that knowledge of the first kind involves possession of a recognitional capacity of the sort Dummett proposes. Additional substantive assumptions are clearly necessary. In Dummett's case, these additional assumptions derive, I believe, from his verificationism together with his commitment to a particular (itself verificationist-inspired) understanding of what has come to be known as the 'manifestation requirement' – roughly, the claim that knowledge of meaning must be fully manifest in linguistic practice.[43] Dummett's discussion of these issues is too rich and complex to do it full justice

[42] In the standard case, the substituend for *b* will be a used instance of the name mentioned by the substituend for *n* – for example: Van knows that 'Gottlob' refers to Gottlob.

[43] See the Preface to Dummett, *The Seas of Language*, pp. xii–xv, but also 'What is a theory of meaning?', pp. 37–8, 46–7, 52, 91–2.

here. Nevertheless, by carefully reflecting on our ordinary name-using practice, a strong case can, I believe, be made to show that possessing knowledge of the kind ascribed by instances of (ii) – at least where this is taken to ascribe competence in the use of a name – neither requires, nor is guaranteed by, the possession of a procedure for recognizing some object as the reference of the name.[44]

An examination of our standard practice of using names reveals, I suggest, that there is little more plausibility in supposing that competent name users always possess a procedure for recognizing some object as the reference of a name they use – or, when they take themselves to have such a procedure, that the object (if any) identified by its deployment must in fact be the reference of the name – than there was in supposing that speakers must be able to individuate descriptively the reference of the names they are competent to use. Consider such names as 'Isaac Newton', 'Kurt Gödel', 'Richard Feynman', 'Kosovo' or, for that matter, the names of many non-famous, ordinary persons or places, picked up here and there in conversation. Many people whom we would ordinarily unhesitantly regard as competent in the use of such names not only fail to possess uniquely individuating descriptive knowledge of their references (as Kripke showed), but also fail to possess – indeed, would deny possessing – a procedure enabling them correctly to recognize (or otherwise uniquely identify) their references.

Whether or not someone takes an object he is presented with to be the reference of a name he is competent to use will surely depend both on his current beliefs and on the manner in which the object is presented. But no matter how an object is presented, it seems always possible that an individual's beliefs may be such that he has good reasons for failing to recognize the object so presented as the reference of a name he uses, *without this in any way undermining his claim to competence in the use of that name*. Nor is it plausible to suppose that, whenever a speaker is competent in the use of the name, he must possess a set of beliefs that guarantees that there is some way of presenting the reference to him such that he cannot reasonably fail to recognize the object so presented as the reference

44 Moreover, I believe, the reasons we have for accepting this are ultimately more compelling than any reasons we might have for accepting Dummett's additional verificationist assumptions – but this is not something I will try to defend here.

of the name. But if this is right, then it cannot be required for a person to be competent in the use of a name that she possess a procedure for recognizing the reference of that name *as* its reference.

Needless to say, a great deal depends on what Dummett takes to constitute possessing a procedure for recognizing an object as the reference of a name – where this counts as grasping the sense of the name – and, in particular, on how such procedures are individuated. Unfortunately, Dummett is not as forthcoming on this matter as one might wish. There may be some plausibility in supposing that any person whom we are willing to grant competence in the use of a name *n* may be said to possess a capacity for recognizing (or identifying) an object as the reference of that name – where such recognition will be manifest in her being able to assert truthfully and with warrant a sentence of the form ⌜*d* is *n*⌝, where *d* is a demonstrative (or, when demonstration is out of the question, a non-circular uniquely individuating description). It may be plausible, for example – though I doubt that even this much is true – that any competent user of a name *n*, simply in virtue of possessing the kind of rationality and the very general sorts of epistemic prowess that we would expect any language user to possess, will have the *capacity* to acquire, either through instruction or directed inquiry, enough knowledge to enable her with warrant to acknowledge of the reference of the name, when presented to her, that *it* is *n*. But possessing this general capacity (even if we assume that every speaker has it) surely cannot be counted as grasping the sense of any *particular* name, for precisely the same general capacity would be deployed in coming to be able to recognize the reference of any name.[45]

[45] I can think of two ways in which one might try to narrow the description of this capacity so as to overcome this objection. On the one hand, one might suppose that the description of an instance of the general capacity that counts as grasping the sense of a particular name will make essential reference to the referent of the name. But even if this could be motivated in a non-ad hoc way, it would not individuate capacities sufficiently finely. For such descriptions will not distinguish the senses of distinct coreferential names. On the other hand, one might suggest that the relevant instance of the general capacity is to be determined by reference to the speaker's beliefs concerning *n*'s reference, the beliefs, say, that would form the starting point of any instruction or inquiry. But this individuates capacities too finely, for not only will different speakers who share a competence with a given name have different beliefs concerning the reference, but the beliefs that an individual speaker has will themselves change substantially over time. It might help here if we could suppose that for every name there is a canonical

The description theorist's principal mistake, then, was not to have had an overly restrictive view of the epistemic resources by means of which speakers can cognitively isolate objects, but rather to have insisted at all that competence with a name is grounded in possessing some specific, independent means of epistemically isolating the object that is in fact its reference from all other objects. And this is a mistake that neo-Fregeans of the Dummettian mould also make. After all, isn't it intuitively obvious that in many cases one's ability to isolate some object cognitively – to think about just that object and not any other – depends upon one's ability to use a name with that object as its reference rather than vice versa?[46] If this is right, then distinguishing knowledge of the sort ascribed by correct instances of (iii) from knowledge of the sort ascribed by correct instances of (ii) – where the latter are taken to ascribe competence in the use of a name – cannot require that competent name users possess the sort of recognitional capacity that Dummett appears to presuppose.

Moreover, once we think about the procedures that individual speakers might actually have (or take themselves to have) for recognizing an object as the reference of a name in their repertoire, it is obvious that these can vary greatly between speakers – speakers,

set of beliefs about its reference that any speaker competent in the use that name possesses, and which will be different for all names differing in sense – even co-referring names. But, so far as I can tell, there is no good reason to suppose that this is the case.

[46] This claim may appear to be inconsistent with the seemingly truistic principle that competence with a name requires knowing its reference. But whether or not one will accept this principle – indeed, its very status as truistic – depends crucially on *what* one takes to be involved in knowing the reference of a name. Our conclusion above is inconsistent with the principle only if one adopts certain non-obligatory views about what counts, in this context, as *knowing the reference* – substantive views that, when made explicit, drain the principle of its truistic status. For a person to count as competent in the use of a name, he must be able to use it in sentences to say true and/or false things about its reference. Of course he could not do this if he could not *in some sense* discriminate the object that is in fact its reference from all others. But this, so far as I can tell, exhausts whatever truth there is to the principle. Nothing, then, precludes us from supposing that the *relevant* discriminative capacity is something the speaker may possess only in virtue of having mastered the name. If so, then mastering a name need not require the speaker to possess independent epistemic capacities of the sort Dummett supposes, but will depend instead on the speaker's being properly situated in a physical and/or social environment that relates her in an appropriate way to the reference. See the end of §3.

for example, with substantially different background beliefs – all of whom we would ordinarily regard as competent users of the same name. The procedures that individual speakers use can also change substantially as they acquire new beliefs about the reference and/or about related features of the world. Thus, even in those cases where speakers can be said to possess a procedure for recognizing an object as the reference of a name they use, the proposal that this is what their grasping the sense of the name consists in does not cohere well with the considerations Dummett raises in the second argument he attributes to Frege.

Dummett is surely right that any account of sense that allows for the sort of widespread inter- and intra-subjective diversity that we have been considering will be of a notion that fails to possess the kind of objectivity Frege insisted upon and, as such, will fail to contribute to an adequate understanding of linguistic communication. But unless Dummett can avoid the criticisms just raised, his own proposal also fails in precisely this regard.

Though Dummett's second argument lends no direct support to his specific proposal concerning the sense of proper names, it does, I believe, establish that an adequate account of the semantic content of a sentence must satisfy Frege's demand for objectivity. Just how best to understand that demand is not, however, altogether clear. Here, of course, I can only scratch the surface of this very large and controversial topic. What will emerge, though, is that a proper understanding of Frege's demand requires that logical considerations be brought back into the forefront of our understanding of sense.

It often seems as though Frege takes the objectivity of thoughts to be evidenced by – perhaps, in some sense, even to consist in – their distinctive inter-subjective availability.[47] For example, we often find him claiming that while two individuals can never have one and the same pain, they may very well entertain one and the same thought.

[47] From the outset, it is important to note that the objectivity Frege insisted upon was always in the first instance the objectivity of *thoughts*. Indeed, his most compelling arguments for their objectivity – viz., those involved in his criticisms of psychologism – depend essentially on the fact that it is thoughts that are his primary object of concern. There is no direct way in which these arguments can be reconstructed where the sole focus of concern is the senses of sub-sentential expressions. Needless to say, Frege does also insist that the senses of sub-sentential expressions are objective too, but their objectivity is essentially derivative from their being possible constituents of thoughts. Appreciating this is, I believe, essential to understanding the issue at stake for Frege.

But remarks like this can easily be unsatisfying, and potentially quite misleading. Isn't there, after all, a fairly straightforward sense in which both you and I *can* have the same pain – the same pain *type*? And so, why isn't our entertaining the same thought also simply a matter of our having mental states of the same type?

This is the point, unfortunately, where many Fregeans are tempted to appeal directly to Frege's view that thoughts are themselves mind-independent objects. The difference between our entertaining the same thought and our feeling the same pain will then be explained by saying, in the case of thoughts, that our distinct mental states count as being of the same type by virtue of their standing in an appropriate relation to the same mind-independent object – the thought – whereas in the case of pains, there is no such object to which our mental states are related and in virtue of which they count as being of the same type. According to this way of understanding the matter, the objectivity of thoughts would seem simply to consist in their *being* mind-independent objects.

This suggestion, however, is not very satisfying. It faces a variety of familiar difficulties – not the least of which are how, on this picture, we are supposed to understand the nature of these mind-independent thoughts, and how, given that nature, we are to understand the special relation between our minds and thoughts in virtue of which the latter become the transparent contents of our thinking. And even though it is true that thoughts are, for Frege, objects, it is noteworthy that whenever he attempts to explain and defend his claim that thoughts are *objective* – where, notice, the relevant contrast is always with the *subjective* contents of mental states he generically calls 'ideas' – he never directly appeals to their status as objects in the manner just suggested. Frege seems instead to have thought that the objective status of thoughts is fully revealed in just those features of our cognitive and linguistic practice that manifest their constitutive connection with *truth* and, so, with logic. The *distinctive* sense in which thoughts may be shared – that is, the sense in which you and I may be said to have the same thought but not the same pain – is revealed by reflecting on these practices.[48]

[48] For an excellent discussion of these and related issues see Thomas Ricketts, 'Objectivity and objecthood: Frege's metaphysics of judgement', in L. Haaparanta and J. Hintikka (eds.), *Frege Synthesized: Essays on the Philosophical and Foundational Work of Gottlob Frege* (Dordrecht: Reidel, 1986), pp. 65–95.

To see what is at issue here, reflect briefly on how, in the most straightforward cases, we use language to communicate agreement and disagreement in judgement. Suppose that both you and I share a language, and that you assertively utter a sentence *S*. Your utterance is sincere, and you intend your words to be taken literally. If I understand your utterance, I find myself immediately in a position to agree or disagree with what you have said.[49] Suppose I find that I agree and, moreover, that *I* now want to assert that about which we agree. The most straightforward way to do this will usually be for me also to utter *S* assertively. It is clearly Frege's view that to count as agreeing – to count as having asserted the *same* thing – we must each, in uttering *S*, express the same thought. And for this, as we have already seen, it will not suffice if what we express merely has the same referential truth conditions. Sentences with the same truth conditions can vary significantly in their inferential potential; and a difference in inferential potential can suffice to undermine a genuine agreement.

Alternatively, suppose that we disagree. Suppose, that is, that I believe the negation of what I correctly understand you to have asserted. The most straightforward way for me to make manifest our disagreement would be for me assertively to utter ⌜¬*S*⌝. Here again, we would not count as disagreeing in the relevant sense – what I asserted would not count as the *contradictory* of what you asserted – unless we both attached the same thought to *S*. Nor will it suffice for my assertion to contradict yours if *S*, as each of us uses it, merely has the same referential truth conditions.

To agree or disagree, we must, in the sense relevant to Frege's concern with their objectivity, entertain the *same* thought. And what is required for this is fundamentally different from anything that may be required in order correctly to say that we feel the same pain. Any sense in which we may be said to agree (or disagree) in the pain we feel is fundamentally different from the sense of agreement (or disagreement) involved when you and I both assert the same thought (or one of us asserts the negation of the thought asserted by the other). For our assertions to express genuine agreement or disagreement in judgement, they must be viewed as *sharing*, either wholly or in part,

[49] For simplicity, I ignore the possibility of simply suspending judgement. I am also assuming that *S* does not contain any relevant indexical devices.

truth-evaluable contents possessing the same inferential potential. In the example, the thought I associate with *S* is true just in case the thought you associate with *S* is true; and whatever logical relationships my thought bears to other thoughts – whether my own or another's – so must yours, and vice versa. When Frege insists that thoughts are objective and not subjective, his intention is to draw attention to a fundamental contrast between the contents constitutive of such mental states as judgements – which, in aiming at the truth, are essentially subject to normative governance by the principles of logic – and the contents constitutive of such mental states as pains, imagings, aesthetic feelings, etc. (i.e. of ideas) – which are not. In other words, to insist that thoughts are objective and not subjective is just to insist that thoughts, but not the contents of ideas, are by their nature subject to evaluation as true or false and, so, to comparative logical assessment. Thoughts count as objective, in the sense relevant to Frege, precisely in virtue of this (constitutive) capacity to stand in determinate *logical* relations to other thoughts, no matter who may be entertaining them or at what time. The only notion of *sameness* relevant to Frege's insistence on the objectivity of thoughts (or, for that matter, thought-constituents) is the notion to which we are obliged to appeal in order to make sense of these logical relations. And there is simply no relevant analogue to this in the case of pains or other mental states whose contents Frege would regard as subjective.

In order for what one asserts to stand in relations of agreement or disagreement – or, more generally, logical relations – to what another speaker asserts by uttering the sentences he does, it is not, of course, essential that we actually be speaking the same language. All that is required is that we attach the same thoughts to the relevant sentences. But if we *are* speaking the same language, and if we are using the same sentences in the manner suggested above, then – assuming the compositionality of sense – to make sense of the possibilities of agreement and disagreement just noted, we shall have to attach the same sense to the sub-sentential expressions out of which the sentences are formed. The trouble, then, with a conception of sense that permits widespread inter- and intra-subjective variability in the senses that speakers attach to the sub-sentential expressions in their languages is that it threatens to seriously undermine our conception of sense as objective in this way. It threatens to cripple our ability to make sense of and do justice to our common, everyday judgements of agreement and disagreement as they concern what we and our fellow speakers say. More generally, it precludes our being able to make sense of and do justice to our ordinary assessments regarding what logical relations hold between our own assertions and those of others.

6 TOWARD A MORE LOGICAL APPROACH TO SENSE

These last considerations bring us right back to our starting point – the idea that Frege introduced the notion of sense in the first instance to subserve purely logical considerations, and that questions of sense identity and difference are ultimately to be answered by the constitutive demands we take logic to place on our cognitive and linguistic practice – demands that the resources provided by a theory of reference alone cannot accommodate. Accordingly – given Frege's commitment to compositionality – two coreferential sub-sentential expressions will differ in sense only if what is expressed by the sentence that results from substituting one for the other has different logical properties from the original. Nothing belongs to the sense of a sub-sentential expression except the systematic contribution it makes to the logical properties possessed by the sentences in which it occurs.

On the present proposal, then, grasping the sense of a proper name will consist in grasping the contribution the name makes to the logical properties of the sentences in which it appears – a contribution not exhausted by the name's having the reference it has. Once we give logical considerations proper pride of place over epistemic considerations in our understanding of sense, a principal rationale for supposing that grasping the sense of a name must consist in or require possessing a capacity to isolate its reference epistemically – either of the sort proposed by the description theorist or by the Dummett-inspired neo-Fregean – loses much of its attraction. And this is just as well, for it should be quite clear by now that no such capacity is in fact required.

This leaves us with two large and intimately related questions. First, how exactly are we to understand the contribution a particular name (or, for that matter, any other sub-sentential expression) makes to the logical properties of the sentence in which it occurs?

For example, when the substitution of distinct coreferential names alters the logical properties of the sentences in which they occur, how are we to think about the differential contribution these names make? And second, in what does the capacity to grasp this contribution consist? Since in order to be credited with a capacity to grasp this contribution, it is not required that a speaker possess an independent capacity to individuate epistemically the reference of a name she is competent to use, what is required?

While the issues raised by these questions are much too large and complicated to be addressed adequately here, I can provide some indication of the theoretical options open to us and, so, of the shape answers to these questions might take once we adopt the logic-based conception of sense being proposed here. As soon as we acknowledge that it is primarily logical and not epistemic considerations that are central to the individuation of thoughts, serious difficulties arise for an all too common assumption concerning the shape that an adequate answer to our questions must take. The problem I have in mind arises as soon as one assumes that, in order adequately to answer them, the account we give of the contribution made by a particular sub-sentential expression to the thoughts expressed by the sentences in which it occurs must provide what I shall call a *compositional explanation* of what is semantically distinctive about those thoughts. In the context of our logic-based reorientation of the notion of sense, those who suppose that providing a theory of sense commits us to engaging in the explanatory project I have in mind will insist that we compositionally *explain* the distinctive logical properties of the sentences in which they occur by appealing to what is distinctive about the senses associated with different sub-sentential expressions. In particular, they will view differences in the logical potential of otherwise referentially isomorphic sentences as *resulting from* specific differences in the senses of the sub-sentential expressions making up the two sentences.[50]

Explanatory strategies of this sort are committed to viewing the senses of sub-sentential expressions as explanatorily prior to the

[50] I am not clear whether it is the temptation to make this assumption that has encouraged adoption of one or another of the epistemic approaches to sense that we have been discussing or vice versa. In any case, this assumption and epistemic approaches to sense seem to be natural bedfellows, and once one gives up on the one, the attraction of the other is substantially diminished.

logical properties of the sentences in which they occur. But an appeal to differences in the senses of sub-sentential expressions will afford us a genuine *explanation* of differences in the logical properties of the sentences in which they occur only if we are offered a substantive account of the notion of sense that tells us what precisely is involved when expressions differ in sense – an account, moreover, that does this *independently of an appeal to the very logical differences* (or anything that depends upon them) *that such differences in sense are invoked to explain compositionally*. Moreover, we must be shown how differences of the sorts *thus specifiable* are – indeed, how they can be – determinative of the logical properties of the sentences containing them.

While I know of no decisive argument to show that a satisfactory account of sense meeting these requirements is impossible, I also know of none that is on offer. And I am deeply sceptical that any such account is likely to be forthcoming.[51] However, even if I am right about this, it would be premature to suppose that the demand for a theory of sense is unsatisfiable and so misguided. For while an adequate theory of sense must, I believe, have something to say in response to the two questions with which we began, once we adopt the approach to sense suggested here, we find that a compositional-explanatory approach is not obligatory.[52]

[51] For a more detailed discussion of some of the reasons for my scepticism here, see William Taschek, 'On ascribing beliefs: Content and context', *Journal of Philosophy*, 95 (1998), pp. 323–53.

[52] It is tempting to suppose that Frege himself was engaged in just the sort of compositional explanatory project that I am criticizing. From this perspective, Frege's introduction of the notion of sense will be understood as a sort of theoretical posit: There must be something about referentially isomorphic sentence pairs such as *Fa* and *Fb* that explains why they express different thoughts. But the only difference is that one has *b* where the other has the distinct but coreferential *a*. So it must be that *b* possesses some semantic property different from that possessed by *a in virtue of which* the sentences express different thoughts. Call this property the 'sense' of these expressions. Whether in fact Frege conceived of his theory of sense in these compositional-explanatory terms is not something I am concerned to settle here – though I think there are good reasons to be suspicious. After all, apart from a few familiar metaphors concerning modes of presentation and the like, Frege had surprisingly little to say about how exactly we are to conceive of the senses of individual words. Given this silence, we certainly cannot credit him with having tried very hard to provide a compositional explanation of how the senses of distinct but coreferential names differentially contribute to determining the logical properties of the sentences in which they appear.

If, as I have been suggesting, the sense of a name is to be thought of as the contribution it makes to the logical properties of the sentences in which it occurs, then a theory of sense for a language containing that name will answer our first question if the logical properties of any sentence in which the name figures are made evident from what the theory states. But now it would seem that a theory of meaning broadly of the sort originally recommended by Davidson, and endorsed more recently by John McDowell and David Wiggins – that is, a theory of meaning possessing the shape of a truth theory more or less in the style of Tarski – may well supply the resources for accomplishing precisely this.[53]

A theory of this form will serve as an adequate theory of sense if, for each sentence of the object language, there is a canonically derived T-sentence the right-hand side of which specifies the content of – the thought expressed by – the sentenced mention the left-hand side.[54] The right-hand sentence will count as content-specifying in this sense just in case it can in turn be used as the content clause of a true (oblique or *de dicto*) ascription of *what is said* when a competent speaker of the object language utters the mentioned object language sentence. And, as I have argued elsewhere, a sentence can only be so used if, when occurring as the content clause in such an ascription, it possesses logical properties corresponding to those possessed by the mentioned object language sentence.[55] In any case, given our discussion so far, if the right-hand sentence of any canonically derived T-sentence possesses not only the same referential truth conditions as the mentioned object language sentence but also corresponding logical properties, then we have no reason to deny

[53] See, for example, the essays in Donald Davidson, *Inquiries into Truth and Interpretation* (Oxford: Clarendon Press, 1984), especially those in the first section; John McDowell, 'On the sense and reference of proper names', *Mind*, 86 (1977), pp. 159–85; and David Wiggins, 'Meaning, truth-conditions, proposition: Frege's doctrine of sense retrieved, resumed, and redeployed in the light of certain recent criticisms', *Dialectica*, 46 (1992), pp. 61–90.

[54] The requirement that the derivation be 'canonical' is to avoid irrelevant detours through logical equivalences that the deductive apparatus will doubtless make possible and, thus, to avoid its appearing that all of the logical equivalents of a sentence have the same sense as it does. See Wiggins, 'Meaning', p. 66, fn. 6; also see Scott Soames, 'Truth, meaning, and understanding', *Philosophical Studies*, 65 (1992), p. 34, fn. 11, for a more sceptical reaction.

[55] See my discussion of the 'logic requirement' in 'On ascribing beliefs'.

that the right-hand sentence exhibits the thought expressed by the object language sentence.

For such a theory to address the first of our two questions, it must enable us to appreciate the contribution that names make to the logical properties of the sentences in which they occur. And if, for the time being, we allow ourselves a host of simplifying assumptions, it is fairly easy to see how such a theory might do this. For each name in the object language, a theory of the kind under consideration will have as axioms instances of something like the following form:

(n) '**ν**' refers to **μ**

where **μ** may – indeed, perhaps, typically will – be replaced by the very name mentioned to the left (but see below). These axioms will be essential to the canonical derivations of T-sentences for any sentence of the object language in which the name is used. The role the axiom plays in the canonical derivation of the relevant T-sentences makes evident not only the contribution the name makes to the referential truth conditions of the sentences in which it appears but also and simultaneously the contribution the name makes to the *inferential* properties of the sentences in which it appears. At least it will do so as long as we properly restrict the allowable substituends for **μ** in axioms that deal with coreferential names.

Suppose, for example, that '*n*' and '*m*' are coreferential names. '*Fn*' and '*Fm*' have the same referential truth conditions but differ in logical properties. In our theory, the axioms dealing with these names will be something like the following:

'*n*' refers to *n*

'*m*' refers to *m*.

The canonically derived T-sentence for '*Fn*' and '*Fm*', respectively, will then be something like:

'*Fn*' is T iff *Fn*

'*Fm*' is T iff *Fm*.[56]

[56] Since we are assuming that the identity '*n* = *m*' expresses an extra-linguistic fact and, so, will not appear among the axioms of our theory, there will be no canonical derivation of the T-sentences '"*Fn*" is T iff *Fm*' or '"*Fm*" is T iff *Fn*'.

Our axioms guarantee that the logical properties of the sentences on the right-hand side of the canonically derived T-sentences corresponds precisely to those of the corresponding object language sentences. For the axioms to guarantee this, it is not, of course, essential that the *same* name be used on the right of 'refers to' as gets mentioned on the left. What *is* essential is that the name that does get used, apart from referring to the same thing as the mentioned name, be distinct (or contextually distinguishable)[57] from whatever name is used in the other axiom. But once this restriction is in place (appropriately generalized and systematically coordinated with the other axioms), it should be clear how the role our axioms play in canonical derivations of T-sentences will make the contribution the object language names make to the logical properties of the sentences in which they appear as evident as one has any right to demand – given our rejection of the compositional explanatory project. It does so by how it makes evident the different ways those names contribute to the referential truth conditions of the sentences in which they appear. And so a theory incorporating axioms of the sort under consideration appears to provide – assuming the rest of the theory is in order – a perspicuous non-*compositional-explanatory* answer to the first of our two questions.

Needless to say, a compelling defence of a proposal of this sort – even restricting our concern to how such a theory might handle proper names – would require considering a variety of more difficult cases than those in which the mere substitution of distinct co-referring names results in sentences with different logical properties from the original. Nevertheless, enough has been said, I believe, to suggest that something along the lines suggested here offers a promising way of systematically capturing and representing precisely

[57] While I believe that it might not yield a pragmatically perspicuous theory, I am inclined to think that there are no substantive semantic reasons for thinking that even the *same* name could not (at least on occasion) be used on the right-hand side of both of these axioms. This will be permitted, though, only if the *use* of this name as it occurs in the one axiom can be systematically distinguished from the *use* of the name as it occurs in the other in such a way that we can keep track of the distinction in our canonical derivations where, as a result of this keeping track, the metalanguage sentences appearing on the right-hand side of canonically derived T-sentences will be understood to have appropriately different logical properties from the metalinguistic sentence appearing on the right-hand side of the counterpart T-sentence canonically derived using the other axiom.

those features of sentences that Frege introduced the notion of sense to capture. Moreover – and, perhaps, most importantly – it does so in a way that reveals how addressing this task does not – contrary to what is so often taken for granted – require that one assume the burdens of the compositional explanatory approach.[58]

According to the picture I have been sketching, to grasp the sense of a proper name is to grasp the contribution it makes to the logical properties of the sentences in which it occurs. In what does the capacity to grasp this contribution consist? A fully adequate positive answer to this question is clearly beyond the scope of this discussion. Nevertheless, notice that once we abandon the compositional explanatory point of view, we are free to give a different kind of answer than we would be expected to give were we still under the spell of that approach. For an account of sense to play a compositional explanatory role, we would need to be able to specify the sense of an expression independently of any given thought. Grasping a sense – in particular, grasping the contribution a particular name makes to the logical properties of the sentences in which it occurs – would have to be viewed as a distinctive instance of a capacity that can be specified independently of our grasping the sense of any sentence in which that name appears, e.g. a capacity to cognitively isolate the reference.

On the present proposal, however, we are under no obligation to suppose that the capacity a speaker has to grasp the contribution a name makes to the logical properties of the sentences in which it occurs is specifiable in this way at all. Rather, on the present view, possessing such a capacity will consists in whatever it is about an individual in virtue of which it is correct to include an axiom of the appropriate sort in a truth theory capable of serving as an adequate theory of sense for her language. But the considerations relevant to this judgement include nothing more nor less than the sorts of

[58] A *prima facie* challenge to this approach – not obviously insurmountable – is provided by Kripke's 'Paderewski' example in 'A puzzle about belief', in A. Margalit (ed.), *Meaning and use* (Dordrecht: Reidel, 1979). Such cases suggest that the content semantically expressed by different utterances of a sentence containing the *same* name may nevertheless differ in their logical properties. See Taschek, 'On ascribing beliefs', for further discussion. Another difficulty concerns how to treat non-referring proper names – what will their axioms look like? See §8 of McDowell, 'On the sense and reference of proper names', for one suggestion.

consideration relevant to correctly interpreting his speech. And we have little reason to hope that in the case of proper names these can be reduced to a neat formula, epistemic or otherwise.

The phenomena that led Frege to introduce his distinction between sense and reference reveal that the considerations that govern the logical appraisal of our assertions and judgements require an appeal to something over and above their referential truth conditions. Sentences with the same referential truth conditions can differ in their logical properties. But how are we to understand this? How best are we to provide an account of the semantic content of our sentences that does adequate justice to this fact? The epistemic approach to sense – combined, as it typically is, with the compositional explanatory point of view – leads us down a blind alley. Once logical considerations are given their proper place in our understanding of Frege's concerns, the challenge to which his doctrine of sense was a response is no less than the challenge to make sense of the relationship between logic and meaning – indeed, between logic and our thought and talk generally. While we have made considerable headway understanding these issues over the last century, the challenge itself remains unmet. Indeed, while it is still far from clear what would constitute adequately meeting this challenge, I hope the considerations offered here point us in a more fruitful direction.

RICHARD HECK

9 Frege and semantics

I FREGE AND THE JUSTIFICATION OF LOGICAL LAWS

In recent work on Frege, one of the most salient issues has been whether he was prepared to make serious use of semantic notions such as reference and truth. Those not familiar with this debate are often surprised to hear of it. Surely, they say, Frege's post-1891 writings are replete with uses of 'true' and 'refers'. But no one wants to deny that Frege makes use of such terms: Rather, what is at issue is how Frege understood them; more precisely, what is at issue is whether Frege employed them for anything like the purposes for which philosophers now employ them. What these purposes are, or should be, is itself a matter of philosophical dispute, and, although I shall discuss some aspects of this issue, my goal here is not to address it directly. My purpose here, rather – one of them, anyway – is to argue that Frege did make very serious use of semantic concepts: In particular, he offered informal mathematical arguments, making use of semantic notions, for semantic claims. For example, he argues that all of the axioms of the Begriffsschrift – the formal system[1] in which he proves the basic laws of arithmetic – are true, that its rules of inference are truth-preserving, and that every well-formed expression in Begriffsschrift has been assigned a reference by the stipulations he makes about the references of its primitive expressions.

Let me say at the outset that Frege was not Tarski and did not produce, as Tarski did, a formal semantic theory, a mathematical

[1] Frege, like Tarski after him, does not clearly distinguish a formal *language* from a formal *theory* formulated in that language. I shall use 'the Begriffsschrift' to refer to the theory, and 'Begriffsschrift', without the article, to refer to the language.

definition of truth.[2] But that is not of any significance here. One does not have to provide a *formal* semantic theory to make serious use of semantic notions. At most, the question is whether Frege would have been prepared to offer such a theory, or whether he would have accepted the sort of theory Tarski provided (or some alternative), had he known of it. On the other hand, the issue is not whether Frege would have accepted Tarski's theory of truth, or Gödel's proof that first-order logic is complete, as a piece of mathematics;[3] it is whether he would have taken these results to have the kind of significance we (or at least some of us) would ascribe to them. Tarski's argument in 'The concept of truth in formalized languages' shows that all axioms of the calculus of classes are true; the completeness theorem shows that every valid first-order schema is provable in certain formal systems. The question is whether Frege could have accepted Tarski's characterization of truth, or Gödel's characterization of validity, or some alternative, *as* a characterization of truth or validity.

The issue is sometimes framed as concerning whether Frege was interested in justifying the laws of logic. But it is unclear what it would be to 'justify' the laws of logic. On the one hand, the question might be whether Frege would have accepted a proof of the soundness of first-order logic as showing that every formula provable in a certain formal system is valid. Understood in this way, the question is no different from that mentioned in the previous paragraph. Another, more tendentious way to understand the issue is as concerning whether Frege believed the laws of logic could be justified *ex nihilo*, whether an argument in their favour could be produced that would (or should) convince someone antecedently sceptical of their truth or, worse, someone sceptical of the truth of *any* of the laws of logic.

If *this* is what is supposed to be at issue,[4] then let me say, as clearly as I can, that neither I nor anyone else, so far as I know, has

2 A. Tarski, 'The concept of truth in formalized langaugues', in J. Corcoran (ed.), *Logic, Semantics, and Metamathematics* (Indianapolis: Hackett, 1958), pp. 152–278.

3 Burton Dreben was fond of making this point.

4 This notion of justification does seem to be the one some commentators have had in mind: See T. Ricketts, 'Generality, sense, and meaning in Frege', *Pacific Philosophical Quarterly,* 67 (1986), pp. 172–95, and J. Weiner, *Frege in Perspective* (Ithaca, N.Y.: Cornell University Press, 1990), p. 277.

ever held that Frege thought logical laws could be justified in this sense. Moreover, so far as I know, no one now does think that the laws of logic can be justified to a logical sceptic – and, to be honest, I doubt that anyone ever has.[5]

So in so far as Frege, or anyone else, thinks the laws of logic can be 'justified',[6] the justification envisaged cannot be an argument designed to convince a logical sceptic. But what then might it be? This is a nice problem, and a very old one, namely, the problem of the Cartesian Circle. I am not going to solve this problem here (and not for lack of space), but there are a few things that should be said about it.[7] The problem is that any justification of a logical law will have to involve some reasoning, which will depend for its correctness on the correctness of the inferences employed in it. Hence, any justification of the laws of logic must, from the point of view of a logical sceptical, be circular. Moreover, even if one were only attempting to justify, say, the law of excluded middle, no argument that appealed to that very law could have any probative force. But, although these considerations do show that no such justification could be used to convince someone of the truth of the law of excluded middle, the circularity is not of the usual sort. One is not assuming, as a premise, that the law of excluded middle is *valid:* If that were what one were doing, then the 'justification' could establish nothing, since one could not help but reach the conclusion one had assumed as a premise. What one is doing, rather, is appealing to certain *instances* of the law of excluded middle in an argument whose conclusion is *that* the law is valid. That one is prepared to appeal to (instances of) excluded middle does not imply that one cannot but reach the conclusion that excluded middle is

[5] I have heard it suggested that Michael Dummett believes something like this. But he writes: '[T]here is no sceptic who denies the validity of all principles of deductive reasoning, and, if there were, there would obviously be no reasoning with him' (M. Dummett, *The Logical Basis of Metaphysics* (Cambridge, Mass.: Harvard University Press, 1991), p. 204).

[6] Note that I am *not* here intending to use this term in whatever sense Frege himself may have used it. I am not concerned, that is, with whether Frege would have said (in translation, of course), 'It is (or is not) possible to justify the laws of logic.' I am concerned with the question whether Frege thought *that* the laws of logic can be justified and, if so, in what sense, not with whether he would have used (a translation of) these words to make this claim. Some commentators have displayed an extraordinary level of confusion about this simple distinction.

[7] There is now a fairly extensive literature on this problem.

valid: A semantic theory for intuitionistic logic can be developed in a classical metalanguage, and that semantic theory does not validate excluded middle. So the mere fact that one uses instances of excluded middle in the course of proving the soundness of classical logic need not imply that the justification of the classical laws so provided is worthless. If one were trying to explain *why* the law of excluded middle is valid, for example, a justification of it that employed instances of that very law might suffice.[8]

That would be one way of understanding what justifications of logical laws are meant to accomplish: They answer the question why a given logical law is valid. It suggests another. The objection that justifications of logical laws are circular depends upon the assumption that their purpose is to show that the laws are *true* (or the rules, truth-preserving). It will be circular to appeal to instances of the law of excluded middle in a justification of that very law only if the truth of instances of the law is what is at issue. But justifications of logical laws need not be intended to demonstrate their truth. We might all be agreed, say, that every instance of the law of excluded middle is, as it happens, *true* but still disagree about whether those instances are *logical* truths.[9] The purpose of a justification of a law of logic might be, not to show that it is true, but to uncover the source of its truth, to demonstrate that it is indeed a law of logic. It is far from obvious that an argument that assumed that all instances of excluded middle were true could not informatively prove that they were *logically* true.[10]

[8] The discussion in this paragraph is heavily indebted to Dummett, *The Logical Basis of Metaphysics*, pp. 200–4. It is also worth emphasizing, with Jamie Tappenden, 'Metatheory and mathematical practice in Frege', *Philosophical Topics*, 25 (1997), pp. 213–64, that an *explanation* of a fact need not amount to a reduction to simpler, or more basic, facts.

[9] For example, intuitionists accept all instances of excluded middle for quantifier-free (and, indeed, bounded) formulae of the language of arithmetic, on the ground that any such formulae can, in principle, be proved or refuted. Now imagine a constructivist who was convinced, for whatever reason, that *every* statement could, in principle, either be verified or be refuted. She would accept all instances of excluded middle as true, but not as logical truths.

[10] More generally, if one is to accept a proof that a particular sentence is logically true, one will have to agree that the principles from which the proof begins are true and that the means of inference used in it are truth-preserving. But one need not agree that the principles and means of inference are *logical:* The proof does not purport to establish that *it is logically true that* the particular sentence is logically true, only that it is logically true. And in model-theoretic proofs of

There is reason to suppose that Frege should have been interested in giving a justification at least of the validity of the axioms and rules of inference of the Begriffsschrift. Consider, for example, the following remark:

I became aware of the need for a Begriffsschrift when I was looking for the fundamental principles or axioms upon which the whole of mathematics rests. *Only after this question is answered* can it be hoped to trace successfully the springs of knowledge upon which this science thrives.[11]

Frege's life's work was devoted to showing that the basic laws of arithmetic are truths of logic, and his strategy for doing this was to prove them in the Begriffsschrift. But no derivation of the basic laws of arithmetic will decide the epistemological status of arithmetic on its own: It will simply leave us with the question of the epistemological status of the axioms and rules used in that derivation. It thus must be at least an intelligible question whether the axioms and rules of the Begriffsschrift are logical in character. What other question could remain?

The discussion that follows the passage just quoted reinforces these points. Frege first argues that epistemological questions about the source of mathematical knowledge are, at least in part, themselves mathematical in character, because the question what the fundamental principles of mathematics are is itself mathematical in character.

In order to test whether a list of axioms is complete,[12] we have to try and derive from them all the proofs of the branch of learning to which they relate. And in doing this it is imperative that we draw conclusions only in accordance with purely logical laws. … The reason why verbal languages are ill suited to this purpose lies not just in the occasional ambiguity of expressions, but above all in the absence of fixed forms for inferring. … If we try to list all the laws governing the inferences which occur when arguments are conducted in the usual way, we find an almost unsurveyable multitude which apparently has no precise limits. The reason for this, obviously, is that these inferences are composed of simpler ones. And

validity, one routinely employs premises that are obviously *not* logically true, such as axioms of set theory.

[11] Frege, 'On Mr Peano's conceptual notation and my own', in *CP*, p. 235 (362). (In this chapter page numbers in parentheses refer to the original publication of Frege's writings.)

[12] Note that Frege uses this term in a way that is close to, but not identical to, how it is standardly used nowadays.

hence it is easy for something to intrude which is not of a logical nature and which consequently ought to be specified as an axiom. This is where the difficulty of discerning the axioms lies: for this the inferences have to be resolved into their simpler components. By so doing we shall arrive at just a few modes of inference, with which we must then attempt to make do at all times. And if at some point this attempt fails, then we shall have to ask whether we have hit upon a truth issuing from a non-logical source of cognition, whether a new mode of inference has to be acknowledged, or whether perhaps the intended step ought not to have been taken at all.[13]

Much of this passage will seem familiar, so strong is the echo of remarks Frege makes in the Preface to *Begriffsschrift* regarding the need for a formalization of logic.[14] But the most interesting remark is the last one, which addresses the question what we should do if at some point we find ourselves *unable* to formalize the proof of a theorem previously proven informally. The most natural next step would be to try to isolate some principle on which the proof apparently depended, which principle would then be a candidate to be added to our list of fundamental principles of mathematics. Once we had isolated this principle, call it NewAx, there would be three possibilities among which we should have to decide: NewAx may be a 'non-logical' truth, one derived from intuition or even from experience; NewAx may be a truth of logic, which is what Frege means when he says that we may have to recognize 'a new mode of inference'; or NewAx may not be true at all, which is what Frege means when he says that the 'intended step ought not to have been taken'. Frege is not just describing a hypothetical scenario here: Frege had encountered just this sort of problem on at least two occasions. I have discussed these two occasions in more detail elsewhere.[15] Here, let me just summarize those discussions.

Frege begins his explanation of the proof of the crucial theorem that every number has a successor by considering a way of attempting

[13] Frege, 'On Mr Peano's conceptual notation', p. 235 (362–3).

[14] Frege, '*Begriffsschrift*: A formula language modeled upon that of arithmetic, for pure thought', in J. van Heijenoort (ed.), *From Frege to Gödel: A Sourcebook in Mathematical Logic* (Cambridge, Mass.: Harvard University Press, 1967), pp. 5–82, at pp. 5–6.

[15] R. G. Heck, '*Grundgesetze der Arithmetik* I §§29–32', *Notre Dame Journal of Formal Logic* 38 (1998), pp. 437–74; G. Boolos and R. G. Heck, '*Die Grundlagen der Arithmetik* §§82–83', in G. Boolos, *Logic, Logic, and Logic* (Cambridge, Mass.: Harvard University Press, 1998), pp. 315–38.

to prove it that ultimately does not work, namely, the way outlined in §§82–3 of *Die Grundlagen.* As part of that proof, one has to prove a proposition[16] that, Frege remarks in a footnote, 'is, as it seems, unprovable ...'[17] It is notable that Frege does *not* say that this proposition is *false,* and there is good reason to think he regarded it as true and so true but unprovable in the Begriffsschrift: It follows immediately from the proposition Frege proves in its place, together with Dedekind's result that every infinite set is Dedekind infinite.[18] Frege knew of Dedekind's proof of this theorem and seems to have accepted it, although he complains in his review of Cantor's *Contributions to the Theory of the Transfinite* that Dedekind's proof 'is hardly executed with sufficient rigour'.[19] Frege apparently expended some effort trying to formalize Dedekind's proof. In the course of doing so, he could hardly have avoided discovering the point at which Dedekind relies upon an assumption apparently not available in the Begriffsschrift, namely, the axiom of (countable) choice. One can thus think of the theorem whose proof we have been unable to formalize either as Dedekind's result or as the unprovable proposition mentioned in section 114 of *Grundgesetze* and of NewAx as the Axiom of Choice.

Remarks of Dummett's suggest he would regard the foregoing as anachronisitc:

> No doubt Frege would have claimed his axioms, taken together with the additional informal stipulations not embodied in them,[20] as yielding a complete theory: to impute to him an awareness of the incompleteness of higher-order theories would be an anachronism.[21]

But I am suggesting only that Frege was prepared to consider the possibility that *his* formalization of logic (or arithmetic) was not complete: It is obvious that particular formalizations can be incomplete.

[16] The proposition in question is that labelled (1) in §82 of *Gl.*

[17] *Gg*, vol. I, §114. Translations are based upon the forthcoming translation due to Philip Ebert, Marcus Rossberg, and Crispin Wright.

[18] R. Dedekind, 'The nature and meaning of numbers', in *Essays on the Theory of Numbers*, trans. W. W. Beman (New York: Dover, 1963), §159

[19] Frege, 'Review of Georg Cantor, *Zum Lehre vom Transfiniten*', in *CP*, p. 180 (271).

[20] These are the stipulations made in §10, which we shall discuss below.

[21] M. Dummett, *The Interpretation of Frege's Philosophy* (London: Duckworth, 1981), p. 423.

What Gödel showed was that arithmetic (and therefore higher-order logic) is *essentially* incomplete, that every consistent formal theory extending arithmetic is incomplete. Of that Frege surely had no suspicion, but that is not relevant here.

In any event, the question whether a given (primitive) principle is a truth of logic is clearly one Frege regards as intelligible. And important. The question of the epistemological status of the basic laws of arithmetic is of central significance for Frege's project: His uncovering the fundamental principles of arithmetic will not decide arithmetic's epistemological status on its own. Though he did derive the axioms of arithmetic in the Begriffsschrift, that does not show that the basic laws of arithmetic are logical truths: That will follow only if the axioms of the Begriffsschrift are themselves logical laws and if its rules of inference are logically valid. The question of the epistemological status of arithmetic then reduces to that of the epistemological status of the axioms and rules of the Begriffsschrift – among other things, to the epistemological status of Frege's infamous Basic Law V, which states that functions $F\xi$ and $G\xi$ have the same 'value-range' if, and only if, they are co-extensional.

It is well known that, even before receiving Russell's letter informing him of the paradox, Frege was uncomfortable about Basic Law V. The passage usually quoted in this connection is this one:[22]

> A dispute can arise, so far as I can see, only with regard to my basic law (V) concerning value-ranges, which logicians perhaps have not yet expressly enunciated, and yet is what people have in mind, for example, where they speak of the extensions of concepts. I hold that it is a law of pure logic. In any event, the place is pointed out where the decision must be made.[23]

Although few commentators have said explicitly that Frege is here expressing doubt that Basic Law V is *true,* the view would nonetheless appear to be very widely held: It is probably expressed so rarely because it is thought that the point is too obvious to be worth

[22] Frege also writes: 'I have never disguised from myself [Basic Law V's] lack of the self-evidence that belongs to the other axioms and that must properly be demanded of a logical law' (*Gg*, vol. II, p. 253). The axiom's lacking self-evidence is reason to doubt it is a *logical* law: Self-evidence can be demanded only of primitive logical laws, not, say, of the axioms of geometry, which are evident on the basis of intuition. Frege does not suggest that he had any doubt about Basic Law V's *evidence.*

[23] *Gg*, vol. I, p. vii.

stating.[24] But we must be careful about reading our post-Russellian doubts about Basic Law V back into Frege: He thinks of Basic Law V as codifying something implicit, not only in the way logicians speak of the extensions of concepts, but in the way mathematicians speak of functions.[25] And there is, so far as I can see, no reason to conclude, on the basis of the extant texts, that Frege had any doubts about the Law's *truth*.

The nature of the dispute Frege expects, and 'the decision which must be made', is clarified by what precedes the passage just quoted:

Because there are no gaps in the chains of inference, every 'axiom'... upon which a proof is based is brought to light; and in this way we gain a basis upon which to judge the epistemological nature of the law that is proved. Of course the pronouncement is often made that arithmetic is merely a more highly developed logic; yet that remains disputable [*bestreitbar*] so long as transitions occur in proofs that are not made according to acknowledged laws of logic, but seem rather to be based upon something known by intuition. Only if these transitions are split up into logically simple steps can we be persuaded that the root of the matter is logic alone. I have drawn together everything that can facilitate a judgement as to whether the chains of inference are cohesive and the buttresses solid. If anyone should find anything defective, he must be able to state precisely where the error lies: in the Basic Laws, in the Definitions, in the Rules, or in the application of the Rules at a definite point. If we find everything in order, then we have accurate knowledge of the grounds upon which an individual theorem is based. A dispute [*Streit*] can arise, so far as I can see, only with regard to my basic law (V) concerning value-ranges ... I hold that it is a law

[24] An exception is Tyler Burge. Though Burge speaks, at one point, of 'Frege's struggle to justify Law (V) as a logical law', what he actually discusses are grounds Frege might have had for doubting its *truth* ('Frege on extensions of concepts from 1884 to 1903', *Philosophical Review*, 93 (1984, pp. 3–34, at pp. 30ff.). Burge claims that Frege's considering alternatives to Basic Law V suggests that he thought it might be false (pp. 12ff.). But given Frege's commitment to logicism, doubts about its epistemological status would also motivate such investigations.

[25] *Gg*, vol. II, §147. Treating concepts as functions then makes Basic Law V sufficient to yield extensions of concepts, too. And there is really nothing puzzling about this treatment of concepts: Technically, it amounts to identifying them with their characteristic functions. For more on this point, see R. Heck, 'The Julius Caesar objection', in R. Heck (ed.), *Language, Thought, and Logic: Essays in Honour of Michael Dummett* (Oxford: Oxford University Press, 1997), pp. 273–308.

of pure logic. In any event, the place is pointed out where the decision must be made.[26]

The dispute Frege envisions would concern the truth of Basic Law V were the correctness of the proofs all that was at issue here. But as I read this passage, Frege is attempting to explain how the long, complicated proofs he gives support his logicism,[27] how he intends to persuade us 'that the root of the matter is logic alone'. The three sentences beginning with 'I have drawn' constitute a self-contained explanation of how the formal presentation of the proofs gives us 'accurate knowledge of the grounds upon which an individual theorem is based', that is, how the proofs provide 'a basis upon which to judge the epistemological nature of' arithmetic, by reducing that problem to one about the epistemological status of the axioms and rules. Of course, someone might well object to Frege's proofs on the ground that Basic Law V is not *true*. But, although Frege must have been aware that this objection might be made, as said above, he thought the Axiom was widely, if implicitly, accepted. Moreover, as we shall see below, Frege took himself to have *proven* that Basic Law V is true in the intended interpretation of the Begriffsschrift.[28] But, in spite of all of this, Basic Law V was not an acknowledged law of logic. The 'dispute' Frege envisages thus concerns what other treatments have left 'disputable' – these words are cognate in Frege's German, too – namely, whether 'arithmetic is merely a more highly developed logic'. The objection Frege expects, and to which he has no adequate reply, is not that Basic Law V is not true, but that it is not 'a law of pure logic'. All he can do is to record his own conviction that it is and to remark that, at least, the question of arithmetic's *epistemological status* has been reduced to the question of Law V's *epistemological status*.

[26] *Gg*, vol. I, p. vii.

[27] This question is, in fact, taken up again in §66. It is unfortunate that this wonderful passage is so little known.

[28] I thus am not saying that Frege nowhere speaks to the question whether Basic Law V is true, even in *Grundgesetze* itself. What I am discussing here is where Frege thought matters stood *after* the arguments of *Grundgesetze* had been given. I am thus claiming that Frege thought he could answer the objection that Basic Law V is not true but would have had to acknowledge that he had no convincing response to the objection that it is not a law of logic.

The general question with which we are concerned here is thus what it is for an axiom of a given formal theory to be a logical truth, a logical axiom.[29] Admittedly, Frege does not say much about this question. One might think that that is because he had no view about the matter, that he had, as Warren Goldfarb has put it, no 'overarching view of the logical'.[30] Goldfarb's point, of course, is not just that Frege *did not* have any general account of what distinguishes logical from non-logical truth. Nor do I. His claim is that Frege's philosophical views precluded him from so much as envisaging, or attempting, such an account. But I find it hard to see how one can make that claim without committing oneself to the view that, for Frege, it is not even a substantive question whether Basic Law V is a truth of logic. Frege insists that Basic Law V is a truth of logic. Suppose that I were to deny that it is. Does Frege believe that this question is one that can be discussed and, hopefully, resolved in a rational manner? If not, then Frege's logicism is a merely verbal doctrine: It amounts to nothing more than a proposal that we should *call* Basic Law V a truth of logic. But if Frege thinks the epistemological status of Basic Law V is subject to rational debate, then any principles or claims to which he might be inclined appeal in attempting to resolve the question will constitute an inchoate (if incomplete) conception of the logical. I for one cannot believe that his considered views could commit him to the former position.

One thing that is clear is that the notion of generality plays a central role in Frege's thought about the nature of logic.[31] According to

[29] Similarly, Frege writes in *Grundlagen* that the question whether a proposition is analytic is to be decided by 'finding the proof of the proposition, and following it all the way back to the primitive truths', those truths 'which … neither need nor admit of proof'. The proposition is indeed analytic if, and only if, it can be derived, by means of logical inferences, from primitive truths that are 'general logical laws and definitions'. An analytic truth is thus a truth that follows from primitive logical axioms by means of logical inferences (*Gl*, §3). The problem is to say what primitive logical truths and logical means of inference are.

[30] Goldfarb expressed the point this way in a lecture based upon W. Goldfarb, 'Frege's conception of logic', in J. Floyd and S. Shieh (eds.), *Future Pasts: The Analytic Tradition in Twentieth-Century Philosophy* (New York: Oxford University Press, 2001), pp. 25–41.

[31] Naturally enough, since his discovery of quantification is so central to his conception of logic. See M. Dummett, *Frege: Philosophy of Language*, 2nd edn

Frege, logic is the most general science, in the sense that it is universally applicable. There might be special rules one must follow when reasoning about geometry, or physics, or history, which do not apply outside that limited area: But the truths of logic govern reasoning of all sorts. And if this is to be the case, it would seem that there must be another respect in which logic is general: As Thomas Ricketts puts the point, 'the basic laws of logic [must] generalize over every thing and every property [and] not mention this or that thing';[32] there can be nothing topic-specific about their content. Thus, the laws of logic are '[m]aximally general truths … that do not mention any particular thing or any particular property; they are truths whose statement does not require the use of vocabulary belonging to any special science'.[33]

So there is reason to think that Frege thought it necessary, if something is to be a logical law, that it should be maximally general in this sense. Some commentators, however, have flirted with the idea that Frege also held the condition to be sufficient.[34] Let us call this interpretation the *syntactic* interpretation of Frege's conception of logic. One difficulty with it is that such a characterization of the logical, even if extensionally correct, would not serve Frege's purposes. For consider any truth at all and existentially generalize on all non-logical terms occurring in it. The result will be a truth

(London: Duckworth, 1981), pp. 43ff.) for a discussion close in spirit to that to follow.

[32] T. Ricketts, 'Objectivity and objecthood: Frege's metaphysics of judgement', in L. Haaparanta and J. Hintikka (eds.), *Frege Synthesized: Essays on the Philosophical and Foundational Work of Gottlob Frege* (Dordrecht: Reidel, 1986), pp. 65–95, at p. 76.

[33] *Ibid*., p. 80. For similar views, see J. van Heijenoort 'Logic as calculus and logic as language', *Synthese,* 17 (1967), pp. 324–30; W. Goldfarb, 'Logic in the twenties', *Journal of Symbolic Logic,* 44 (1979), pp. 351–68; and B. Dreben and J. van Heijenoort, 'Introductory note to 1929, 1930, and 1930a', in S. Feferman *et al.* (eds), *Collected Works*, 3rd edn (New York: Oxford University Press, 1986), vol. I, pp. 44–59.

[34] Ricketts, for example, speaks of 'the identification of the laws of logic with maximally general truths' ('Objectivity and objecthood, p. 80). He goes on to quote Frege's remark that 'logic is the science of the most general laws of truth' (Frege, 'Logic', in *PW*, p. 128) and then glosses it as follows: 'To say that the laws of logic are the most general laws of truth is to say that they are the most general truths'. But whence the identification of the most general laws of truth with the most general truths?

that is, in the relevant sense, maximally general and so, on the syntactic interpretation, should be a logical truth. Thus, '$\exists x \exists y(x \neq y)$' should be a logical truth, since it is the result of existentially generalizing on all the non-logical terms in 'Caesar is not Brutus'. But the notion of a truth of logic plays a crucial epistemological role for Frege. In particular, logical truths are supposed to be analytic, in roughly Kant's sense: Our knowledge of them is not supposed to depend upon intuition or experience. Why should the mere fact that a truth is maximally general imply that it is analytic? Were there no way of knowing the truth of '$\exists x \exists y(x \neq y)$' except by deriving it from a sentence like 'Caesar is not Brutus', it certainly would not be analytic. More worryingly, consider '$\exists x \forall F(x \neq \grave{\varepsilon} F\varepsilon)$', which asserts that some object is not a value-range. This sentence is maximally general – if it is not, that is reason enough to deny that Basic Law V is a truth of logic – and, presumably, Frege regarded it as either true or false. But surely the question whether there are non-logical objects is not one in the province of logic itself.

Still, we need not be attempting to explain what it is for any truth at all to be a truth of logic, only what it is for *a primitive* truth,[35] an axiom, to be a truth of logic. So perhaps the condition should apply only to primitive truths: The view should be that a primitive truth is logical just in case it is maximally general. And it is eminently plausible that maximally general primitive truths must be analytic, for it is very hard to see how our knowledge of such a truth could depend upon intuition or experience. Intuition and experience deliver, in the first instance, truths that are *not* maximally general but that concern specific matters of fact. Hence, in so far as they support our knowledge of truths that are maximally general, they apparently must do so by means of inference. But then maximally general truths established on the basis of intuition or experience are not primitive.[36]

It might seem, therefore, that semantic concepts will play no role in Frege's conception of a truth of logic, that his conception is essentially syntactic. This, however, would be a hasty conclusion, for there are two respects in which the syntactic interpretation is

[35] See *Gl*, §3.

[36] Something like this line of thought is suggested by Ricketts, 'Objectivity and objecthood, p. 81.

incomplete, and these matter. First, our earlier statement of what maximally general truths are needs to be refined. Ricketts writes that '[m]aximally general truths … do not mention any particular thing or any particular property'. But reference to some specific concepts will be necessary for the expression of any truth at all, logical or otherwise. Frege himself remarks that 'logic … has its own concepts and relations; and it is only in virtue of this that it can have a content':[37] The universal quantifier refers to a specific second-level concept; the negation-sign, a particular first-level concept; the conditional, a first-level relation. And when Frege offers his 'emanation of the formal nature of logical laws' – an account not unlike a primitive version of the model-theoretic account of consequence, according to which logical laws are those whose truth does not depend upon what non-logical terms occur in them – the main problem he discusses is precisely that of deciding which notions are logical ones, whose interpretations must remain fixed: 'It is true that in an inference we can replace Charlemagne by Sahara, and the concept *king* by the concept *desert* … But one may not thus replace the relation of identity by the lying of a point in a plane.'[38]

The problem of the logical constant – the question which concepts belong to logic – is, for this reason, central to Frege's account of logic. His inability to resolve this problem may well have been one of the sources of his doubts about Basic Law V: Unlike the quantifiers and the propositional connectives, the smooth breathing – from which names of value-ranges are formed – is not obviously a logical constant. It is clear enough that what we now regard as logical constants have the generality of application Frege requires them to have: They appear in arguments within all fields of scientific inquiry, arguments that are, at least plausibly, universally governed by the laws of (the logical fragment of) the Begriffsschrift. It is far less clear that the smooth breathing – and the set-theoretic

[37] Frege, 'On the foundations of geometry: second series', in *CP*, pp. 293–340 (428).

[38] *Ibid.* It is hard to see that the question which concepts are logical is likely to admit of an answer in non-semantic terms. For some contemporary work, see G. Sher, *The Bounds of Logic: A Generalized Viewpoint* (Cambridge, Mass.: MIT Press, 1991). Sher's theory relies crucially on model-theoretic notions, such as preservation of truth-value under permutations of the domain. Dummett considers a similar proposal (*Frege: Philosophy of Language*, p. 22, fn.).

reasoning in which it would be employed – is similarly ubiquitous. It would therefore hardly have been absurd for one of Frege's contemporaries to insist that the smooth breathing and Basic Law V are peculiar to the 'special science' of *mathematics*.

The second problem with the syntactic interpretation is that it places a great deal of weight on the notion of primitiveness, and we have not been told how that is to be explained. Our modification of the syntactic interpretation – which consisted in claiming only that maximally general *primitive* truths are logical – will be vacuous unless there are restrictions upon what can be taken as a primitive truth. Otherwise, we could take '$\exists x \forall F(x \neq \grave{\varepsilon} F\varepsilon)$' as an axiom and its being a logical truth would again follow immediately. One might suppose that Frege's remarks on the nature of analyticity, mentioned above, committed him to some such notion. But it would be a mistake to think that Frege is committed to thinking that certain truths, of their very nature, admit of no proof: He is perfectly aware that, although some rules of inference and some truths must be taken as primitive, it may be a matter of choice which are taken as primitive. And since it is not obvious that there are any rules or truths that must be taken as primitive in every reasonable formalization, there need be none that are *essentially* primitive.[39] So, if the notion of primitiveness is to help at all here, we need an account of what makes a truth a candidate for being a primitive truth in some formalization or other. A natural thought would be that the notion of self-evidence should play some role,[40] but Frege says almost nothing directly about this question, either.[41]

One way to approach this issue would be via Frege's claim that logical laws are fundamental to thought and reasoning, in the sense that, should we deny them, we would 'reduce our thought

[39] Thus, Frege writes: '[I]t is really only relative to a particular system that one can speak of something as an axiom' (Frege, 'Logic in Mathematics', in *PW*, p. 206). See also *Bs*, §13, where Frege says, in effect, that he could have chosen other axioms for the theory and, indeed, that it might be essential to consider other axiomatizations if all relations between laws of thought are to be made clear.

[40] See *Gg*, vol. II, p. 253.

[41] There has been some recent work on this matter: see T. Burge 'Frege on knowing the foundation', *Mind*, 107 (1998), pp. 305–47; and R. Jeshion, 'Frege's notion of self-evidence', *Mind*, 110 (2001), pp, 937–76.

to confusion'.[42] I have no interpretation to offer of this claim. But I want to emphasize that it is not enough for Frege simply to *assert* that his axioms cannot coherently be denied. What Frege would have needed is an account of why the particular statements he thought were laws of logic have this privileged status. The semantic concepts Frege uses in stating the intended interpretation of the *Begriffsschrift*, which I shall discuss momentarily, also pervade his mature work on the philosophy of logic: And it is a nice question why Frege should have turned to the study of semantic notions at just this time. My hunch, and it is just a hunch, is that he did so because he was struggling with the very questions about the nature of logic we have been discussing and that he was developing a conception of logic in which they would play a fundamental role. Frege argues, in the famous papers written around the time he was writing *Grundgesetze*, that semantic concepts are central to any adequate account of our understanding of language, of our capacity to express thoughts by means of sentences, to make judgements and assertions, and so forth.[43] So, if Frege could have shown that negation, the conditional, and the quantifier were explicable in terms of these semantic concepts – and he might well have thought that the semantic theory for the *Begriffsschrift* shows just this – he could then have argued that they are, in principle, available to anyone able to think and reason, that is, that these notions (and the fundamental truths about them) are, in that sense, implicit in our capacity for thought. Unfortunately, such an argument would not apply to Basic Law V: The notion of a value-range does not seem to be fundamental to thought in this way, and, as we shall see, the semantic theory does not treat it the same way it treats the other primitives. So that might have provided a second reason for Frege to worry about its epistemological status. But I shall leave the matter here, for we are already well beyond anything Frege ever discussed explicitly.

[42] Frege, *Gg*, vol. I, p. vii; see also *Gl*, §14.

[43] See, for example, Frege, 'On sense and meaning', in *CP*, pp. 157–77 (34), where Frege argues that the truth-values 'are recognized, if only implicitly, by everybody who judges something to be true'. See also Frege's flirtation with a transcendental argument for the laws of logic (*Gg*, vol. I, p. xvii).

2 FORMALISM AND THE SIGNIFICANCE OF INTERPRETATION

The discussion in the preceding section began with the question what it might mean to justify the laws of logic. I argued that justifications of logical laws intended to establish their truth must be circular. But the argument for that claim depended upon an assumption that I did not make explicit, namely, that the logical laws whose truth is in question are *the thoughts expressed by* certain sentences. It is quite possible to argue, without circularity, that certain *sentences* that in fact express, or are instances of, laws of logic are true, say, to argue that every instance of '$A \vee \neg A$' is true. I do just that in my introductory logic classes. Of course, the arguments carry conviction only because my students are willing to accept certain claims that I state in English using sentences that are themselves instances of excluded middle. But that discloses no circularity: My purpose is just to convince them of the truth of all sentences of a certain form.

Semantic theories frequently have just this kind of purpose. A formal system is specified: a language is defined, certain sentences are stipulated as axioms, and rules governing the construction of proofs are laid down. The language is then given an interpretation: The references of primitive expressions of the language are specified, and rules are stated that determine the references of compound expressions from those of their parts. It is then argued – completely without circularity – that all of the sentences taken as axioms are true and that the rules of inference are truth-preserving. Of course, the argument carries conviction only because we are willing to accept certain claims stated in the metalanguage, the language in which the interpretation is given. But that discloses no circularity: The purpose of the argument is to demonstrate the truth of the sentences taken as axioms and the truth-preserving character of the rules. Its purpose is to show not that the thoughts expressed by certain formal sentences are true but only that those sentences are true.

The semantic theory Frege develops in Part I of *Grundgesetze* has the same purpose. In the case of each of the primitive expressions of Begriffsschrift, he states what its interpretation (that is, its reference) is to be. Thus, for example:

'$\Gamma = \Delta$' shall denote the True if Γ is the same as Δ; in all other cases it shall denote the False.[44]

'$\Phi(a)$' is to denote the True if, for every argument, the value of the function $\Phi(\xi)$ is the True, and otherwise it is to denote the False.[45]

Some of Frege's stipulations – which I shall call his *semantic stipulations* regarding the primitive expressions – do not take such an explicitly semantic form. Thus, for example, in connection with the horizontal, Frege writes:

I regard it as a function-name, as follows:

$- \Delta$

is the True if Δ is the True; on the other hand, it is the False if Δ is not the True.[46]

Frege wanders back and forth between the explicitly semantic stipulations and ones like this: But the point, in each case, is to say what the reference of the expression is supposed to be, and Frege argues in § 31 of *Grundgesetze* that these stipulations do secure a reference for the primitives and, in § 30, that the stipulations suffice to assign references to all complex expressions, too, if they assign references to all the primitive expressions.[47]

Frege goes on to argue that each axiom of the Begriffsschrift is true. Thus, about Axiom I he writes:

By [the explanation of the conditional given in] § 12,

$\Gamma \rightarrow (\Delta \rightarrow \Gamma)$

could be the False only if both Γ and Δ were the True while Γ was not the True. This is impossible; therefore

$\vdash \Gamma \rightarrow (\Delta \rightarrow \Gamma)$[48]

And, similarly, in the case of each of the rules of inference, he argues that it is truth-preserving. Thus, regarding transitivity for the conditional, he writes:

[44] *Gg*, vol. I, §7.

[45] *Ibid.*, §8.

[46] *Ibid.*, §5.

[47] For discussion of these arguments, see Heck, '*Grundgesetze der Arithmetik* I §§29–32'; R. Heck, '*Grundgesetze der Arithmetik* I, §10', *Philosophia Mathematica*, 7 (1999), pp. 258–92; and Ø. Linnebo, 'Frege's proof of referentiality', *Notre Dame Journal of Formal Logic*, 45 (2004), pp. 73–98.

[48] *Gg*, vol. I, §18.

From the two propositions

$\vdash \Delta \rightarrow \Gamma$

$\vdash \Theta \rightarrow \Delta$

we may infer the proposition

$\vdash \Theta \rightarrow \Gamma$

For $\Theta \rightarrow \Gamma$ is the False only if Θ is the True and Γ is not the True. But if Θ is the True, then Δ too must be the True, for otherwise $\Theta \rightarrow \Delta$ would be the False. But if Δ is the True then if Γ were not the True then $\Delta \rightarrow \Gamma$ would be the False. Hence the case in which $\Theta \rightarrow \Gamma$ is not the True cannot arise; and $\Theta \rightarrow \Gamma$ is the True.[49]

These arguments – which, for the moment, I shall call *elucidatory demonstrations* – tend by and large not to be explicitly semantic: That is, Frege usually speaks not of what the premises and conclusion denote, but rather of particular objects *being* the True or the False. One might suppose that this shows that Frege's arguments should not be taken to be semantic in any sense at all. But, to my mind, the observation is of little significance: What it means is just that Frege is not being as careful about use and mention as he ought to be.

It is sometimes said that Begriffsschrift is not an 'interpreted language', in the sense of a syntactic object – a language, in the technical sense – that has been given an interpretation. Rather, it is a 'meaningful formalism', something like a language in the ordinary sense, but one that just happens to be written in funny symbols – something in connection with which it would be more appropriate to speak, as Ricketts does, of 'foreign language instruction' than interpretation.[50] If so, then one might suppose that Frege could not have been interested in 'interpretations' of Begriffsschrift because, in his exchanges with Hilbert, he seems to be opposed to any consideration of varying interpretations of meaningful languages. But, as Jamie Tappenden has pointed out, Frege's own mathematical work involved the provision of just such reinterpretations of, for example, complex number theory. What Frege objected to was Hilbert's claim that content can be bestowed upon a sign *simply* by indicating a range of alternative interpretations.[51] In some sense, it seems to me,

[49] *Ibid.*, § 15.

[50] Ricketts, 'Generality, sense and meaning', p. 176.

[51] J. Tappenden, 'Geometry and generality in Frege', *Synthese,* 102 (1995), pp. 319–61. For further consideration of this kind of question, see J. Tappenden, 'Frege on

Frege thought that the concept of an interpreted language was more basic than that of an uninterpreted one – and it is hard not to be sympathetic. But it simply does not follow that one cannot intelligibly consider other interpretations of the dis-interpreted symbols of a given language.

In any event, Frege was certainly aware that it would be possible to treat Begriffsschrift as an uninterpreted language, with nothing but rules specifying how one sentence may be constructed from others. For the central tenet of Formalism, as Frege understood the position, is precisely that arithmetic ought to be developed as a Formal theory,[52] in the sense that the symbols that occur in it have no meaning (or that their meaning is somehow irrelevant). Such a theory need not be lacking in mathematical interest: It can, in particular, function as an object of mathematical investigation. There could, for example, be a mathematical theory that would prove such things as that this 'figure' (formula) can be 'constructed' (derived) from others using certain rules – or that a given figure cannot be so constructed.[53] One can, if one likes, stipulate that certain figures are 'axioms', which specification one might compare to the stipulation of the initial position in chess, and take special interest in the question what figures can be derived from the 'axioms'.[54] Frege's fundamental objection to Formalism is that it cannot explain the

axioms, indirect proof, and independence arguments in geometry: Did Frege reject independence arguments?', *Notre Dame Journal of Formal Logic*, 41 (2000), pp. 271–315. And even if we were to accept this objection, it still would not follow that Frege was uninterested in semantics. See J. Stanley, 'Truth and metatheory in Frege', *Pacific Philosophical Quarterly*, 77 (1996), pp. 45–70, at p. 64.

[52] For a discussion of this notion of a formal theory, see Frege, 'Formal theories of arithmetic', in *CP*, pp. 112–21 I shall capitalize the word 'Formal' when I am using it in the sense explained here.

[53] *Gg*, vol. II, §93.

[54] *Ibid.*, §§90–1. Frege's discussion explicitly concerns the rules of arithmetic, not those of logic: but, of course, for Frege, arithmetic is logic, and his formal system of arithmetic, the Begriffsschrift, contains no axioms or rules that are (intended to be) non-logical. His discussion of what requirements the rules of arithmetic must meet therefore applies directly to the axioms and rules of inference of the Begriffsschrift itself. Thus, he writes: 'Now it is quite true that we could have introduced our rules of inference and the other laws of the Begriffsschrift as arbitrary stipulations, without speaking of the reference and the sense of the signs. We would then have been treating the signs as figures' (*Gg*, vol. II, §90). That is to say, we should then have been adopting a Formalist perspective on the Begriffsschrift.

applicability of arithmetic, and this needs to be explained, for 'it is applicability alone which elevates arithmetic from a game to the rank of a science'.[55] An examination of Frege's development of this objection will thus reveal what he thought would have been lacking had Begriffsschrift been left uninterpreted – and so what purpose he intended his semantic stipulations to serve.

Frege distinguishes 'Formal' from 'Significant'[56] arithmetic. He characterizes Significant arithmetic as the sort of arithmetic that concerns itself with the references of arithmetical signs, as well as with the signs themselves and with rules for their manipulation. Formal arithmetic is interested only in the signs and the rules: It treats Begriffsschrift as an uninterpreted language. On the Formalist view, the references of, say, numerals are of no importance to arithmetic itself, though they may be of significance for the application of arithmetic.[57] And, according to Frege, this refusal to recognize the references of numerical terms is what is behind another of the central tenets of Formalism, that the rules[58] of a system of arithmetic are, from the point of view of arithmetic proper, entirely arbitrary: 'In Formal arithmetic we need no basis for the rules of the game – we simply stipulate them.'[59] Though Formalists recognize that the rules of arithmetic cannot really be arbitrary, they take this fact to be of significance, not for arithmetic, but only for its applications:

Thomae ... contrasts the arbitrary rules of chess with the rules of arithmetic ... But this contrast first arises when the applications of arithmetic are in question. If we stay within its boundaries, its rules appear as arbitrary as those of chess. This applicability cannot be an accident – but in

[55] *Gg*, vol. II, §91.

[56] The German term is '*inhaltlich*', which Geach and Black translate in the first edition of Translations as 'meaningful'. While this was a reasonable translation then, it is now dangerous, since the cognate term 'meaning' has become a common translation of Frege's term '*Bedeutung*'. In the third edition, they translate '*inhaltlich Arithmetik*' as 'arithmetic with content'; a literal translation would be 'contentful arithmetic'. Both of these sound cumbersome to my ear.

[57] *Gg*, vol. II, §88.

[58] Frege speaks, throughout these passages, of the 'rules' of the Formal game, thereby meaning to include, I think, not just its 'rules of inference', but also its 'axioms' – though he does tend to focus more on the 'rules permitting transformations' than on the stipulation of the initial position or 'starting points' (*ibid.*, §90). The reason is that he tends to think even of the axioms of a Formal theory as rules saying, in effect, that certain things can always be written down. See here *ibid.*, §109.

[59] *Ibid.*, §89.

Formal arithmetic we absolve ourselves from accounting for one choice of the rules rather than another.[60]

It is important to remember that, throughout this discussion, Frege is *contrasting* Formal and Significant arithmetic. When he speaks of 'absolv[ing] ourselves from accounting for one choice of the rules rather than another', he is not just saying that the rules of arithmetic are non-arbitrary; he is implying that, if we are to formulate a system of Significant arithmetic, we must ourselves 'account ... for one choice of the rules rather than another'.

Frege does not think of this account as a mere appendage to Significant arithmetic, but as a crucial part of the work of the arithmetician:

It is likely that the problem of the usefulness of arithmetic is to be solved – in part, at least – independently of those sciences to which it is to be applied. Therefore it is reasonable to ask the arithmetician to undertake the task ... This much, it appears to me, can be demanded of arithmetic. Otherwise it might happen that, while [arithmetic] handled its formulas simply as groups of figures without sense, a physicist wishing to apply them might assume quite without justification that they expressed thoughts whose truth had been demonstrated. This would be – at best – to create the illusion of knowledge. The gulf between arithmetical formulas and their applications would not be bridged. In order to bridge it, it is necessary that the formulas express a sense and that the rules be grounded in the reference of the signs.[61]

The rules must be so grounded because arithmetic is expected to deliver truths – not just truths, in fact, but knowledge. As Frege concludes the passage: 'The end must be knowledge, and it must determine everything that happens'.[62]

On the Formalist view, the numerals and other signs of a system of arithmetic can have no reference, as far as arithmetic itself is concerned: 'If their reference were considered, the ground for the rules would be found in these same references.'[63] What is most important, for present purposes, is Frege's conception of how the references of the expressions ground the rules:

[60] *Ibid.*, §89.
[61] *Ibid.*, §92.
[62] *Ibid.*, §92.
[63] *Ibid.*, §90.

The question, 'What is to be demanded of numbers in arithmetic?' is, says Thomae, to be answered as follows: In arithmetic we require of numbers only their signs, which, however, are not treated as being signs of numbers, but solely as figures; and rules are needed to manipulate these figures. We do not take these rules from the reference of the signs, but lay them down on our own authority, retaining full freedom and acknowledging no necessity to justify the rules.[64]

Thus, not only do the references of the signs ground the rules that govern them, but, unless we are Formalists, we must recognize an obligation to justify these rules, presumably by showing that they are grounded in the references of the signs. Frege elsewhere specifies what condition rules of inference, in particular, must be shown to satisfy:

Whereas in Significant arithmetic equations and inequations are sentences expressing thoughts, in Formal arithmetic they are comparable with the positions of chess pieces, transformed in accordance with certain rules without consideration for any sense. For if they were viewed as having a sense, the rules could not be arbitrarily stipulated; they would have to be so chosen that, from formulas expressing true thoughts, only formulas likewise expressing true thoughts could be derived.[65]

Thus, the rules of inference in a system of Significant arithmetic must be truth-preserving. And this condition – that the rules should be truth-preserving – is not arbitrarily stipulated, either. It follows from arithmetic's ambition to contribute to the growth of knowledge:

If in a sentence of Significant arithmetic the group '3 + 5' occurs, we may substitute the sign '8' without changing the truth-value, since both signs designate the same object, the same actual number, and therefore everything which is true of the object designated by '3 + 5' must be true of the object designated by '8' … It is therefore the goal of knowledge that determines the rule that the group '3 + 5' may be replaced by the sign '8'. This goal requires the character of the rules to be such that, if in accordance with them a sentence is derived from true sentences, the new sentence will also be true.[66]

[64] *Ibid.*, §94.
[65] *Ibid.*, §94.
[66] *Ibid.*, §104.

Derivation must preserve truth, for only if it does, and only if the axioms are themselves true, will the theorems of the system be guaranteed to express true thoughts` it is only because the thoughts expressed by these formulas are true … and, indeed, are known to be true … that their application contributes to the growth of knowledge, rather than producing a mere •illusion of knowledge•. [XJ]

Since Frege is interested in developing a system of Signi cant arithmetic, he in particular owes some account of why the rules of the *Begriffsschrift* are non-arbitrary, that is, a demonstration that they are truth-preserving (and a similar demonstration that its axioms are true). Unless Frege | agrantly failed to do just what he is criticizing the Formalists for failing to do, he must somewhere have provided such an account. There is no option but to suppose that he does so in part I of *Grundgesetze* and that the elucidatory demonstrations in particular are intended to show that the rules of the system are truth-preserving and that the axioms are true. Indeed, since Frege himself speaks of a need to justify the rules and of their being grounded in the references of the signs, we may dispense with our euphemism and speak, not of elucidatory demonstrations, but of Frege•s semantic justi cations of the axioms and rules.

Z : • ;"

I have argued that Frege•s semantic justi cations of the axioms and rules of his system are intended to establish that, under the intended interpretation of the *Begriffsschrift* … this being given by the semantic stipulations governing the primitive expressions … its axioms are true and its rules are truth-preserving. But, according to Ricketts, they cannot have been intended to serve this purpose, because Frege•s •conception of judgment precludes any serious metalogical perspective• from which he could attempt to justify his axioms and rules. [X] His philosophical views •preclude ineliminable uses of a truth-predicate,

[XJ] See *ibid*., §§ , . Note that Frege is arguing here not only that the rules are required to be truth-preserving if arithmetic is to deliver knowledge but, conversely, that the substitution of terms having the same reference is permissible because the goal of arithmetic is knowledge. Substitution of coreferential terms … indeed, even of terms with the same sense … is not permitted everywhere? It is not permitted in poetry or in comedy, for example.

[X] Ricketts, •Objectivity and objecthood•, p. JX. Van Heijenoort goes so far as to claim that Frege•s •rules are void of any intuitive logic• (van Heijenoort, •Logic as including uses in bona fide generalizations', such as would be necessary were one even to be able to *say* that a rule of inference is valid. Ricketts is not, of course, unaware of what goes on in part I of *Grundgesetze,* but he claims that Frege's sole purpose in part I is to[69]

> teach his audience Begriffsschrift. Frege's stipulations, examples, and commentary function like foreign language instruction to put his readers in a position to know what would be affirmed by the assertion of any Begriffsschrift formula. The understanding produced by Frege's elucidatory remarks should have two immediate upshots. First, it should lead to the affirmation of the formulas Frege propounds as axioms; second, it should prompt the appreciation of the validity of the inference rules Frege sets forth.[70]

Frege's elucidations thus enable his reader to know what is expressed by any Begriffsschrift formula; so knowing, the reader can determine whether the formulae expressing the axioms are true by asking herself whether she is prepared to assert what they express. She may be aided by Frege's examples, commentary and so forth, but this heuristic purpose is the only purpose they serve: The semantic justifications are not demonstrations of the truth of the axioms, nor of the validity of the rules, but are meant to *persuade*.

But it is unclear why, if Frege's only purpose were to teach his audience Begriffsschrift, he should make use of such notions as that of an object, or of a truth-value or of reference, and why

calculus', p. 326). But Frege simply spends too much time explaining the intuitive basis for his rules for this claim to be plausible; and, if that weren't enough, if correct, it would make Frege a formalist. The following passage is often cited as expressing Frege's opposition to meta-perspectives:

> We have already introduced a number of fundamental principles of thought in the first chapter in order to transform them into rules for the use of our signs. These rules and the laws whose transforms they are cannot be expressed in the Begriffsschrift, because they form its basis. (*Bs*, §13)

But it would be absurd for Frege to suggest that the axioms cannot be expressed in Begriffsschrift. He is speaking here simply of rules, in particular, of rules of inference, and noting that they cannot be so expressed: Frege has, in the first chapter, only introduced the system's rules. He goes on to explain that he is out, in the second, to find axioms from which all 'judgements of pure thought' will follow by means of those rules. Frege is thus making the distinction between rules and axioms here, not expressing opposition to meta-perspectives.

[69] I have capitalized 'Begriffsschrift' in both occurrences. I do not intend to consider here Ricketts's reasons for ascribing this view to Frege. For an extended discussion of his interpretation, see Stanley, 'Truth and metatheory'.

[70] Ricketts, 'Generality, sense, and meaning', pp. 176–7.

his 'explanations' should be, in the usual sense, compositional. It would do as well (and be far simpler) just to explain how to *translate* a proposition of Begriffsschrift into English (or German).[71] But Frege does not say simply that '$\Gamma = \Delta$' expresses the thought that Γ is the same as Δ: He says that it 'shall denote the True if Γ is the same as Δ [and] in all other cases … shall denote the False'.[72] One might reply that natural languages do not perspicuously express what Frege wishes to express in the Begriffsschrift. But while this is fine so far as it goes, it suggests merely that some technical vocabulary might be needed to 'teach Begriffsschrift'. It does not explain why that vocabulary should be semantic.

Moreover, Frege's semantic justifications become a great deal more complicated than those cited so far, particularly in cases in which free variables – which he calls Roman letters – occur in the premises and conclusion of an inference.[73] But this has been obscured by an almost universal misunderstanding of Frege's use of Roman letters. I just said that they are free variables, but it is widely held that there really aren't any free variables in Begriffsschrift: that Roman letters are tacitly bound by invisible, initial universal quantifiers. Frege does say that the scope of a Roman letter 'shall comprise everything that occurs in the proposition',[74] which amounts to his stipulating that a formula containing free variables is true just in case its universal closure is true. But he rejects the interpretation of Roman letters as tacitly bound almost immediately thereafter:[75]

Our stipulation regarding the *scope* of a *Roman letter* is to set only a lower bound upon the scope, not an upper bound. Thus it remains permissible to extend such a scope over several propositions, and this renders the Roman

[71] The contrast between a semantic theory and a translation manual is, of course, emphasized in D. Davidson, 'Radical interpretation', in *Inquiries into Truth and Interpretation* (Oxford: Clarendon Press, 1984), pp. 125–139, at pp. 129–30.

[72] *Gg*, vol. I, §7.

[73] The interpretive claims made in the remainder of this section and the next are developed in more detail, and defended, in Heck, '*Grundgesetze der Arithmetik* I §§29–32'. That paper limits itself to discussion of the technical details of Frege's arguments in §§29–32 and does not, as the present paper does, discuss the bearing of my interpretation on questions about Frege's conception of logic. This paper and that one are, therefore, companion pieces, to some extent, although the discussion here is independent of the messy details encountered there. A more unified discussion will appear in a book on Frege now in preparation.

[74] *Gg*, vol. I, §17.

[75] I am silently converting Frege's notation to ours.

letters suitable to do duty in inferences, which the Gothic letters, with the strict closure of their scopes, cannot. If we have the premises '$\vdash x^2 = 1 \rightarrow x^4 = 1$' and '$\vdash x^4 = 1 \rightarrow x^8 = 1$' and infer the proposition '$\vdash x^2 = 1 \rightarrow x^8 = 1$', in making the transition we extend the scope of the 'x' over both of the premises and the conclusion, in order to perform the inference, although each of these propositions still holds good apart from this extension.[76]

There is, for Frege, an important difference between a proposition of the form '$\vdash \Phi(x)$' and its universal closure '$\vdash \forall \Phi(x)$'.[77] The nature of this difference, however, is puzzling: what could Frege mean by saying that, in making certain inferences, we must 'extend the scope of the "x" over both of the premises and the conclusion'? Surely he cannot mean that something like

$$\forall x\{\vdash (x^2 = 1 \rightarrow x^4 = 1) \wedge \vdash (x^4 = 1 \rightarrow x^8 = 1) \rightarrow \\ \vdash (x^2 = 1 \rightarrow x^8 = 1)\}$$

is supposed to be well-formed!

Frege is concerned here with what licenses us to make the inference under discussion. There is a rule in his system, rule (7), that permits it.[78] That rule – transitivity for the conditional – allows the inference from '$\vdash \Delta \rightarrow \Gamma$' and '$\vdash \Theta \rightarrow \Delta$' to '$\vdash \Theta \rightarrow \Gamma$'. But if Roman letters were treated as tacitly bound, rule (7) would not apply: rule (7) does not allow an inference from '$\vdash \forall x(x^2 = 1 \rightarrow x^4 = 1)$' and '$\vdash \forall x(x^4 = 1 \rightarrow x^8 = 1)$' to '$\vdash \forall x(x^2 = 1 \rightarrow x^8 = 1)$'. The point is not that this formal rule could not be made to apply: It can, if we introduce a notation in which initial universal quantifiers can be suppressed; some formal systems treat free variables in just that way. Nor is there any substantive worry about whether the inference is in fact valid. Rather, the problem is that we are at present without

[76] *Gg*, vol. I, §17.

[77] Compare this remark: 'Now when the scope of the generality is to extend over the whole of a sentence closed off by the judgement stroke, then as a rule I employ Latin letters ... But if generality is to extend over only part of the sentence, then I adopt German letters ... Instead of the German letters, I could have chosen Latin ones here, just as Mr Peano does. But from the point of view of inference, generality which extends over the content of the entire sentence is of vitally different significance from that whose scope constitutes only a part of the sentence. Hence it contributes substantially to perspicuity that the eye discerns these different roles in the different sorts of letters, Latin and German' (Frege, 'On Mr. Peano's conceptual notation and my own', p. 378; I have altered the translation slightly).

[78] The rules of the system are listed in *Gg*, vol. I, §48.

any argument that inferences of this form *are* valid when the premises and conclusion contain free variables:[79] The semantic justification of rule (7) given in *Grundgesetze,* § 15 (and quoted earlier), did not allow for the possibility that 'Γ', 'Δ' and 'Θ' might contain free variables. That justification, which is essentially a justification in terms of truth-tables, presupposes that 'Γ', 'Δ' and 'Θ' *have truth-values* and, moreover, that the truth-values they have, when they occur in one premise, are the same as those they have when they occur in the other or in the conclusion. Only if we may speak of the truth-value of the occurrence of '$x^2 = 1$' in the first premise, and only if it has the same truth-value in all of its occurrences, will the justification apply. And we cannot so speak.

Nowadays, what we would say is that the inference is valid because, whenever we make a simultaneous assignment of objects to free variables in the premises and the conclusion, the usual argument on behalf of transitivity – the argument in terms of truth-tables – still goes through, if we replace occurrences of 'true' with occurrences of 'true under that assignment': That is to say, that argument can be adapted to show that, if the premises are both true under the *same* assignment, the conclusion must also be true under that same assignment. When Frege says that the scope of 'x' is to be 'extend[ed] … over several propositions', he is attempting to express the relevant notion of simultaneous assignment: The idea is that, as we perform the inference, we treat the variable as (in Frege's terminology) indicating the same object in every one of its occurrences, whether in one of the premises or in the conclusion.

What Frege has said to this point speaks only to the notion of simultaneity and not to the notion of an assignment itself. But what follows the passage just discussed are further remarks on the nature of free variables and inferences involving them, including rule (5) of the Begriffsschrift, the rule of universal generalization:

A Roman letter may be replaced at all of its occurrences in a proposition by one and the same Gothic letter … The Gothic letter must then at the same

[79] It is worth emphasizing that free-variable reasoning is distinctive of Frege's new logic (polyadic quantification theory). There is no need for such reasoning in syllogistic logic (which is not to deny that monadic quantification theory can be formulated as a sub-theory of polyadic).

time be inserted over a concavity in front of a main component outside which the Gothic letter does not occur.[80]

Decoding Frege's terminology, what the rule says is that one can infer '$A \rightarrow \forall x B(x)$' from '$A \rightarrow B(x)$', if '$x$' is not free in A.[81] Frege's semantic justification of this rule is contained in *Grundgestze* I, § 17, and is in three stages. First, he notes that '$\vdash \Gamma \rightarrow \Phi(x)$' is equivalent to '$\vdash \forall x[\Gamma \rightarrow \Phi(x)]$', since a formula containing a Roman letter is true just in case its universal closure is true. Secondly, he argues that, if 'x' is not free in Γ *and no other variables are free in either* Γ *or* $\Phi(x)$, then '$\vdash \forall x[\Gamma \rightarrow \Phi(x)]$' is equivalent to '$\vdash \Gamma \rightarrow \forall x \Phi(x)$': That is, he shows, by means of what is now a familiar argument, that '$\forall x(p \rightarrow Fx)$' is equivalent to '$p \rightarrow \forall x Fx$'. The final stage of the argument is contained in the following passage:

> If for 'Γ' and '$\Phi(x)$', combinations of signs are substituted that do not refer to an object and a function respectively, but only indicate, because they contain Roman letters, then the foregoing still holds generally if for each Roman letter a name is substituted, whatever this may be.[82]

It is important to see how odd this final stage of the argument is. What Frege wants to show is that, if 'x' is not free in A, then '$\forall x(A \rightarrow B(x)$' is equivalent to '$A \rightarrow \forall x B(x)$'. But what he says is that, if we substitute names for all free variables, other than 'x', in A and $B(x)$, the argument that establishes that '$\forall x(p \rightarrow Fx)$' is equivalent to '$p \rightarrow \forall x Fx$' will go through.

It is not immediately obvious why that should suffice. What we would say nowadays is that, if we make a simultaneous assignment to the free variables other than 'x' in A and $B(x)$, that same argument will go through, 'true' again being replaced by 'true under the assignment'. The only difference between this argument and Frege's is that, where we speak of assignments, he speaks of substitutions. Frege does not, however, mean to speak here of

[80] *Gg*, vol. I, §48.

[81] A Roman letter is a free variable; a Gothic letter, a bound one; and the concavity, the universal quantifier. To say that the quantifier must appear 'in front of a main component outside which the Gothic letter does not occur' is to say that it need not contain the antecedent of the conditional in its scope if the Roman letter in question does not occur in the antecedent.

[82] *Gg*, vol. I, §17.

substitutions of *actual terms* of Begriffsschrift for the variable,[83] but of *auxiliary names* assumed only to denote some object in the domain. What Frege is assuming, in the argument at which we have just looked, is that the inference from '$\phi(x)$' to '$\psi(x)$' will be valid just in case ⌜$\psi(\Delta)$⌝ is true whenever ⌜$\varphi(\Delta)$⌝ is true, Δ being a name new to the language and subject only to the condition that it must denote a member of the domain. This idea can be made precise: applied to quantification, it constitutes a perfectly coherent alternative to Tarski's treatment in terms of satisfaction.[84] It is a mark of the depth of Frege's understanding of logic that he realized that the presence of free variables in the language means that even the validity of rules of inference belonging to its *propositional* fragment – rules like *modus ponens* and transitivity for the conditional – cannot be justified simply in terms of the truth-tables. It is all the more remarkable that, in thinking about this problem, he was led to produce this alternative to Tarski's treatment of the quantifiers. And I, for one, find it hard to believe that the arguments at which we have just looked are but part of an attempt to 'teach Begriffsschrift'. The argument Frege gives in favour of the validity of universal generalization is surely not intended merely to encourage the reader not to object to the applications he makes of it. If that were all he wanted, he could have had it far more easily.

4 *GRUNDGESETZE DER ARITHMETIK* I, §§30–1

Matters become yet more complicated with Basic Law V.[85] The semantic stipulation governing the smooth breathing is not like the stipulations Frege gives for the other primitives: He does not

[83] If he were so to speak, the argument would show only (to put the point in Tarskian language) that the conclusion is true whenever the premise is, when objects *denoted by terms in the language* are assigned to the free variables. Compare Dummett, *Frege: Philosophy of Language*, p. 17). For detailed discussion of how Frege's argument leads to the conclusion that the inference is valid, see Heck, '*Grundgesetze der Arithmetik* I §§29–32'.

[84] See *ibid.*, Appendix, for a sketch of such a theory and for references. A similar treatment of quantification is given in Benson Mates's textbook *Elementary Logic*, 2nd edn (Oxford: Oxford University Press, 1972).

[85] The discussion in this section summarizes some of the results of Heck, '*Grundgesetze der Arithmetik* I §§29–32', and Heck, '*Grundgesetze der Arithmetik* I §10', which should be consulted for defences of claims that are not defended here.

directly stipulate what its reference is to be. Of course, it would not have been difficult for him to do so: He need only have said that a term of the form $\grave{\varepsilon}\Phi(\varepsilon)$ denotes the value-range of $\Phi(\xi)$;[86] he could then have argued that, since the value-range of $\Phi(\xi)$ is the same as the value-range of $\Psi(\xi)$ just in case the same objects fall under $\Phi(\xi)$ and $\Psi(\xi)$, Basic Law V holds. Now, in fact, Frege does consider such a stipulation at one point,[87] but all we are told about value-ranges is that the value-range of the function $\Phi(\xi)$ is the same as that of $\Psi(\xi)$ just in case they have the same values for the same arguments.[88] In effect, then, the only stipulation Frege makes about the smooth breathing, and the one he uses in his arguments, is that $\grave{\varepsilon}\Phi(\varepsilon) = \grave{\varepsilon}\Psi(\varepsilon)$ has the same truth-value as $\forall x(\Phi(x) = \Psi(x))$. Frege notes, in *Grundgesetze* I, §20, that the truth of Basic Law V follows immediately from this stipulation (or from the combined effect of those made in *Grundgesetze* I, §3, 9). But it will do so only if the stipulation is in good order, only if it suffices to assign a reference to the smooth breathing.

But the stipulation does *not* directly assign a reference to the smooth breathing. And unless it somehow succeeds in doing so indirectly, as it were, Basic Law V cannot be justified in terms of it: Officially, the axiom ought then to be declared neither true nor false, on the ground that it contains an expression that has no reference. Frege therefore needs to argue that his stipulation, augmented by others to be mentioned shortly, does indeed secure a reference for the smooth breathing; his argument comprises most of §31 of *Grundgesetze.* Had it been successful, Frege would have *proven* that Basic Law V is true in the intended interpretation of the system. That is why I said earlier that Frege could have had no real doubts about the *truth* of Basic Law V.[89]

The question whether the smooth breathing has been assigned a reference is made pressing by the peculiar nature of the semantic stipulation governing it. But Frege still argues that a reference

[86] I'll write quotation marks and corner quotes with invisible ink in this section, to avoid cluttering the exposition.

[87] *Gg*, vol. I, §9.

[88] *Ibid.*, §3.

[89] Frege also speaks of the 'legitimacy' of the semantic stipulation as having been 'established once for all' and makes reference to his intention to 'develop the whole wealth of objects and function treated of in mathematics out of the germ of the eight functions whose names are enumerated in vol. I, §31' (*Gg*, vol. II, §147).

has been assigned to the other primitives of Begriffsschrift.[90] The complete argument of §§30–1 has a more general conclusion: that the stipulations provide every well-formed expression with a reference – and not just *a* reference, at least one reference, but a *unique* reference. Since Frege argues in §31 that a reference has been assigned to each of the primitive expressions, he need only show that, if every primitive expression of the language has a reference, then every expression that can be formed from these primitives also has a reference. That argument – it may be the first proof by induction on the complexity of expressions ever given – is contained in §30. In fact, the section contains two things, which are not separated in Frege's exposition: A reasonably precise account of the syntax of Begriffsschrift and a demonstration that every expression correctly formed from referring expressions refers. Frege explains that complex names are formed by applying certain combinatorial operations to the primitive expressions of the language and that every name is formed by successive applications of these operations. This 'closure clause' serves to define the class of well-formed expressions by means of the ancestral and so implies the validity of proof by induction on the complexity of expressions:[91] It is this that allows Frege to argue that, if all primitive expressions of Begriffsschrift refer, then every well-formed expression refers, by arguing that the two ways of forming complex expressions from simpler ones preserve referentiality. The proof is not trivial: The argument that complex predicates – such as '$\xi = \xi$' – denote is both subtle and elegant.[92]

The primitive expressions of the Begriffsschrift are indeed listed in §31, but it is hard to believe that Frege refers to it at this point simply for that reason: Rather, the argument given in §31 is what shows that all of these expressions refer and that is what *makes* them legitimate.

[90] That Frege should *argue* for this claim contradicts Weiner's view that, for Frege, 'no work is required to show that primitive terms have *Bedeutung*' (Weiner, *Frege in Perspective*, p. 129). To be sure, not much work is required to show that most of them refer, but a *lot* of work is required to show that the smooth breathing does.

[91] Thus, Weiner's objection that the induction principle employed in this proof is never stated (*ibid.*, p. 240) is met, since no special induction principle needs to be stated here.

[92] A complex predicate is one formed by omitting occurrences of one term from another, leaving argument-places in its wake. Thus, one can form the complex predicate '$\xi = \xi$' by omitting both occurrences of t from '$t = t$'. See, again, Heck, '*Grundgesetze der Arithmetik* I §§29–32', for discussion of the argument.

Frege's argument that the smooth breathing denotes is complex and difficult to interpret. For present purposes, we do not need to discuss its details, but there is a feature of the argument that is worth mentioning. Frege takes it to be enough to prove:[93]

(I) If $\Phi(\xi)$ and $\Psi(\xi)$ denote, then $\ulcorner \grave{\varepsilon}\Phi(\varepsilon) = \grave{\varepsilon}\Psi(\varepsilon)\urcorner$ denotes;

(II) If $\Phi(\xi)$ denotes, and if p denotes a truth-value, then $\ulcorner p = \grave{\varepsilon}\Phi(\varepsilon)\urcorner$ denotes.

Claim (I) is supposed to follow from the semantic stipulation governing the smooth breathing, that $\grave{\varepsilon}\Phi(\varepsilon) = \grave{\varepsilon}\Psi(\varepsilon)$ has the same reference as $\forall x[\Phi(x) = \Psi(x)]$, the latter formulae itself having a reference because the expressions from which it is constructed do. To establish (II), Frege needs to specify whether the truth-values are value-ranges and, if so, which ones they are: If they are not value-ranges, $p = \grave{\varepsilon}\Phi(\varepsilon)$ will be false (and so will denote); if they are, then p will have the same reference as some expression of the form $\grave{\varepsilon}\Psi(\varepsilon)$, whence $p = \grave{\varepsilon}\Phi(\varepsilon)$ will have the same reference as some sentence of the form $\grave{\varepsilon}\Phi(\varepsilon) = \grave{\varepsilon}\Psi(\varepsilon)$, and (II) will reduce to (I). In § 10, Frege argues that it is consistent with the other semantic stipulations that the truth-values are their own unit classes and then stipulates that they are.

It is often said that Frege needs to make this stipulation because he requires every predicate to denote a total function, one that has a value for every argument. This is right, but we are now in a position to appreciate the reason for this requirement: It is imposed by the purpose of the proof being given in § 31 and, more generally, by the fact that Begriffsschrift is supposed to have a classical semantics. The truth-values of complex sentences are specified in terms of the references of their simpler components, by means of the truth-tables and the usual sorts of (objectual) stipulations for the

[93] What Frege needs to show is that $\Delta = \grave{\varepsilon}\Phi(\varepsilon)$ denotes, so long as Δ and $\Phi(\xi)$ do. His assumption that these two cases are the only ones that need to be considered involves a tacit restriction of the domain to truth-values and value-ranges. If the domain contains only such objects, then each of them is either the value of p, for some assignment of a truth-value to p, or the reference of $\grave{\varepsilon}\Phi(\varepsilon)$, for some assignment of a function to $\Phi(\xi)$, since every value-range is the value-range of some function. Thus, the oft-heard claim that, for Frege, the quantifiers always have an unrestricted range is false. For further discussion of this matter, see Heck, '*Grundgesetze der Arithmetik I* §10'.

quantifiers. If $\Delta = \grave{\varepsilon}\Phi(\varepsilon)$ did not have a reference, when Δ denotes a truth-value, $\forall x(x = \grave{\varepsilon}\Phi(\varepsilon))$ would not have a reference, and the argument would collapse.[94] The stipulation that the truth-values are their own unit classes thus plays an essential role in Frege's proof that every well-formed expression denotes, and it is not mentioned outside § 10 – except in § 31 and a handful of sections that themselves refer to § 31. In particular, the stipulation *is not embodied in the axioms and rules* of the Begriffsschrift. The sentence stating that the truth-values are their unit classes is neither provable nor refutable in the Begriffsschrift, as Frege essentially shows in § 10. Of course, he could have adopted this sentence as an additional axiom: But the reason Frege needs to make the stipulation has nothing to do with the syntax of the formal theory but rather concerns its semantics, so Frege does not bother to make such a stipulation.[95]

The purpose of §§ 30–1 is thus to prove[96] that every well-formed expression in Begriffsschrift refers (and, in particular, that the smooth breathing does). It follows (or would follow, were the argument not fatally flawed) that Basic Law V is true and, moreover, that the system is consistent, since all axioms of the theory are true, the rules are truth-preserving, and there is a sentence – the sentence '$\forall x(x \neq x)$' will do – that is assigned the value False by the stipulations and so is not a theorem, since every theorem has the value True. As we have seen, the argument makes heavy use of semantic notions, in particular, the notion of reference. Moreover, although the argument that the smooth breathing refers is flawed,

94 And its reason for collapsing would be quite independent of whether any *term* of the language – let alone any primitive term – denotes a truth-value. For the significance of this fact, see *ibid*.

95 Parallel remarks could be made about Frege's stipulation, in *Gg*, vol. I, §11, concerning the references of improper descriptions.

96 Weiner has argued that 'there are serious obstacles to reading §§28–31 as the presentation of a proof' (Weiner, *Frege in Perspective*, p. 240). She notes that the conclusion of the proof is not used in Frege's proofs of the axioms of arithmetic (p. 242). But the proof is metatheoretic: Its conclusion is a claim *about* Begriffsschrift; there is no reason that appeal need be made to it in later proofs. She also says that the argument does not meet the standards for 'a metatheoretic proof in an introductory logic course' (p. 240). But it should not be surprising if Frege is unclear about the conceptual underpinnings of the argument, since it is likely the first metatheoretic argument ever given. And, with the exception of the failed proof that the smooth breathing denotes, I'd have to disagree: It's a *very* sophisticated proof, especially the part concerning complex predicates.

there is nothing wrong with the remainder of the proof. The remainder of §§ 30–1 constitutes a correct proof that the semantic stipulations governing the primitive logical expressions suffice to assign each of them a unique reference – and so suffice to assign a unique reference to every expression properly formed from them. Since the semantic justifications really do show that the axioms and rules of the Begriffsschrift, other than Basic Laws V and VI, are true and truth-preserving, respectively, part I of *Grundgesetze* contains a correct proof that the logical fragment of the Begriffsschrift – that is, Frege's formulation of second-order logic – is sound, that is, that all of its theorems are true.

5 CLOSING

We have thus seen that, in *Grundgesetze*, Frege gives a number of arguments whose purpose is to show that the axioms and rules of the Begriffsschrift are, respectively, true and truth-preserving. There are the semantic justifications of the axioms and rules, found scattered throughout part I; and there is the argument of § 30–1, which is not only supposed to show that Basic Law V is true, but that every well-formed expression has a reference. These arguments have explicitly semantic conclusions, and they make heavy use of semantic notions. Their character makes it extremely unlikely that they are intended merely as a peculiar sort of foreign language instruction. Such oft-heard claims as that 'Frege never raises any metasystematic question'[97] or, more strongly, that 'metasystematic questions as such … could not meaningfully be raised' by him[98] are therefore doubtful, at best.

One could yet question how seriously these apparently semantic arguments are to be taken, on the ground that, if they are to be understood as 'properly scientific', rather than as 'elucidatory', they would have to be formalizable in the Begriffsschrift itself. And perhaps, for some reason or other, Frege would have denied that semantic arguments *could* be formalized in the Begriffsschrift. But why?[99] Of course, it follows from Tarski's theorem that, since

[97] Van Heijenoort, 'Logic as calculus', p. 326.

[98] Dreben and van Heijenoort, 'Introductory note', p. 44.

[99] One might have thought that the concept horse problem would pose technical difficulties: But that problem does not arise when the argument is carried out in

the Begriffsschrift formalizes a classical theory sufficient for arithmetic, if its own truth-predicate is definable in it, it is inconsistent. But *Frege* had no reason to think this and so no reason to think that the semantic arguments he gives in *Grundgesetze* could not be formalized in the Begriffsschrift. Indeed, the natural view would surely have been that such reasoning can be reproduced within the Begriffsschrift – which, indeed, it can.[100] So the Begriffsschrift is inconsistent. Again.

Such terms as 'metalogical perspective', 'semantic metaperspective' and 'metasystematic standpoint' – these being the buzzwords of a now familiar tradition in Frege scholarship – are deeply misleading:[101] There is an almost subliminal suggestion that semantic reasoning requires a perspective beyond the Begriffsschrift, that such reasoning *cannot* be carried out within it. But the mere fact that the conclusion of an argument concerns the semantic properties of a particular theory does not show that it cannot be formalized within it: though not all arguments for semantic claims concerning Peano Arithmetic can be formalized within Peano Arithmetic, many can be.[102] Nor are semantic claims about PA the only ones that cannot be proven in PA: That PA is consistent is a *syntactic* claim, purely syntactic proofs of which (for example, Gentzen's) cannot be carried out in PA.

But we do need to be careful here. Ricketts claims, at one point, that 'anything like formal semantics, as it has come to be understood in the light of Tarski's work on truth, is utterly foreign to Frege'.[103] This claim I think I have shown to be untenable. But I

a higher-order formal theory, but only when one is attempting to talk about the semantics of Begriffsschrift in natural language.

100 Tarski shows us how to formulate a definition of truth for a second-order language in a third-order language. But Basic Law V can be used to reduce quantification over third-level concepts to quantification over second-level concepts – or, indeed, objects.

101 Tappenden, in 'Metatheory', has well documented the extent to which certain forms of argument have become something akin to secret handshakes among the members of this tradition.

102 For example, a materially adequate definition of truth for Σ_n sentences, for any n you like, can be formulated within PA, and using these definitions one can then give a semantic proof of the consistency of Σ_n arithmetic, for every n. But there is no way to paste all these definitions together in PA to get a definition of truth for the whole of the language of arithmetic. Fortunately.

103 Ricketts, 'Objectivity and objecthood', p. 67.

have not argued that semantics, seen in the light of Tarski's work on *logical consequence,* is not foreign to Frege. The mathematical work at which we have looked is concerned with such questions as whether the axioms are *true,* or whether the rules are *truth-preserving,* or whether the primitive expressions of *Begriffsschrift refer.* None of the work at which we have looked addresses such questions as whether the axioms are *logically* true or the rules are *logically* valid. And although I have argued that Frege ought to have been, and was, interested in these questions, it is unclear whether he thought mathematical work might bear upon them, let alone whether he would have accepted Tarski's characterization of the notion of logical consequence (or some alternative).[104] Though there are indications that, a few years after the publication of *Grundgesetze*, Frege was beginning to think about logical consequence in mathematical terms,[105] we do not, in my opinion, yet know enough to decide this interpretive issue.[106]

[104] It is perhaps worth remarking that some contemporary philosophers have also rejected Tarski's characterization of consequence, notably, John Etchemendy, *The Concept of Logical Consequence* (Cambridge, Mass.: Harvard University Press, 1990).

[105] The relevant discussion is in Frege, 'On the foundations of geometry: second series', part III. For discussion of these passages, see T. Ricketts, 'Frege's 1906 foray into meta-logic', *Philosophical Topics*, 25 (1997), pp. 169–88; and Tappenden, 'Metatheory'.

[106] I am fortunate to have had very helpful comments on earlier drafts of this paper: Thanks are due to George Boolos, Tyler Burge, Warren Goldfarb, Michael Kremer, Ian Proops, Thomas Ricketts, Alison Simmons and Joan Weiner.

I have two very large debts, which I decided not to acknowledge at every point at which they were felt, as that would have cluttered the paper. The first is owed to Jamie Tappenden. While he was visiting at Harvard, during the 1994–5 academic year, we had an extraordinarily fruitful, year-long discussion about Frege and, in particular, the issues with which this paper is concerned. The second is to Jason Stanley. Much of the first half of this paper was born in conversation with him; it is difficult to remember which ideas originated with whom. But I am reasonably certain that he is responsible for all the mistakes.

The first draft of this paper was written in the summer of 1995; it reached essentially its current form in the summer of 1997. That it remained unpublished for so long is due to circumstances over which I had no control. That said, I wish nonetheless that I had felt able to take more serious account of papers published or written since, but Frege scholarship is extremely active nowadays, and that would essentially have required me to rewrite the entire paper. I have therefore added references to some recent work, but otherwise have chosen to be silent. That silence should not itself be interpreted.

MARK WILSON

10 Frege's mathematical setting

I have not yet any clear view as to the extent to which we are at liberty arbitrarily to create imaginaries, and to endow them with supernatural properties.

John Graves, 'Letter to William Hamilton'

INTRODUCTION

Although Gottlob Frege was a professional mathematician, trained at one of the world's greatest centres for mathematical research, it has been common for modern commentators to assume that his interests in the foundations of arithmetic were almost entirely 'philosophical' in nature, unlike the more 'mathematical' motivations of a Karl Weierstrass or Richard Dedekind. As Philip Kitcher expresses the thesis:

> The mathematicians did not listen [to Frege because] ... none of the techniques of elementary arithmetic cause any trouble akin to the problems generated by the theory of series or results about the existence of limits.[1]

Indeed, Frege's own presentation of his work easily encourages such a reading. Nonetheless, recent research into his professional background reveals ties to a rich mathematical problematic that, *pace* Kitcher, was as central to the 1870s as any narrow questions about series and limits *per se*. An appreciation of the basic facts involved (which this essay will attempt to describe in non-technical terms)

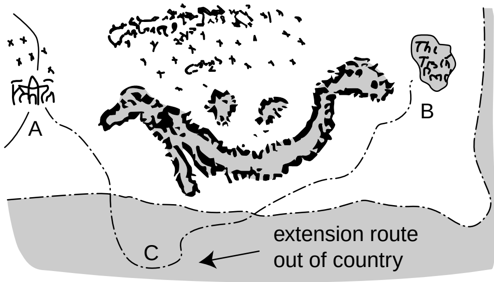
[1] Philip Kitcher, 'Frege's epistemology', *Philosophical Review*, 88 (1992), pp. 235–62. The prevalence of such attitudes is surveyed by Jamie Tappenden, 'Extending knowledge and "fruitful concepts"', *Noûs*, 29 (1995), pp. 427–67.

can only heighten our appreciation of the depths of Frege's thought and of the persistent difficulties that any adequate philosophy of mathematics must confront. Although it may be possible to appreciate Frege's approach to language on its own terms, some awareness of the rather unusual *examples* that he encountered in the course of his mathematical work can enhance our understanding of his motivations within linguistic philosophy as well.

In the most general terms, the ontological world of nineteenth-century mathematics expanded far beyond its traditionally circumscribed boundaries, a phenomenon that first became evident in the *extension element problems* that we shall emphasize in this essay. In response, a philosophy of *relative logicism* emerged that sought to explain the mysterious new entities as *logical constructions* of some sort or other. The *absolute logicism* that Frege proposed with respect to the regular number systems can be viewed as a natural outgrowth from, and improvement upon, these established logicist traditions. Many of Frege's methodological remarks enjoy a sharper piquancy, I believe, if they are examined against this richer mathematical backdrop (Frege rarely draws explicit notice to such issues, but his central examples ('the direction of a line') often represented commonplaces within the prior discussions). Beyond its relevance to Frege's thinking, an acquaintance with the extension element problem can revive our own appreciation of the weird wonders of the philosophy of mathematics, lest we forget about the *unexpected factors* that frequently force mathematics to alter its courses in uncharted ways.

EXTENSION ELEMENTS WITHIN GEOMETRY

In working on a mathematical problem, we often find it hard to reason rigorously directly from point A to point B, due to some barrier (imagine a large mountain as a metaphor) lying between the points. However, we can sometimes espy another location C *outside the borders of our native country* that would sustain an easy path A → C → B. An early illustration of this phenomena dating to the 1530s can be found in the problem of extracting the roots of cubic and quartic equations: mathematicians such as Gerolama Cardano uncovered algebraic techniques that eventually led to the real roots desired, but their computational pathways wandered through



strange intermediate values such as '$-3 + \sqrt{-2}$'. As time wore on, these intervening 'imaginary' (or 'complex') numbers gradually assumed a vital importance within mathematical practice generally, but the exact rationale of their employment, beyond raw expediency, remained hazy.

In the first half of the nineteenth century, a host of 'foreign elements' invaded a plethora of traditional mathematical subjects and it became clear that some newer philosophy of mathematics was needed to rationalize their employment. Once analytic geometry (= the use of *algebraic equations* to represent geometrical facts such as 'line L crosses circle C') had been invented in the 1600s by Descartes and Fermat, it was quickly observed that the conclusions of standard Euclidean argumentation could be often replicated by swift manipulation on formulae. Furthermore, the algebraic pathway to a geometrical conclusion often reaches its results without engaging in all of the delicate fussing about subcases that one finds in Euclid (proofs in his traditional diagram-based, constructive style are commonly called 'synthetic', in contrast to an algebraic 'analytic' approach). What secret power permits this dramatic algebraic simplification? An inspection of the 'analytic' proof indicates that its reasoning pathways often travel through intermediate 'points' bearing strange coordinate locations such as $<2.5, -3 + \sqrt{-2}>$.

But how can reasoning developed for thinking about *numerical values* produce such a surprising unity within *geometry*? Various English mathematicians of the early nineteenth century articulated a somewhat mystical faith that the blind application of algebraic algorithms must always lead to correct results, even if the paths pursued seem completely unintelligible – this point of view was often called *the generality of algebra*. On this reckoning, the imaginary points arrive, in the phrase of the mathematician E. Hankel, as 'a gift from algebra'. Bertrand Russell expressed the obvious objection to this facile manner of thinking:

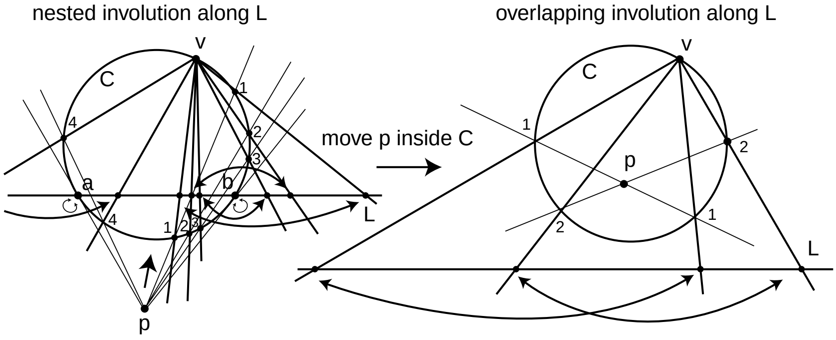
> As well might a postman presume that, because every house in a street is uniquely determined by its number, therefore there must be a house for every imaginable number.[2]

Indeed, a brute appeal to algebraic formalism is plainly an inadequate rationale for introducing novel entities into mathematics – as Russell suggests, one would quickly reach ridiculous conclusions if one applied 'the generality of algebra' in all walks of life.

In the 1820s, a number of synthetic geometers, led by J.-L. Poncelet, decided that the simplifications offered by algebraic proof must spring from a deeper source: namely, the world of standard geometry could be greatly improved if mathematicians would tolerate a variety of 'extension elements' lying just outside the limited perimeters drawn within traditional Euclidean thinking (in the manner that convenient location C lies beyond the borders of our mountainous country). These early 'projective geometers' typically justified their unseen supplements through appeals to '*persistence of form*', a methodological doctrine we shall explore in a moment. However, this thesis is greatly troubled by its own inherent vagueness and, by Frege's time, most rigour-minded mathematicians had replaced appeals to 'persistence of form' by *relative logicism*: the claim that the extension elements can be justified using *purely logical resources* alone. The methodological appeal of this newer point of view has diminished over time, its percepts having become displaced in turn by Hilbert-style axiomatics at the turn of the twentieth century (in a manner to be considered at the end of this essay). Frege's methodological motivations have often been misunderstood, I believe, through a failure to properly locate their placement within these forgotten relative logicist traditions.

Let us now examine how the original 'persistence of form' thinking operated, because relative logicism should be seen as an adaptation of its basic contours (here, and in several other portions of the essay, I will include details that a casual reader may wish to

[2] Bertrand Russell, *An Essay on the Foundations of Geometry* (New York: Dover, 1956), p. 44.



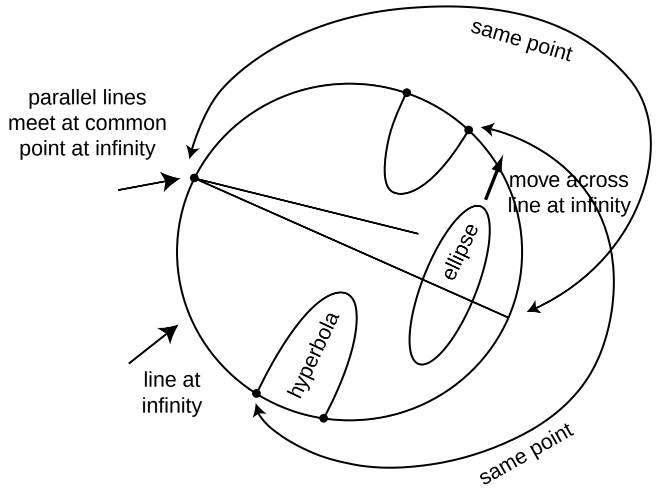
skip – they are provided to help interested parties find their way through the standard history of mathematics literature).[3] Let's begin with a simple circle **C** and a line **L** running through **C**. Their intersections engender two regular points **a** and **b**. We will now tell a story of how **a** and **b** can *become invisible* if their surrounding geometrical relationships become altered in a natural way. We will set up a somewhat complicated web of characteristic geometric construction around **a**, **b**, **C** and **L** and then slowly adjust their internal relationships so that **a** and **b** seem as if they have been merely 'pushed off the page'. First locate the exterior point **p** such that rays emerging from **p** intersect the circle tangentially at **a** and **b** (see the left-hand side of our diagram). This **p** is called the *pole* of the line **L** relative to **C** while **L** reciprocally serves as the *polar* of **p** (such 'pole and polar' arrangements possess many interesting geometrical properties). Note that an arbitrary ray from **p** will typically cross **C** in two spots, linking together pairs of points on the circle in the match-ups indicated by the small numbers. Let's now choose an arbitrary new point **v** upon **C** and use it to project our **C**-based match-ups onto the line **L**. Here we should think of **v** as acting like the lamp in a film projector that projects the circular match-ups registered within **C**'s 'film' onto the 'screen' represented by line **L** (this variety of mathematics is called 'projective geometry' precisely because it studies the transfer of ordering arrangements from one surface to another). The projected image seen on **L** will be a *nested* mapping of points to

[3] A good source is Jeremy Gray, *Worlds Out of Nothing* (London: Springer, 2007). For an important philosophical examination of 'proof unity', see Ken Manders, 'The Euclidean diagram' in P. Mancosu (ed.), *The Philosophy of Mathematical Practice* (Oxford: Oxford University Press, 2008), pp. 112–83.

one another known as an *involution.* Such mappings display a host of special geometric properties prized by the ancient geometers, including the fact that the distances x and x^* of the paired points will obey the relationship $x.x^* = +D$ within a suitable coordinate system. Plainly, our original **a** and **b** points serve as the two *centres* of this nested involution in an obvious way (their locations x satisfy the 'self-correspondent' condition $x.x = +D$).

Here are the considerations that encourage 'persistence of form' thinking. Picture our diagram's maze of lines as if they constitute a little mechanism whose moveable parts are linked to one another. Let's now adjust those parts by gradually pushing the pole point **p** *inside* **C** (conceive of **p** as a lever that forces the other parts of the diagram to shift positions). We can easily 'see' what will happen: as **p** moves towards **C**, its polar **L** will shift in the opposite direction, eventually leading to the situation pictured on the right-hand side of the diagram, with **p** inside **C** and **L** outside. Our 'lamp' **v** continues to project an involution match up onto **L** but its point-wise associations will become *overlapping* after **p** crosses into **C**. And, finally, our two erstwhile centres **a** and **b** seem to vanish.

Or do they? Why not assume **a** and **b** remain present, but have merely become *invisible* through being *pushed off the page*? That is, the collective geometrical 'mechanisms' on the left and right sides of our diagram should be regarded as essentially the same, except that their **a** and **b** parts can no longer be 'seen' in left-hand circumstances. Thus 'persistence of form': we conclude that our diagram's missing **a** and **b** are still present in left-hand circumstances, because the same geometrical unities (= 'form') preserve themselves as we gradually adjust our diagrams. Revisiting algebra's ability to simplify traditional proofs from this new point of view, we recognize that algebra obtained its unificatory advantages by automatically supplying imaginary coordinate names to extension elements that, properly speaking, should have been added to traditional geometry through 'persistence of form' considerations. That is, 'the generality of algebra' achieved its apparent successes within a geometrical context only through a happy accident: its computational procedures just happened to provide names for the auxiliary elements required to keep the organic 'mechanisms' of Euclidean geometry intact under adjustments in 'form'. Considering our involution equation $x.x^* = +D$, we find that, as **p** moves inside **C**, an originally



positive D gradually shrinks to 0 and then becomes negative once **p** moves inside **C**. Solving the 'self-correspondent' condition *x.x* = –D for its 'centres', we find that our missing centres **a** and **b** take up the imaginary coordinate locations +√–D and –√–D along **L**. So our invisible **a** and **b** are not truly 'gifts from algebra'; their real sources are the *invariant properties* contained in our family of geometrical constructions.

Once this extraordinary ontological gambit is accepted, we realize that traditional Euclidean proofs were often complicated because they could not appeal to vital parts of a geometrical construction that had been pushed into 'invisibility'. This exclusion forced traditional argumentation to work around the missing pieces by dividing a proof into a large number of subcases, distinguished from one another according to their missing parts. Restoring the invisible ingredients allows us to treat a multitude of cases in a unitary manner.

An allied simplification of Euclidean proof can be also achieved by tolerating a supplementary *line at infinity* such that, if an ellipse is moved across its bounds, the full figure will reappear in our 'local space' as an hyperbola (such an identification of seemingly different figures again permits a great simplification in our proofs). In the illustration, our regular local Euclidean plane has been contracted to lie inside the central circle, so that the manner in which an ellipse moves across the line at infinity can be observed (note that diametrical opposite points along the line at infinity are to be identified). Observe that normally parallel lines will now intersect at points upon this infinitely distant line – we will revisit these 'points at infinity' in our discussion of Frege's *Grundlagen*.

In 1851 the mathematician H. J. S. Smith explained the 'persistence of form' doctrine as follows:

> [I]f we once demonstrate a property for a figure in any one of its general states, and if we then suppose the figure to change its form, subject of course to the conditions with which it was first traced, the property we have proved, though it may become unmeaning, can never become untrue, even if every point and every line, by means of which it was originally proved, should wholly disappear.[4]

Because our diagram's 'pole and polar' persist as point **p** continuously moves inside **C**, we may postulate 'unmeaning' (= without representation in intuition) *ideal points* to support the continuation of such properties. It is plain to see that such an unrefined methodological doctrine can easily lead to gross error if one improperly posits 'persistent elements' through its aid. Indeed, relative logicism attempts to provide such a corrective. It should be noted that most programs of this nature continue to work with the *evaluative concepts* that stand at the centre of the standard 'persistence of form' judgements. With respect to our invisible **a** and **b** case, the crucial property 'sets up an involution along **L** based upon **C** and **v**' serves as the central 'evaluative concept' that most relative logicist treatments exploit in introducing their own **a** and **b** surrogates in a better way.

Although we can't survey such issues here, many of the rival methodologists that Frege criticized (e.g. Hermann Schubert)[5] maintained that unsupplemented appeals to 'persistence of form' can

[4] H. J. S. Smith, 'On some of the methods at present in use in pure geometry', in *Collected Papers*, vol. I (New York: Chelsea, 1965), p. 4.

[5] H. Schubert, *Mathematical Essays and Recreations* (Chicago: Open Court, 1910). Schubert interprets Kronecker's celebrated pronouncement, 'God created the integers; all else is the work of man' as expressing the thesis that the other entities of mathematics are engendered from the natural numbers through 'persistence of form'-like 'free creativity'. Jeremy Heis has recently directed my attention to allied remarks in the influential *Logik* by Wilhelm Wundt (Stuttgart: Ferdinand Enke, 1880–3). For an excellent discussion of the general manner in which Frege's uncharitable readings of his rivals have distorted our modern appreciation of their merits, see W. W. Tait, 'Frege versus Cantor and Dedekind: On the

provide an adequate defence for conceptual innovation within mathematics (including the introduction of the natural numbers in the first place). In such critics' behalf, we might observe note that 'persistence of form' doctrines directly highlight the *epistemological considerations* that actually inspired the postulation of the extra mathematical entities whereas the *motives* that drive conceptual development within mathematics are often left obscure in logicist accounts.

EXTENSION ELEMENTS WITHIN NUMBER THEORY

Relative logicism's career can't be completely appreciated without some knowledge of parallel developments that arose in connection with the 'ideal numbers' of algebraic number theory (once again, the unconcerned reader may skim this section).

The original impetus for introducing 'ideal numbers' came when C. F. Gauss wrote of 'complex integers' in his *Disquisitiones Arithmeticae* in 1801.[6] In themselves, 'complex integers' are nothing new; they simply comprise numbers of the form $a + b\mathrm{i}$ where a and b are normal integers (= 'whole numbers') and $\mathrm{i} = \sqrt{-1}$. One of the most salient facts about the regular integers is that they break *uniquely* into *prime factors* (i.e., 24 can be only expressed as 2 x 2 x 2 x 3) whereas a more general number such as π or $6 - 2\mathrm{i}$ can be decomposed into myriad sets of factors. The great advantage of possessing prime factors is that they allow a great deal of control over the integers that gets typically lost within the more amorphous realms of number. But Gauss realized that if we remain within a *restricted orbit* of complex numbers (his 'complex integers'), then a variety of unique factorization persists within this enlarged realm, with all the advantages to be gained therefrom. Unique factorization then allowed Gauss to answer certain important questions in number theory easily, e.g. how to characterize all integers whose fourth powers give a remainder of n when divided by p. As in our geometrical case, once we envision the regular integers as enriched with a slightly extended halo of 'complex integers', the commonalities of

concept of number', in W. W. Tait (ed.), *Early Analytic Philosophy: Frege, Russell, Wittgenstein* (Chicago: Open Court, 1997).

[6] K. F. Gauss, *Disquisitiones Arithmeticae*, trans. A. A. Clark (New Haven: Yale University Press, 1965).

behaviour amongst the regular integers become more transparent. This fact impressed Gauss greatly:

It is simply that a true basis for the theory of the biquadradic residues [i.e., the questions about fourth powers] is to be found only by making the field of the higher arithmetic, which usually covers only the real whole numbers, include also the imaginary ones, the latter being given full equality of citizenship with the former. As soon as one has perceived the bearing of this principle, the theory appears in an entirely new light, and its results become surprisingly simple.[7]

In the 1840s, treating matters related to Gauss's investigations and to Fermat's 'last theorem', E. E. Kummer realized that unique factorization becomes lost again as we move out to further collections of generalized 'integers'. Consider the 'algebraic integers' that arise when $\sqrt{15}$ is added to the rational numbers. In this range of numbers, 10 breaks into irreducible factors in two distinct ways: as 2.5 *and* $(5 + \sqrt{15})(5 - \sqrt{15})$. *If* we only had further factors to work with, e.g. $\sqrt{5}$ and $\sqrt{3}$, unique factorization could be restored in this realm because $2 = (\sqrt{5} + \sqrt{3})(\sqrt{5} - \sqrt{3})$, $5 = (\sqrt{5})^2$, $5 + \sqrt{15} = \sqrt{5}(\sqrt{5} + \sqrt{3})$ and $5 - \sqrt{15} = \sqrt{5}(\sqrt{5} - \sqrt{3})$. In such terms, 10 can be seen as 'really' decomposing into $(\sqrt{5})(\sqrt{5})(\sqrt{5} + \sqrt{3})(\sqrt{5} - \sqrt{3})$ – the apparent non-unique factorizations of 10 arising as these four basic ingredients get paired up in different ways. But we can't remedy the situation by simply including $\sqrt{5}$ and $\sqrt{3}$ along with the numbers generated by $\sqrt{15}$, because that closure will include a lot of values we don't want. Kummer finessed these difficulties in an intriguing way. He would only add an unspecified 'ideal number' to the $\sqrt{15}$ field to capture the *highest commonality* between 2 and $5 + \sqrt{15}$, without identifying the missing 'factor' concretely with '$\sqrt{5} + \sqrt{3}$' or any other concrete representation of that type. Instead, he let the *pairing*[8] '$(2, 5 + \sqrt{15})$' name his desired 'ideal number' and observed that other pairs such as '$(4, 10 + 2\sqrt{15})$' must denote the same 'ideal factor'. He wrote:

In order to secure a sound definition of the true (usually ideal) prime factors of complex numbers, it was necessary to use the properties of prime factors of complex numbers which hold in every case and which are entirely

[7] Quoted in L. W. Reid, *The Elements of the Theory of Algebraic Numbers* (New York: MacMillan, 1910), p. 208.

[8] '(n,m)' is standard notation for the greatest common divisor of n and m.

> independent of the contingency of whether or not actual decomposition takes place; just as in geometry, if it is the question of the common chords of two circles even though the circles do not intersect, one seeks an actual definition of these ideal common chords which shall hold for all positions of the circles. There are several such permanent properties of complex numbers which could be used as definitions of ideal prime factors ... I have chosen one as the simplest and most general ... One sees therefore that ideal prime factors disclose the inner nature of complex numbers, make them transparent, as it were, and show their inner crystalline nature.[9]

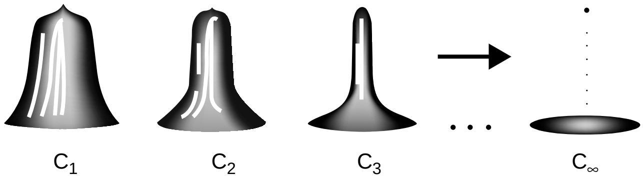
In other words, a specific group of algebraic numbers may cry out for supplementary 'ideal factors' to consolidate their behaviours into a fully satisfactory domain. In a famous letter to Kronecker, Kummer compares this enlargement process to the postulation of unseen elements in chemistry (an apt comparison because, in the chemical doctrine of Kummer's time, such 'elements' were never supposed to appear in 'naked' form in nature, rather like the quarks of modern science). Note that Kummer also aligns his practices with the geometrical circumstances we have surveyed.

'FREE CREATIVITY' AND RELATIVE LOGICISM

For such reasons, most mathematicians had concluded by 1860 that mathematics no longer needed to confine its researches to the more or less fixed domains characteristic of classical thinking (Euclidean geometry and the real numbers). To be sure, earlier investigators such as Euler had explored the properties of the complex numbers intently, but such researches had been largely ignored by methodologists such as Kant. Under the influence of nineteenth-century Romanticism, it became common to assert that the 'free creativity' of mathematicians allows them to explore whatever domains they may wish.

But, clearly, unbridled appeal to 'free creativity' will easily engender potential problems with respect to rigour and reliability, especially in situations where one's 'free creativity' extends to infinitary domains and processes. A celebrated illustration of these dangers arose in the context of G. F. B. Riemann's celebrated work in complex function

[9] D. E. Smith, *A Source Book in Mathematics* (New York: Dover, 1985), vol. I, pp. 120–4.



theory (an episode presumably familiar to Frege, as his teacher Alfred Clebsch had laboured to render Riemann's results mathematically respectable). Riemann had argued that the behaviour of such functions can be better understood if they are aligned with so-called 'Riemann surfaces', which are spaces that cannot always be understood in regular spatial terms. To prove key facts about his 'surfaces', Riemann relied upon an existence criterion he dubbed 'Dirichlet's Principle': if a collection of functions can be graded by positive number assignments, then some minimal function must exist within this set.[10] Here's a simple illustration of what is at issue. Take a wire rim of arbitrary shape and apply a soap film to it. Such a membrane stores internal energy according to its degree of bending; so a calculation of the energy stored within a particular coating will grade that shape in the 'positive number assignment' manner required by Dirichlet's principle. In real life, we intuitively expect that the film will eventually assume an equilibrium configuration that stores energy in a minimal way (sometimes there will be several placements that manage this). Dirichlet's principle simply converts these intuitive expectations into a general principle. But Karl Weierstrass showed that this assumption cannot be true in general. Let our 'rim' consist of a regular oval *plus* a single point above its centre. Now consider the sequence of bell-like patterns illustrated, where our film attaches to our oval and point in the manner required. As we progressively examine the sequence of shapes $\mathbf{C}_1$, $\mathbf{C}_2$, ..., we find that their total degree of bending continuously decreases but never reaches a minimum. Their limit $\mathbf{C}_\infty$ displays a discontinuous jump that disqualifies $\mathbf{C}_\infty$ from qualifying as a true soap film altogether. We have thus constructed a descending set of positive energy films whose lower bound does not represent a mathematical object of the same type as the $\mathbf{C}_i$, contrary

[10] A. F. Monna, *Dirichlet's Principle* (Utrecht: Oostheok, Scheltema and Holkema, 1975).

to Dirichlet's principle. Without some deep repair, brute appeals to Dirichlet's Principle cannot be regarded as reliable.

Such failures of intuitive expectation when infinite collections are concerned led many nineteenth-century mathematicians to decide that only *logic* could properly settle what occurs when such limits are reached. In particular, Richard Dedekind observed that normal Euclidean ruler and compass constructions will not install all the points we wish upon a straight line, but will only carry us to positions such as $\sqrt{2}$, $2\sqrt{2}$, etc. But Kantian spatial intuition can only certify the presence of points of this limited ilk, leaving a line with a lot of unfilled gaps in it. Dedekind maintained that the plugging of these 'holes' was tacitly prompted by *logical thinking* on the part of the mathematician, not by any variety of true geometrical intuition:

> All constructions that occur in Euclid's *Elements*, can, so far as I can see, be just as accurately effected [in an algebraically constructed discontinuous] space; the discontinuity of this space would not be noticed in Euclid's science, would not be felt at all ... All the more beautiful it appears to me that without any notion of measurable quantities and simply a finite number of simple thought-steps man can advance to the creation of the pure continuous number domain; and only by this means in my view is it possible for him to render the notion of continuous space clear and definite.[11]

In such doctrines, the thesis I have dubbed 'relative logicism' was born: logical thinking has a capacity to create further entities to fill in unwanted gaps within some independently given domain. The doctrine is a *relative* logicism, because logic requires properties within the preexisting domain to guide its creation of the supplementary entities.[12] Thus *logical construction* becomes viewed as the crucial methodology that allows the 'free creativity' of the mathematician to explore enlarged domains of objects unreachable by 'intuitive' consideration. Clearly, a logic-based approach might

[11] Richard Dedekind, *Essays on the Theory of Numbers*, trans. W. W. Beman (New York: Dover, 1963), p. 38.

[12] It should also be considered an *elective logicism*, in the sense that the mathematician can *choose* the specific sub-range of potential 'logical objects' she favours in order to frame a *closed extension domain* with nice properties (e.g. a ring with unique factorization). Pace those interpreters who maintain that Frege is an absolutist with respect to quantifier ranges, I believe that he is an electivist at heart. But such interpretative issues would take us too far afield here.

also avoid the vagaries of 'persistence of form' doctrine in tackling the extension element problems we have surveyed.

To be sure, both Dedekind and Frege were also *absolute logicists* with respect to the sundry number systems, which is not surprising, given that such thinking represents a natural extension of the relative logicist point of view (albeit not an obligatory move, for the latter position was accepted by many mathematicians who rejected absolute logicism itself). We shall see that Frege's own 'absolute logicist' thinking was influenced by several relative logicist programmes popular in his era.

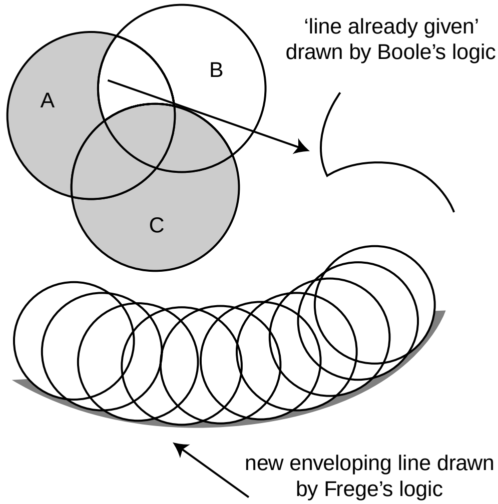
Frege plainly views the limit-fixing capacities of a proper 'logic' in a manner similar to Dedekind's. In a revealing passage where he compares the merits of his own system of logic to schemes such as George Boole's, he turns this theme to his advantage:

> If we look at the [concepts that can be defined in a logic like Boole's], we notice that … the boundary of the concept … is made up of parts of the boundaries of concepts already given … It is the fact that attention is primarily given to this sort of formation of new concepts from old ones … which is surely responsible for the impression one easily gets in logic that for all our to-ing and fro-ing we never really leave the same spot … [But i]f we compare what we have here with the definitions contained in our examples of the continuity of a function and of a limit and again that of following a series which I gave in §26 of my *Begriffsschrift*, we see that there's no question there of using the boundary lines we already have to form the boundaries of the new ones. Rather totally new boundary lines are drawn by such definitions – and these are the scientifically fruitful ones.[13]

Start with simple concepts **A, B** and **C**. From these, traditional formal logic could only construct simple compounds like **(A & ~B)** ∨ **C**, which corresponds to the grey region within the illustrated trio of Euler's circles (a.k.a. 'Venn diagrams'). Note that the boundary of

[13] 'Boole's logical calculus and the concept-script', in *PW*, p. 34. A similar passage, other aspects of which are discussed in Tappenden, 'Extending concepts', can be found in *Gl*, §88.

In the context of Fourier series, A.-L. Cauchy had mistakenly assumed that the limit of continuous functions must be continuous, when this property obtains only if the generating functions are 'uniformly continuous'. This distinction, introduced by Stokes and Weierstrass, hinges upon distinctions of quantifier scope. Cauchy's error, which was widely discussed in the 1870s, is probably what Frege has in mind here (although there are analogous examples within the calculus of variations that would have also been familiar to him).



(A & **~B)** ∨ **C** is comprised of arcs from the circles **A,B,C**. If logic could range no further from home base than that, its powers would truly prove as circumscribed as critics like Kant had assumed. Employing Frege's richer logic, utilizing both relations and second-order quantifiers, we can define the wholly distinct line that serves as the *envelope* of all the basic circles upon which it depends. This boundary is 'totally new', not coincident with any arcs of its spawning circles.

However, we must distinguish between two different logical capacities here: the ability to *express* what is required in the hypothetical bounding curve and the ability to prove that such a curve actually *exists*. The first task is accomplished by the system laid down in the *Begriffsschift*, but the second requires some supplementary doctrine about the existence of 'logical objects', in the manner, say, of the notorious 'Axiom V' of *Grundgesetze*. But there were several contemporaneous doctrines about 'logical object existence' afloat within the general relative logicist tradition and to these we shall now turn.

DEFINITION BY ABSTRACTION AND EQUIVALENCE CLASSES

In 1871, Richard Dedekind suggested both an improvement and a rationalization of Kummer's approach in a famous supplement to

Dirichlet's lectures on number theory.[14] He asks, in effect, 'What does Kummer want his "ideal numbers" to do?' *Answer*: to serve as divisors of a certain collection of algebraic numbers. 'Why,' Dedekind then proposed, 'don't we let the *entire set of numbers* we want divided *comprise* the missing "ideal number"?' That is, let us simply replace Kummer's posited 'ideal number' $(2, 5 + \sqrt{15})$ with the infinite *set of numbers* it needs to factor $\{2, 3, 3 + \sqrt{15}, 5 + \sqrt{15}, 4, \ldots\}$, a single gizmo which avoids the multiple representations to which Kummer appeals. Dedekind explained:

> [I]t has seemed desirable to replace the ideal number of Kummer, which is never defined in its own right ... by a noun for something that actually exists.[15]

Dedekind's sets (which he dubbed 'ideals') are distinguished by the fact that they are closed under the property that if elements λ and μ are already in the ideal, then so is $\alpha\lambda + \beta\mu$, where α and β are any rational numbers. Dedekind suggested that we reinterpret Kummer's procedure as follows: rather than *adding* 'ideal numbers' into an original range of numbers $\mathcal{N}$, we should instead climb from $\mathcal{N}$ to a new range of objects $\mathcal{N}^*$ formed by considering all the 'ideal' *sets that can be manufactured from* $\mathcal{N}$. The original members of $\mathcal{N}$ become replaced at the $\mathcal{N}^*$ level by their 'principle ideal' surrogates, viz., those sets that simply consist of all multiples of a single $\mathcal{N}$ element. The advantage of working within this higher domain of sets is that, unlike in $\mathcal{N}$, unique factorization obtains within $\mathcal{N}^*$. This basic format for interrelating structures, where one domain is built from another through set-theoretic processes, is now standard in modern algebra courses, although, historically, it took some time before the equivalence class approach became canonical.

In a similar vein, we could 'jump up' to a new realm of geometry $\mathcal{G}^*$ by considering as its 'points' all sets of involution mappings operating over our old-fashioned geometry $\mathcal{G}$. An old-fashioned

[14] In Richard Dedekind, *Gesammelte mathematische Werke*, vol. III, ed. R. Fricke, E. Noether and O. Ore (Braunscheweig: F. Vieweg und Sohn, 1930). Dedekind's first use of the equivalence class idea emerges, almost in passing, to introduce some modular arithmetics in 'Abriss einer Theorie der Höheren Kongruezen in Bezug auf einen Reellen Primzahl-Modulus' in vol. I of the same collection.

[15] Richard Dedekind, *Theory of Algebraic Numbers*, trans. John Stillwell (Cambridge: Cambridge University Press, 1996), p. 94. This 1877 work still provides an excellent introduction to the subject and its motivations.

point in $\mathcal{G}$ will reappear within $\mathcal{G}^*$ as the centre of a nested $\mathcal{G}$-involution, whereas the new 'points' will correspond to overlapping $\mathcal{G}$-involutions.

The basic trick displayed here – manufacturing 'new' entities by forming sets of old objects – is, of course, employed by Frege in his own construction of the natural numbers, which are treated as equivalence classes of concepts whose extensions can be mapped to one another in one-one fashion. As we shall see later, the *rationale* Frege offers for this process is rather different than that suggested by Dedekind. Nonetheless, both men regarded these set-theoretic constructions as *sanctioned by logic*. If the 'laws of thought' can build the missing elements needed to bring a mathematical domain to satisfactory ontological completion, it appears that we have finally reached a satisfactory resolution to the puzzle of the extension elements that does not upset mathematics' claims to be both a priori and grounded within intuitive sources of knowledge.

Appeals to equivalence classes will seem quite natural if one regards the novel elements as formed by *conceptual abstraction* in a traditional philosophical mode: one first surveys a range of concrete objects and then *abstracts* their salient commonalities. It is possible (but not certain)[16] that Dedekind viewed his invocation of set theory as simply a mathematical precisification of the 'abstraction' process described by earlier logicians. The notion of replacing Kummer's ideal number $(2, 5 + \sqrt{15})$ by the set 'ideal' $\{2, 3, 3 + \sqrt{15}, 5 + \sqrt{15}, 4, \ldots\}$ will seem natural because the latter set represents the source objects from whose shared features Kummer 'abstracted' his ideal factor. Indeed, the noted geometer Fredrigo Enriques explicitly rationalized Dedekind's procedures in exactly this vein:

> For it can be admitted that entities connected by such a relation [of equivalence class type] possess a certain property in common, giving rise to a concept which is a logical function of the entities in question and which is in this way *defined by abstraction*.[17]

[16] For an excellent survey of Dedekind's opinions, see Jeremy Avigad, 'Methodology and metaphysics in the development of Dedekind's theory of ideals', in J. Ferreirós and J. Gray (eds.), *The Architecture of Modern Mathematics* (Oxford: Oxford University Press, 2006). It is fairly common to employ 'abstraction' as a means of rendering a subject matter 'representation independent'.

[17] Fredrigo Enriques, *The Historical Development of Logic*, trans. Jerome Rosenthal (New York: Holt, Rinehart and Winston, 1929), p. 132. He employs the 'direction

In fact, Dedekind pursued the traditional abstractionist story a step further by recommending that, once one has 'jumped up' into the required set theoretic realm, we complete the abstractive process by *replacing* these sets by 'freely created' mathematical objects that retain only the properties we really need inside the enlarged realm itself (our set theoretic construction merely serves as a disposable ladder to lift us safely into the autonomous higher realm we seek). In an often quoted letter to Heinrich Weber, Dedekind wrote, referring to his famous articulation of real numbers as sets (= 'sections' or 'cuts') of rational numbers:

> You say that the irrational number ought to be nothing other than the section itself, whereas I prefer it to be created as something new (different from the section) which corresponds to the section and produces the section. We have the right to allow ourselves such a power of creation and it is more appropriate to proceed thus, on account of treating all numbers equally.[18]

Dedekind's 'throw away your constructive ladder after you have climbed it' represented a fairly common theme within the abstractionist tradition.

As it happens, in his own 'logical' approach to simple arithmetic, Dedekind does not bother with equivalence classes *per se*, but only employs set theory to build a specific *exemplar* of arithmetical structure. Because of this difference, Frege is often portrayed within the folklore of modern philosophical commentary as the thinker who tried to argue 'philosophically' that numbers *had to be identified* with sets of equinumerous concepts on the grounds that such identification was the only proposal that *abstracts properly from all of number's potential applications*, whereas the more 'mathematical' Dedekind sought only to articulate 'freely created' objects sufficient 'to do a mathematical job'.[19]

I find little textual evidence for attributing such motivations to Frege. He is critical of 'abstractionist' views generally and often

of L_o' example more or less as Frege did, citing the geometers Vailati and Burali-Forti in this regard.

[18] Richard Dedekind, 'Letter to Hermann Weber', in William Ewald (ed.), *From Kant to Hilbert*, vol. II (Oxford: Oxford University Press, 1996).

[19] Paul Bernacerraf, 'What the numbers could not be', in P. Bernacerraf and H. Putnam (eds.), *Readings in the Philosophy of Mathematics: Selected Readings*, 2nd edn (Cambridge: Cambridge University Press, 1983). Howard Stein, 'Logos, logic and logistiké', in W. Aspray and P. Kitcher (eds.), *History and Philosophy of Modern Mathematics* (Minneapolis: University of Minnesota, 1988).

observes that extension elements can be acceptably introduced in a wide variety of ways. Generally, the remarks that are often misread as Fregean expressions of a 'supply a unique abstractionist story for justifying the numbers' philosophy merely express the formal requirement that, however the new mathematical entities are handled, their introduction must be executed in a manner that ensures that the new objects will be properly *counted* (so that, in whatever manner we define the complex points **a** and **b**, there must be exactly two of them).

RELATIVE LOGICISM WITHOUT USING SETS

There were a number of alternative approaches to relative logicism within Frege's era that have become largely forgotten today but which seem to have influenced his own philosophical thinking. In particular, it was often emphasized that concepts should be given conceptual priority over their extensions. Christoph Sigwart wrote in an influential logic primer of the period:

> [Some logicians believe] that concepts are gained by abstraction, i.e., by a process which separates the particular objects from those by which they are distinguished from each other, and gathers the former together into a unity. But the supporters of this view forget that, in order to resolve an object of thought into its particular characteristics, judgments are necessary which have for their predicates general ideas … and as these concepts make the process of abstraction possible, they must have been originally obtained in some other way … [To try to] form a concept by abstraction in this way is to look for the spectacles we are wearing by aid of the spectacles themselves.[20]

In this regard, it was often remarked that predicative concepts can be directly converted into a species of 'concept-object', as when we frame the abstract object *motherhood* from the everyday trait … *is a mother*.

In mid-century, the German geometer Karl von Staudt[21] tacitly relied upon this observation when he proposed an influential programme for converting 'persistence of form' considerations into

[20] Christoph Sigwart, *Logic*, vol. I, trans. Helen Dendy (London: Swan Sonnenschein, 1895), pp. 248–9.

[21] K. von Staudt, *Geometrie der Lage* (Nuremberg: Bauer and Raspe, 1847) and *Beiträge der Lage* (Nuremberg: Bauer and Raspe, 1856).

more respectable patterns of definitional extension within standard geometry. Following the pattern of our *motherhood* example, he observes that we are citing a similar concept-object when we speak of *the common direction* of two parallel lines. That is, starting with the *relational concept* 'x is parallel to line L_0', logic allows us to speak instead of an abstract *object* 'the direction of line L_0'. Von Staudt then made the remarkable suggestion that these commonplace concept-objects could serve as adequate *replacements* for Poncelet's 'points at infinity' – we simply let the direction of L_0 *become* the missing 'point' that sits at the far end of L_0. In an allied vein, he suggests that we convert 'x maps to y under a right-handed overlapping involution' into a concept-object and let it replace one of the missing complex points that serve as the centres of this involution (he utilizes 'x maps to y under a left-handed overlapping involution' to instantiate the other missing centre).

Historically, the suggestion that *concepts-treated-as-objects* could be substituted in place of otherwise problematic entities was quite unprecedented in mathematical practice,[22] but, once this unexpected pill was swallowed, von Staudt found he could rationalize all of projective geometry's manoeuvres through a straightforward, if tedious, programme of redefinition. The trick is to amalgamate the new concept-objects into the old world of geometry by *redefining* our old geometrical notions to suit the new elements. Thus we must redefine the original Euclidean notion of 'lying upon' (call it 'lies upon_0') so that our new 'points at infinity' can be meaningfully held to 'lie upon_1' the line L_0 (obviously, no concept-objects can lie upon_0 L_0 if 'lies upon_0' is understood in the old sense). This process of carefully crafted redefinition must be repeated several times before von Staudt can work his way to the full conceptual world needed within extended geometry. Observe that von Staudt's programme employs simple *concept-objects* directly as replacements for the entities sought, rather than collecting together infinite *equivalence classes* in Dedekind's manner. In fact, the infinities Dedekind's technique blithely evokes were often rejected as extravagant by critics in this

[22] Hans Freudenthal, 'The impact of von Staudt's foundations of geometry', in P. Plaumann and K. Strambach (eds.), *Geometry – von Staudt's Point of View* (Dordrecht: Reidel, 1981). Mark Wilson, 'Frege: The Royal Road from geometry', in William Demopoulos (ed.), *Frege's Philosophy of Mathematics* (Cambridge, Mass.: Harvard University, 1995).

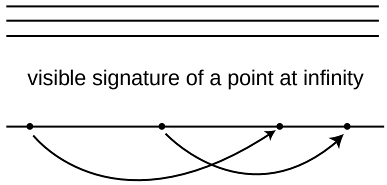
period.[23] Although Frege himself employs extensions in his own constructions, he may have originally intended to utilize simpler concept-objects in von Staudt's manner (I'll suggest how in the next section). However this may be, Frege discusses the 'direction of L_0' case in §§64–8 of the *Grundlagen* (as well as the concept von Staudt substitutes for the 'line at infinity') without remarking upon their utilization within the prior geometer's work (with which Frege was undoubtedly familiar).

As noted above, in performing these extension element introductions, we must carefully circumscribe the concepts employed so that the *requisite number* of new objects will be engendered in the conversion to concept-objects. Thus there should normally be *exactly two* complex points acting as the centres of an overlapping involution. Von Staudt distinguishes between right-hand and left-hand mappings simply as a trick for getting this 'object count' to come out properly. Frege's worries about the proper 'criterion of identity' for a specific class of 'objects' are closely tied to sharp mathematical demands such as this.

We might observe that von Staudt's approach and Dedekind's share a common theme: they both return to the original *motives* that inspired the introduction of the extra elements and search for the *evaluative concepts* that sparked such postulation. Thus the points at infinity were inspired by the evaluative concept 'x is parallel to y', for we evaluate lines L_0 and L_1 as meeting in the same infinite point only if L_0 and L_1 are parallel to one another. Both men then construct a suitable 'logical object' from the evaluative concept highlighted: von Staudt proposing that we replace a point at infinity by 'the direction of L_0' while Dedekind's approach suggests that the set $\{L \mid L \text{ is parallel to } L_0\}$ be employed.

As Frege conceptualized these issues, the evaluative concepts selected must adequately serve as the core of the 'recognition judgements' that indicate how our newly introduced elements display their handiwork within concrete mathematical circumstances (according to the philosophy 'by their fruits, you shall know them'). Furthermore, the highlighted judgements must embody proper *standards of identity* for their corresponding concept-objects.

23 Harold Edwards, 'The genesis of ideal theory', *Archive for the History of the Exact Sciences*, 23 (1980), pp. 321–78.



Question: how do we know that we are considering *the same points at infinity* in a given context? *Answer*: only if the lines with which they are associated lie parallel to one another. *Question*: how do we determine whether $(2, 5 + \sqrt{15})$ represents the *same ideal factor* as $(3, 3 + \sqrt{15})$? *Answer*: only if they satisfy the same divisibility tests for the regular numbers in the base ring.

As other essays in this volume make evident, there has been much contemporary philosophical interest in a revived 'neo-logicism' that attempts to base all invocation of 'abstract objects' upon unsupplemented 'recognition judgements' similar to the 'Hume's Principle' of *Grundlagen*, §73. I doubt that such a programme would have enjoyed Frege's philosophical imprimatur, for his central intention in highlighting 'recognition judgements' in the *Grundlagen* is to isolate the precise traits that earlier mathematicians had utilized in their vague appeals to 'persistence of form'. As such, these evaluative concepts merely provide the raw material with which a proper program for introducing the desired objects in an absolute logicist fashion might begin. The notion that claims like Hume's principle alone could constitute an adequate method for handling questions of mathematical existence would have almost certainly struck Frege as an unhappy return to the methodological vagaries of earlier times.[24]

PLÜCKER'S RECARVING OF CONTENT AND THE CONTEXT PRINCIPLE

There was a particular recasting of von Staudt's work in an algebraic vein that drew much attention during Frege's student days at Göttingen. It merits a brisk survey here, for it potentially casts a

[24] Crispin Wright and Bob Hale, *The Reason's Proper Study* (Oxford: Oxford University Press, 2004).

revealing light upon many of Frege's puzzling claims about his celebrated but obscure 'context principle'. This technique carves out simple 'concept-objects' in von Staudt's manner through reversing the *direction of functionality* within target mathematical claims. For technical reasons, Frege eventually employed equivalence class constructions in the *Grundlagen*, but the discussion preceding often suggests a sympathy for the 'reversing functionality' approach. This technique was developed in the early 1870s by the mathematicians Otto Stolz and Felix Klein,[25] following the percepts of their teacher, Julius Plücker, often regarded as 'the father of algebraic geometry' today. Plücker had introduced a revolutionary perspective into the subject by carving up previously understood 'geometrical contents' in novel ways. In so-called 'homogeneous coordinates' (see any college geometry text), the equation of a planar straight line assumes the form $Ax + By + Cz = 0$. When we first consider this equation, we naturally regard the list of constants $[A,B,C]$ as acting upon the range of variability (x,y,z). That is, we read the equation as claiming that the *function* $[A,B,C]$ carves out the *range of points* (a,b,c) that lie upon a common straight line. But what happens if we instead hold a specific point (a,b,c) fixed and allow let the erstwhile $[A,B,C]$ 'constants' to vary, i.e., we consider the reversed equation $Xa + Yb + Zc = 0$? Here we let '(a,b,c)' act as the *function* which then carves out a *range* of lines $[A,B,C]$. In fact, the locus of this new 'range' comprises a natural geometrical entity: it represents the *pencil of all lines* running through (a,b,c), whose individual rays are now distinguished by the varying 'line coordinates' $[A,B,C]$. To highlight these symmetries better, we might rewrite the claim that 'point (a,b,c) lies upon the line $[A,B,C]$' as '$[A,B,C]^T(a,b,c) = 0$' where standard matrix multiplication is employed. Then, according to whether we select the [] block or the () block as open to *variation*, we will parse our original proposition as representing the actions of distinct 'unsaturated' functions acting upon distinct ranges of saturated 'objects' (borrowing Frege's terminology from 'On concept and object'). From this point of view, a given curve can be carved with equal justice

[25] Otto Stolz, 'Die geometrische *Bedeutung* der complexen Elemente in der analytischen Geometrie', *Mathematische Annalen*, 4 (1871); Felix Klein, *Elementary Mathematics from an Advanced Standpoint: Geometry* (New York: Dover, 1941). See my 'Ghost World' in M. Beaney and E. H. Reck (eds), *Gottlob Frege: Critical assessments of leading philosophers* (London: Routledge, 2005), vol. III, pp. 157–75.

into either the union of its range of *points* or the intersection of its range of *tangent lines*, depending upon the direction of functionality chosen.[26] Readers of the Ricketts chapter in this volume will note the immediate affinities of this Plückerian point of view with Frege's own thinking upon 'range' and 'variation'.

The older geometer's work inspired a large number of contemporaneous attempts to reconfigure geometrical intuition by carving space into various choices of primitive 'elements'. The most famous of these investigations was Sophus Lie's 'sphere geometry', but Frege himself worked upon a decomposition where the 'elements' were pairs of points treated as fused unities.[27] Such examples provide a concrete (and rather startling) significance to Frege's frequent assertions that propositional contents can be 'carved up' in unexpected ways, e.g.

> [I]nstead of putting a judgement together out of an individual as subject and an already previously formed concept as a predicate, we do the opposite and arrive at the concept by splitting up the content of a possible judgement ... [T]he ideas of these properties and relations are [not] formed apart from objects: on the contrary they arise simultaneously with the first judgement in which they are ascribed to things.[28]

Returning to relative logicism, Stolz and Klein applied Plücker's 'recarving of content' point of view to von Staudt's extension element programme in an interesting fashion. When we treat the '(a,b,c)' piece of '$[A,B,C]^{\mathrm{T}}(a,b,c) = 0$' as a function, we find that no triple beginning with a zero (i.e. $(0,b,c)$) will carve out a true pencil of intersecting lines – the various lines whose coordinates $[A,B,C]$ algebraically satisfy '$[A,B,C]^{\mathrm{T}}(0,b,c) = 0$' will run parallel to one another, rather than sharing a common point. Ah ha, Stolz and Klein recognized, isn't this exactly the algebraic feature we require in a *point at infinity*? So why don't we redefine our old '(a,b,c) lies upon$_0$ $[A,B,C]$' claim so that $(0,b,c)$ becomes meaningfully permitted

[26] If we write down a formula with respect to the line coordinates $[A,B,C]$ belonging to a curve, we typically get a new equation: the 'point equation' $x^3 - y^2z = 0$ converts to the 'line equation' $4X^3 + 27Y^2Z = 0$. The latter formula reveals singularities that we might not have noticed in its 'point equation' garb. The striking revelations possible through functional recarving probably made a deep impression upon Frege's philosophical thinking.

[27] Frege, 'Lecture on the geometry of pairs of points in the plane', in *CP*, pp. 103–7.

[28] Frege, 'Boole's logical calculus', in *PW*, p. 17.

to lie upon$_1$ [A,B,C]? We only need to guarantee that we set up the right number of new 'points' when we proceed in this way (the trick is to follow the 'recognition judgement' that (0,a,b) and (0,c,d) qualify as the same 'point' if and only if they are exact multiples of one another). In other words, through a mixture of Plückerian recarving and definitional extension, we extend the reach of the expression '$[A,B,C]^T(a,b,c) = 0$' to cover point at infinity situations. Utilizing allied tricks with involutions, Stolz and Klein handled the complex points nicely as well.

Many of Frege's characteristic remarks about 'recognition judgements' and 'contextual definition' fit the Stolz/Klein techniques closely. We first extract suitable concept-objects out of a family of claims through reversed function recarving and then expand these assertions into a fuller range by installing identity conditions upon these new 'objects' through suitable 'recognition judgements'. In this regard, Frege elsewhere remarks that, if we wish, the complex geometrical points could be introduced as the (finite) 'commonalities' between an arbitrary circle **C** and any line **L** not intersecting **C**.[29] Since many distinct circle/line pairs correspond to the same imaginary points, we face the problem of finding a 'recognition judgement' that will resolve when (**C, L**) and (**C*, L***) represent the same complex points. It is only because addressing this question directly proves a bit tricky that most geometers favour involution mappings as the canonical means for introducing the complex points.

It is quite conceivable that Frege began the *Grundlagen* with a plan to introduce the integers through an allied 'functional recarving' pattern. Begin with the claim 'Concept C maps in 1–1 fashion to C_n', where C_n is some canonical concept that, logically, is satisfied by exactly n members (for 0, such a canonical concept could be '$x \neq x$'). Now reverse the direction of the functionality within our mapping claim to obtain 'The concept-object corresponding to "maps in 1–1 fashion to C_n" belongs to concept C' (or, more briefly, 'the number belonging to C_n belongs to the concept C'). 'Hume's principle' will then serve as the requisite 'recognition judgement' that determines whether two of these newly introduced 'numbers' qualify as the same or not. Under this approach, we do not require infinite Dedekind-style

[29] Frege, 'On a geometrical representation of imaginary forms in the plane', in *CP*, p. 2.

sets, but only simple concept-objects obtained through functional reversal.

However, closer analysis shows that such ploys can only supply context-dependent 'objects' that qualify only as 'incomplete symbols' in Bertrand Russell's sense and cannot behave as the entirely self-sufficient manner that naive Plücker-like thinking first suggests. Many commentators have noted that Frege's deliberations in the *Grundlagen* take an abrupt turn in §68, when, without preparation, extensions suddenly enter the scene.[30] If Frege had originally expected to apply a Plücker-like strategy to his numbers but recognized their 'incomplete symbol'-like features by the time he came to §68, his initial friendliness towards 'definitions in context' and his stress upon 'Context Principle' recarvings of content would appear better motivated. Such a mid-stream shift in strategy would explain his puzzling remark in §68:

> I believe that for 'extension of the concept' we could simply write 'concept'. But this would be open to the two objections:
>
> 1. that this contradicts my earlier statement that individual numbers are objects, as is indicated by the use of the definite article in expressions like 'the number two' and by the impossibility of speaking of ones, twos, etc. in the plural, as also by the fact that the number constitutes only an element in the predicate of a statement of number;
> 2. that concepts can have identical extensions without themselves coinciding.
>
> I am, as it happens, convinced that both these objections can be met; but to do this would take us too far afield for present purposes. I assume that it is well known what the extension of a concept is.[31]

Certainly, Plücker-like recarvings provide a more vivid application for Frege's Context Principle than do the equivalence class techniques he actually adopts (in the latter, the existence of the needed 'logical objects' must be established through axiom V-like *postulation*, rather than simple 'conceptual recarvings').

I hasten to add that it is hard, on the basis of the available texts, to establish that Frege ever had such a strategy in view. I have devoted

[30] Michael Dummett, *Frege: Philosophy of Mathematics* (London: Duckworth, 1991).

[31] Translated by J. L. Austin as *The Foundations of Arithmetic* (New York: Harper and Row, 1960), p. 80.

a fair amount of space to these precedents because (1) given his own mathematical work and training, Frege was plainly aware of these proposals and (2) Plückerian examples cast a potentially revealing light upon his often elusive remarks about 'propositional content'.

With respect to the latter, the recarving techniques suggest that modern geometers continue to traffic in the same fixed realm of *Euclidean facts* as the ancients, but over time that original domain has become progressively recarved into ever richer ranges of novel geometrical *objects* (i.e., the holistic 'propositional content' of the underlying facts do not alter under the recarvings, but their *ontological parsing* adjusts considerably). From this point of view, science should not regard a proposition's 'objective content' as altered even when its surface expression gets reconfigured in quite unexpected ways. Such themes emerge in Frege's writings in a variety of ways. For example, he often argued that, insofar as objective science was concerned, a holistically conceived proposition does not lose its 'scientific content' if it loses (or gains) some 'intuitive garb' it had previously displayed (or lacked). In his earliest mathematical work, Frege experimented with methods for aligning claims about the (affine) complex points on a plane with imagery comprised of entanglements of 3D lines above the plane.[32] The purpose of this exercise was to associate an artificial 'intuitive presentation' with the claims about the complex point facts. Frege did not regard the 'propositional content' of the original claims as altered by this annexation; the supplementation was viewed merely as a convenient tool to help the geometer *reason more easily* about the 'unintuitive' matters at hand. In §26 of the *Grundlagen*, Frege describes two imaginary creatures whose limited projective 'intuitions' correspond to different aspects of geometrical reality in classic 'inverted spectrum' fashion:

> Over all geometrical theorems they would be in complete agreement, only interpreting the words differently in terms of their respective intuitions. With the word 'point' for example, one would connect one intuition and the other another. We can therefore still say that this word has for them an objective meaning, providing only that by this meaning we do not understand any of the peculiarities of their respective intuitions.[33]

[32] Frege, 'Imaginary forms in the plane' in *CP*. J. L. Coolidge, *The Geometry of the Complex Domain* (Oxford: Clarendon Press, 1924), surveys the history of allied investigations.

Once again the implication seems to be: insofar as scientific communication is concerned, their sundry theorems traffic in the same 'objective content,' despite the different intuitive trappings in which the two creatures privately cloak these 'contents'.

Such considerations suggest the following picture of truth in mathematics. Within Euclidean geometry, the original fixed set of holistic facts is delivered to us through Kantian 'intuition', although the modern geometer can displace these original 'intuitive presentations' at will and supplement the geometrical domain with sundry 'logical objects'. Arithmetic, at first blush, seems to have its fundamental contents supplied by intuition in an allied way, but closer analysis shows that numbers secretly serve as purely logical evaluators and can be safely applied to any subject matter whatsoever. (I'll enlarge upon this reasoning in the next section.)

However, Frege's writings are not sufficiently explicit upon many of these issues although they all constitute natural responses to the scientific dilemmas of his time. Modern commentators frequently discuss Frege's notions of 'propositional content' in a manner decoupled from the rather radical methodological policies that he adopts within his own mathematical projects. I suggest that this policy of divorcement may overlook vital clues to his actual thinking.

ABSOLUTE LOGICISM

This essay has been largely devoted to the thesis of *relative logicism* as an account of how long-established mathematical domains might spawn satellite 'logical objects' to aid in understanding the original setting. *Absolute logicism*, as advocated by Frege and Dedekind, claims that various traditional mathematical domains can themselves be regarded as comprised as 'logical objects' engendered by the need to understand the structure of *non-mathematical* realms.[34] Once again, such doctrines were not spawned by philosophical musings alone, but by a mathematical need to understand

[33] *Gl*, p. 35. He presumes that these hypothetical individuals do not 'intuit' any of the metrical characteristics that break the formal duality between planar 'line' and 'point' for the likes of us.

[34] Frege also insisted that our 'numbers' must be directly applicable to *mathematical situations* as well, for we want to gauge the size of various collections of

more precisely the range of cases in which number-like evaluators could be profitably employed.

For example, the regular complex numbers can nicely compute how repeated rotations will compose *within a plane* (if we can independently manipulate an adjustable rod to reach positions **a** and **b** through operations A and B, where will the rod reach if operation B is applied after A? *Answer*: **a**.**b**). Can we find more general complex number-like gizmos that can capture *three-dimensional movements* in a comparable vein? Such research led to the sundry 'dual numbers', quaternions and allied number-like systems that were widely studied in Frege's era (such inquiries have become important once again in the context of modern robotics). Alternatively, one might try to tackle these representational problems by applying the regular complex numbers in unexpected ways. In fact, in early work[35] Frege experimented with grading a restricted class of functional representations correlated with infinitesimal rotations in this fashion, somewhat in the manner of Sophus Lie. Frege's interest in the *application problem* for the various number systems may have emerged from these background concerns: under what conditions can a calculus historically devised for purpose $\mathscr{P}$ be successfully transferred to novel purpose $\mathscr{Q}$? And the natural answer suggests itself: only if a certain *logical structure* within the local realm of traits under consideration is present. One can ascertain this vein of thinking most clearly in Frege's approach to the real and complex numbers. Although he never completed the intended developments, Peter Simons[36] has supplied a plausible delineation of how the scheme would have worked: a real number **r** is treated as an *evaluator* of a given property P*'s position within a linearly ordered family of properties P. More explicitly, to claim that '*a* is π metres long' indicates that '*a* possesses that length property L* which occupies the πth place within a broader family of length traits <$\mathscr{L}$, L_1, Abut>, where this collection represents the smallest family of traits that contains L_1 (= the property of having the same length as the

'natural numbers' through the application of these very same 'numbers' (Russell's type-based 'number' constructions, notoriously, could not do this).

[35] Frege, 'Methods of calculation based upon an extension of the concept of quality', in *CP*, pp. 56–92.

[36] Peter Simons, 'Frege's theory of real numbers', *History and Philosophy of Logic*, 8 (1987), pp. 25–44.

standard metre bar in Paris) and is also closed under end-to-end composition (L_{i+j} represents the length property framed when two objects possessing length properties L_i and L_j are abutted end-to-end). Considering π as a concept-object that marks a property's position within such a relational family, π gets identified with the set of all quadruples $<P^*, \mathscr{P}, P_1, R>$ that can be mapped onto a canonical non-empty family of properties constructed with logical materials alone (Frege would have employed his already defined integers to build up (and complete) a suitable canonical family of fraction-like properties).

Prescinding from these technical complexities, the natural numbers serve as logical evaluators of the *cardinal size* of a concept C, whereas the real numbers evaluate its *comparative position* within a linear family C of related concepts (if such a family is pertinent to C). Our philosophical task in setting up the sundry number systems is to elucidate the underlying logical basis for the relevant evaluation of the concept C and to then employ some method for providing logic-based concept-objects able to capture the assessment under review. We are thereby adopting the same basic methodology as pertains within geometry's circumstances, but our real number evaluations needn't rest upon any underlying range of intuitively supplied facts comparable to those required in geometrical assessment, simply because only the logical structure of the family $\mathcal{C}$ is wanted for their applicability. Thus we obtain an absolute logicism for the number systems that is impossible within geometrical circumstances. In rejecting the support of Kantian 'intuition' for his number systems, Frege conformed to opinions commonly shared by investigators then exploring the application range of number-like evaluators.

AXIOMATIC POSTULATION

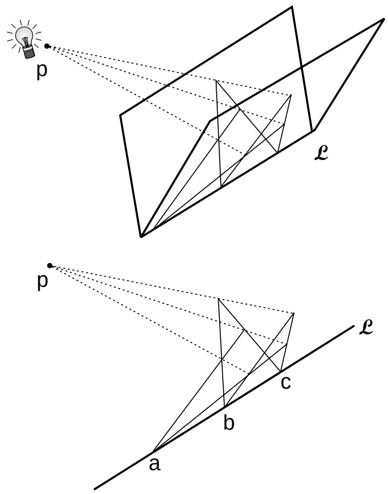
Until 1900 or so, von Staudt's programme was commonly regarded as providing the 'right explanation' for why the strange extension elements could be legitimately added to traditional geometry, although few studied his techniques in detail simply because the work involved was so tedious. However, this methodological consensus vanished virtually overnight with the rise of the axiomatic approach followed by David Hilbert and his school. Under their approach, Euclidean and projective geometry become regarded as 'implicitly defined' quite independently by their own parochial collections of axioms, leaving the question of their interrelationships to be determined by now standard model-theoretic techniques ('Can a model $\mathcal{M}$ of the Euclidean group be extended to frame a projective model $\mathcal{M}^*$?', etc.). If so, von Staudt's tiresome stagewise constructions can often be avoided: if you think that Euclidean geometry might be better understood in a system with imaginary points, *directly* specify axiomatically the richer structure you desire and then indicate how Euclidean geometry can be profitably *embedded* within it. *Don't* waste your time trying to construct what you seek out as strange and improbable concept-object slicings of your original domain.[37] The almost instantaneous popularity of this new point of view drove von Staudt's method of relying upon 'recognition judgement'-based concepts into intellectual oblivion. Post-Hilbertian commentators often sarcastically dismissed von Staudt's efforts as motivated by antiquated, 'extra-mathematical' demands upon mathematical existence, not unlike the criticisms of Frege considered at the head of this article. In this vein, the irrepressible E. T. Bell wrote:

> In proving that geometry could, conceivably, get along without analysis, von Staudt simultaneously demonstrated the utter futility of such a parthnogenetic mode of propagation, should all geometers ever be singular enough to insist upon an exclusive indulgence in unnatural practices.[38]

Essentially, Hilbert's appeals to *independent axiomatization* provided a fresh methodology for rigorously implementing the philosophy of 'free creativity' enunciated earlier: mathematicians are free to cook up any internally consistent realm they please, unfettered

[37] This is the policy recommended in O. Veblen and J. W. Young, *Projective Geometry* (Boston: Ginn, 1910). To be sure, many of von Staudt's techniques will return as methods of constructing extensions to old models, but the relative logicist suggestion that they capture the 'recognition judgements' that conceptually prompted the enlargements is abandoned. In this same regard, Hilbert helped to popularize the modern employment of Dedekind's ideals within algebra, without any pretension that the equivalence classes somehow 'abstract' from the original domain.

[38] E. T. Bell, *The Development of Mathematics* (New York: McGraw-Hill, 1940), p. 349. Allied attitudes are expressed, in a more philosophical context, in the two historical articles reprinted in Ernest Nagel, *Teleology Revisited* (New York: Columbia University Press, 1982).



by foundational tethers to more familiar mathematical territory. As Hilbert wrote Frege in a celebrated exchange:

Of course I must also be able to do as I please in the matter of positing characteristics; for as soon as I have posited an axiom, it will exist and be 'true' … If the arbitrarily posited axioms together with all their consequences do not contradict one another, then they are true and the things defined by these axioms exist. For me, this is the criterion of truth and existence.[39]

As a relative logicist, Frege would have been heartily opposed to 'formalism' of this ilk.

Despite these obvious differences in philosophical attitude, Frege's tone in his exchanges with Hilbert and Hilbert's amanuensis A. Korselt often seems excessively harsh, as if Frege were writing from a conservative and antiquated geometrical methodology that he does not adopt within his own mathematical work. Perhaps our ruminations on relative logicism suggest some deeper reasons for his unhappiness. In his 1898 lectures on geometry,[40] Hilbert claims that his attention to subgroups of axioms allows us to 'diagnose the structure of our spatial intuition'. he had in mind situations such as the following. A certain restricted group of non-metrical axioms

[39] E.-H. W. Kluge (ed.), *On the Foundations of Geometry and Formal Theories of Arithmetic* (New Haven and London: Yale University, 1971), p. 12.

[40] D. Hilbert, *David Hilbert's Lectures on the Foundations of Geometry 1891–1902*, ed. M. Hallett and U. Majer (Berlin: Springer, 2004). Hallett's editorial comments are particularly helpful.

$\mathcal{F}$ about points and lines *within 3D space* are sufficient to establish the 2D claim known as *Desargues' theorem*:

If two triangles are placed so that the straight lines connecting corresponding sides meet in a point, then the points of intersection of corresponding sides will lie upon a common line.

Essentially, the relevant proof proceeds by collapsing the 3D upper diagram of a triangle projected from one plane into the 2D situation displayed in the lower half of the illustration. But Hilbert proved the surprising fact that a purely 2D analog of the facts $\mathcal{F}$ could not logically force the truth of Desargues' theorem alone. Results of this type instantly made Hilbert's work widely celebrated and Frege's reluctance to extend him any credit for his discoveries seems quite uncharitable. Some of Frege's discomfort may trace to the fact that Hilbert's conception of 'logically forces' is tacitly first-order in nature (more or less), whereas Frege's basic approach to 'logic' tolerates the liberal invocation of extra 'logical objects', whether they arise as abstracted sets or as Plückerish recarvings. But from the recarving point of view, the dimensionality of the plane is not a fixed matter, for a plane will change its dimensions if it is carved into *circles* as its primitive elements rather than *points*. Starting within a 2D group of facts $\mathcal{F}$, it might be possible to devise extra 'logical objects' through recarving that will permit a reinstatement of the standard 3D proof of Desargues' theorem. From this point of view, Hilbert's claim that Desargues' theorem is 'logically independent' of the 2D $\mathcal{F}$ group may lack clear significance. In his final essay on geometry, Frege attempted to render greater justice to Hilbert's independence results. As various commentators have observed,[41] that essay articulates what is, in effect, a model-theoretical account of first-order logical consequence. This similarity does not show that Frege himself has adopted a modern 'semantic approach to logic'; it is more likely that he remained loyal to the nineteenth-century traditions which had assumed that 'logic' must somehow validate appeals to novel 'objects' of a 'direction of line L_0' variety.

To the modern reader, this old-fashioned appraisal of logic's 'creative' capabilities may seem startling, as we no longer expect that 'logic' alone can erect mighty layers of supplementary 'objects'

[41] William Demopoulos, 'Frege, Hilbert, and the conceptual structure of model theory', *History and Philosophy of Logic*, 15.2 (1994), pp. 211–25.

above an originally limited domain. But this limited 'first order view' of logic's capacities did not become canonical until the 1930s and Frege's underlying objections to Hilbert's point of view may trace tacitly to this divergence.

After 1904, Hilbert hoped that the consistency and completeness of his free-standing axiomatic schemes could sometimes be established by elementary means (otherwise, the direct construction of a suitable model was required).[42] But Kurt Gödel's famous second incompleteness theorem showed that the consistency of a sufficiently rich axiomatic system can be authenticated only if the consistency of some yet stronger theory is assumed. So the problems that originally bedevilled the naive 'free creativity' thesis return again: how can we ascertain that our 'free creativity' does not depend upon an inconsistently described structure? In light of the unavailability of elementary checks upon consistency, modern mathematical orthodoxy has settled upon the following resolution: *mathematics is free to study any subject that can be legitimated as a well-defined class within the set theoretical hierarchy*. This view can be called *set theoretic absolutism*, for it makes reduction to set theory the final arbiter of mathematical existence. Although such ontological absolutism represents 'official policy' today, some mathematicians and philosophers (who often don't like set theory much) harbour in their bosoms opinions closer to naive 'free creativity': *mathematics should be free to study the properties of any self-consistent, free-standing construct*. But we currently lack any well-developed philosophy of mathematics that can support this hope (which seems to rely upon an unsupported faith that Dirichlet Principle-like problems will never visit us again). In these respects, we are still confronted with the same task of reconciling 'safe procedure' with 'free creativity' that had troubled Frege and the other relative logicists of the nineteenth century.[43]

[42] Hilbert was never one-sided in his thinking and served, in fact, as a great advocate for algebraic construction in Dedekind's vein. He could have readily accepted that there might be natural mathematical objects ('differentiable manifolds') that can be readily constructed but are not easily captured within an axiomatic frame.

[43] First written 1998; rewritten 2008. Thanks to Jeremy Avigad, Bill Demopoulos, Michael Friedman, Jeremy Heis, Penelope Maddy, Tom Ricketts, Jamie Tappenden and Michael Thompson for their helpful comments.

MICHAEL HALLETT

11 Frege and Hilbert

INTRODUCTION

Between 1897 and 1902, there took place a brief correspondence between Frege and Hilbert, consisting of four letters from Frege, and two letters and three postcards from Hilbert.[1] It centres on Frege's reactions to Hilbert's classic *Grundlagen der Geometrie*, first published in 1899, and Hilbert's restatements in his letters to Frege of the foundational positions which that work, sometimes only implicitly, embodies.[2] Despite the obvious richness of common purpose between Frege and Hilbert, the correspondence is especially instructive because of the strong disagreements expressed. For example, the two disagreed on the form and function of definitions, the nature, purpose and formulation of axioms, the nature of (axiomatized) mathematical theories, the method of independence

[1] This is all the correspondence which is extant. It is published in Frege, *Wissenschaftlicher Briefwechsel*, ed. G. Gabriel, H. Hermes, F. Kambartel, F. Kaulbach, C. Thiel and A. Veraart (Hamburg: Felix Meiner, 1976), with English translations in *PMC*, E.-H. W. Kluge (ed.), *Gottlob Frege on the Foundations of Geometry and Formal Theories of Arithmetic* (New Haven and London: Yale University Press, 1971), and also M. Resnik, *Frege and the Philosophy of Mathematics* (Ithaca: Cornell University Press, 1980). I will refer to the letters just by dates; they can then be found easily in any of the works containing them.

[2] D. Hilbert, 'Grundlagen der Geometrie', in *Festschrift zur Feier der Enthüllung des Gauss-Weber-Denkmals in Göttingen* (Leipzig: B. G. Teubner, 1895). Hilbert's *Grundlagen* is henceforth cited by its famous nickname, the *Festschrift*, to distinguish it from Frege's *Grundlagen*. The *Festschrift* went through six further editions in Hilbert's lifetime, and has experienced eight more since Hilbert's death. The original edition has recently been republished as ch. 5 of *Hilbert's Lectures on the Foundations of Mathematics and Physics*, ed. M. Hallett and U. Majer, vol. I: *David Hilbert's Lectures on the Foundations of Geometry, 1891–1902* (Heidelberg, Berlin and New York: Springer, 2004).

proofs in geometry, the role and form of consistency proofs and the nature of mathematical existence. Many of the articles of disagreement, especially those on axioms and independence proofs, also reveal or underline significant differences in their respective conceptions of logic. Frege followed the correspondence with two polemical, and wider-ranging, articles[3] on similar or related themes, Hilbert himself having apparently declined Frege's suggestion that their exchange of views be published.[4] These two papers help to fill out the picture on Frege's side, first by restating Frege's opposition, and then by presenting his insights into the formal structure of Hilbert's position. Especially important are Frege's attempts in his second article to render central results of Hilbert's project as read through his own system.

The issue of who was right and who wrong is a complex one, and it is not the purpose of this paper to attempt any final judgement; rather, its point is to discuss one of the central disagreements, namely the importance (as Frege saw it) of fixed reference. At the root of Hilbert's foundational investigation was a way of regarding mathematics which distanced mathematical theories themselves from fixed interpretations of them, an approach which Frege found incoherent. One of the points I will try to bring out is that Frege struggled with what he saw as philosophical difficulties which others, e.g., Dedekind and Hilbert, saw as simply intrinsic to the nature of mathematics. Hilbert not only absorbed this point, but turned this 'difficulty' into a powerful methodological tool, and, in doing so, effected a transformation in the conception of mathematics.

This, of course, is not the whole story. On many points of formulation, Frege was clearly in the right, and in effect pushed Hilbert for answers, not least on questions of logical formulation, answers which were not forthcoming until much later, if at all. And just as

[3] Frege, 'Über die Grundlagen der Geometrie', *Jahresbericht der deutschen Mathematiker-Vereinigung*, 12 (1903), pp. 319–24, 368–75, reprinted in Frege, *Kleine Schriften*, ed. Ignacio Angelelli (Hildesheim: Georg Olms Verlag, 1967), pp. 262–72; and Frege,'Über die Grundlagen der Geometrie', *Jahresbericht der deutschen Mathematiker-Vereinigung*, 15 (1906), pp. 293–309, 377–403, 423–30, reprinted in *Kleine Schriften*, pp. 281–323. Hereafter GG 1903 and GG 1906 respectively.

[4] See GG 1903, p. 319.

importantly, Frege's later considered reactions to Hilbert's study of, and approach to, geometry served to emphasize formal elements in his own conception of logic which make it seem a lot closer to the modern conception than might at first be thought. But these must be topics for another paper.

1 DEFINITIONS AND THE FIXING OF REFERENCE

1.1 The purpose of definitions for Frege

Frege takes over 'the traditional notion of axioms' (see, e.g., GG 1906, p. 295); to employ his mature terminology, axioms are true Thoughts, and hence must have a determinate sense and a determinate reference (truth-value). According to the doctrine enunciated in 'On sense and reference', sentences/propositions express Thoughts, and a sentence/proposition can only have a truth-value if all the terms in it have a fixed reference. Thus, as he says to Hilbert, the 'axioms, basic laws, theorems'

> ... ought to contain no word and no sign whose sense and reference or whose contribution to the Thought expressed does not already stand fully fixed, so that there is no doubt as to the sense of the proposition, of the Thought therein expressed. It can only be a question of whether the Thought is true, and then on what its truth rests. (Frege to Hilbert, 27 December 1899.)[5]

Axioms and basic laws are distinguished from theorems in that they neither can be proved nor are in need of proof;[6] conversely, whatever can be proved from more fundamental principles should not be taken as an axiom.[7] Intrinsic to the foundational project for Frege (certainly in his treatment of arithmetic) is the search for, and establishment of, the ultimate ground on which the given truths rest. Frege seems to have thought that there are really only two fundamental sources of truth in mathematics: logical truth, which

[5] As is the case with all the non-English quotations given here, the translations are my own unless otherwise explicitly noted.

[6] See *Gl*, end of §3.

[7] Though see M. Dummett, 'Frege and the consistency of mathematical theories', in *Frege and Other Philosophers* (Oxford: Clarendon Press, 1991), p. 10.

grounds arithmetic in the widest sense, and geometrical intuition, which grounds geometry.[8] As Frege says to Hilbert, the axioms of geometry

> are true, and however are not proved because the knowledge of them flows from a completely different source from that of logic, a source which one can call intuition of space. (Frege to Hilbert, 27 December 1899)

Whatever the source of truth of the primitive principles, the logical system will ensure that truth flows down to all the theorems.

Definitions play an important role in this. Their first function is to provide a reference for terms which apparently have none:

> I divide the totality of mathematical propositions into definitions and the remainder (axioms, basic laws, theorems). Every definition contains a sign (an expression, a word) that before had no reference, and which acquires one only through the definition. (Frege to Hilbert, 27 December 1899.)

After the definition has been set down, there corresponds to it a trivial truth (in the language expanded to contain the previously empty sign), i.e., the proposition expressing the equivalence between the sign defined and the expression giving its sense and reference.[9]

However, the stress on reference fixing alone does not fully capture the importance of definitions in Frege's project. In his late 'Logic in mathematics',[10] Frege distinguishes two kinds of definitions. The first is what he calls 'constructive definition', where a new sign is introduced, and given sense and reference, by stipulation. This kind of definition is strictly speaking unnecessary, but in practice indispensable. The second kind is what Frege calls 'analytical definition', where the meaning of a sign with 'a long-established use' is analysed and this meaning is (re)constituted from simpler (known) meanings. Frege then remarks that it would perhaps be better not to call analytic definitions definitions at all, since the equivalence stated has more the nature of an axiom than

[8] Incidentally, Frege never challenges Hilbert's choice of axioms, but rather their formulation, what they can be considered as saying.

[9] See *Gl*, §67.

[10] Frege, 'Logik in der Mathematik', in *Nachgelassene Schriften*, ed. H. Hermes *et al.* (Hamburg: Felix Meiner, 1969), pp. 219–70; see also E. Reck and S. Awodey, *Frege's Lectures on Logic: Carnap's Student Notes, 1910–14* (LaSalle, Ill.: Open Court Publishing Company, 2005).

an 'arbitrary stipulation'. Moreover, where analysis furnishes a clear sense in place of an uncertain, vague or informal one, the established sign can be regarded *de facto* as new, whereby the 'analytical definition' becomes a constructive one.[11] From what was said above about the nature of the Fregean foundational project, it is clear that analytical definitions have an essential purpose, for they show how to render provable central propositions which before the analysis, and subsequent definition, were not so. Hence, they form a central part of the search for a genuinely fundamental level. This kind of definition corresponds to what Frege in the *Grundlagen* calls 'fruitful definitions'.[12] The two sorts of definition go hand in hand. In GG 1906, p. 303, Frege writes:

> The proper meaningfulness of definitions lies in the logical construction out of urelements. And because of this, one could not dispense with them even in such a case [where mutual understanding is guaranteed]. The insight into the logical structure which definitions afford is not only in itself valuable, but it is also a condition for the insight into the logical linkage of truths ... The intellectual activity which leads to the setting-up of a definition can be one of two kinds, analytical or constructive, just like the activity of the chemist, who either analyses a substance into its elements or combines given elements into a new substance. In either case, one sees the synthesis of a substance. Likewise, one can also here achieve something new through logical construction, and set down a sign for this.

Examples of the employment of analytical definitions can be found in Dedekind's work,[13] and, unsurprisingly, their role is canonically illustrated by Frege's *Grundlagen*, where analysis succeeded by constructive definition is at the heart of the procedure. The concepts of following in an *R*-series for any relation *R* (the *R*-ancestral), number and successor generally, some of the individual numbers,

[11] See Frege, 'Logik in der Mathematik'. For discussion, see M. Dummett, 'Frege and the paradox of analysis' in *Frege and Other Philosophers*, pp. 17–52; E. Picardi, 'Frege on definition and logical proof', in *Atti del Congresso Temi e Prospettive della logica e della filosofia della scienza contemporanee. Cesean 7–10 gennaio 1987* (Bologna: CLUEB, 1988), vol. I, pp. 227–30.

[12] See *Gl,* p. ix, and then §4.

[13] R. Dedekind, *Stetigkeit und irrationale Zahlen* (Braunschweig: Vieweg und Sohn, 1872), translated in W. Ewald (ed.), *From Kant to Hilbert*, 2 vols. (Oxford: Oxford University Press, 1996), pp. 765–79; and Dedekind, *Was sind und was sollen die Zahlen?* (Braunschweig: Vieweg und Sohn, 1969), translated in W. Ewald (ed.), *From Kant to Hilbert*, pp. 787–833.

then natural number generally, are all defined after suitable analysis. The basic arithmetical laws (essentially, the second-order Peano axioms) are then proved (in second-order logic), making fundamental use of Hume's Principle (*HP*):

$$\forall F,G[NxF = NxG \leftrightarrow F \approx \mathrm{G}] \qquad (HP)$$

where '*NxF*' is a term-forming operator which takes any sortal concept *F* and assigns it an object, its number, and where '$F \approx G$' is a *defined* notion. The conceptual role of *HP* is discussed in the next section.

The epistemological importance of 'fruitful definitions' for Frege is therefore clear, given their obvious role in revealing genuine axioms and genuine (direct) denotations. Definitions construct denotations out of urelements, and therefore they ensure that the only objects referred to in the (reconstructed) theory are those explicitly dealt with or introduced by the assumptions at the outset, in the process revealing logical path to the fundamental assumptions and thus 'the logical linkage of truths'. It follows that the semantic role of definitions is fundamental, for in deriving the basic principles of arithmetic from purely logical ones, the fixity of reference secured through definitions is absolutely necessary, since reference must be secured to the right sort of things. Clearly there is an element of convention as to the choice of objects (primitives) to be used in the definitions, particularly the definition of number. Extensions are primitive in the *Grundlagen*, though value-ranges of functions are taken as primitive in the *Grundgesetze*. The important things, though, are: (i) what Frege sees as *logical objects* must be chosen; and (ii) reference is fixed to those objects. We will return to these points. At the moment it is important to realize why the reference-fixing role of definitions is stressed so heavily in Frege's letter to Hilbert.

In his *Grundgesetze*, Frege devoted considerable attention to a careful statement of what is required of correct definition. In the first volume,[14] there is a short section (§33), referred to in Frege's letter to Hilbert of 27 December 1899, which sets out in sober fashion various criteria for proper definition. This section concerns, not definition in the wider sense, but rather what is permissible as correct

[14] *Gg*, vol. I.

definition with respect to the fundamentals of the system set out in this first volume. However, the second volume[15] contains a more expansive, and certainly more polemical, discussion of definition. (The key sections are §§56–67.) Here the emphasis is on definition of 'sharply delimited concepts', and indeed §56 begins:

> A definition of a concept (possible predicate) must be complete; it must unambiguously determine, for every object, whether it falls under the concept (whether the predicate can be truly asserted of it) or not. There can be no object for which, according to the definition, it remains doubtful whether it falls under the concept, even though it may not always be possible for us humans with our defective knowledge to decide the question.

Unique definitions are important, to avoid ambiguity of reference; and Frege also declares firmly against partial or piecemeal definitions. Frege sees the gradual widening in the nineteenth century of the number concept, and the subsequent redefinitions, as running counter to these strictures. He admits that the 'scientific progress' which was behind such widening rendered piecemeal definitions perhaps 'unavoidably necessary'. But he goes on to say that, properly speaking, old signs or terms should have been supplemented or replaced by new ones, instead of the use of the old signs to serve a new, divided purpose, and indeed 'logic demands this', the reluctance to introduce new signs being 'the cause of many unclarities in mathematics' (§58). He goes on:

> All the more must it be stressed that logic cannot recognise as concepts conceptually similar formations which are still in flux, and which have not yet received final and sharp boundaries.

A second (putative) definition of the same term either draws the *same* boundaries, in which case we do not have a new *definition*, but rather an assertion which has to be proved, or it does not, in which case we did not have a definition in the first place. Moreover, these piecemeal definitions are never final ('for who can know if we have arrived [with such a definition] at a finished proposition?'), and, Frege says:

> Without final definitions, one does not have final theorems. We never emerge from vacillation and unfinishedness. (§61)

[15] *Gg*, vol. II.

If it is never clearly fixed exactly what is being talked about, then it is simply not clear what is being asserted. For Frege,

> [In] mathematics, a word without fixed meaning has no meaning at all. (GG 1906, p. 303)

He also argues against 'conditional definitions', for example, a definition of the form 'If *a*,*b* are numbers, then "*a* + *b*" denotes …'. Frege states:

> But the addition sign is only defined when the meaning of every possible combination of signs of the form '*a* + *b*' is determined, whatever meaningful proper names one sets in for '*a*' and '*b*'.[16]

In other words, under conditional definitions, generally clear answers to arbitrary questions of identity ('Does $a = b$?') will not be theoretically available, and the central condition for defining a concept, namely that the definition assigns 'sharp boundaries' to the concept, cannot be fulfilled.

The concern with proper definition which is expressed so forcefully in the second volume of the *Grundgesetze* of 1903 is in fact clearly foreshadowed in the philosophical difficulties with referential indeterminacy with which Frege grapples in his *Grundlagen*, and these in turn explain much of Frege's central criticisms of Hilbert's *Festschrift*. Let us turn now to Frege's work.

1.2 *Failing to define the concept of number*

The project pursued in the *Grundlagen* was to show that arithmetic is *analytic* by showing how it can be deduced from second-order logical laws and definitions. Frege's analysis of arithmetical statements results in the claims that numbering (e.g., in the assertion of certain canonical numerical statements such as 'Jupiter has four (Galilean) moons') consists in the assignment of a number-object (say four) to a sortal concept ('Galilean moon of Jupiter'). It follows that there must be a concept, Number, which has the numbers falling under it, and that there must also be a general principle governing

[16] *Ibid.*, §65.

the association of number-objects with concepts. Frege isolates *HP* (pp. 4–18) as that general principle, but what about the concept of Number with which *HP* is implicitly concerned? As Frege says (*Grundlagen*, §4), in the attempt to reduce the fundamental propositions of number theory to general logical laws,

> above all it is the concept of number which must be either defined or recognized as indefinable, and that is the central problem of this book. On its solution rests the decision as to the nature of arithmetical laws.[17]

Frege dismisses the idea that the numbers can be 'given to us' directly as individuals through any special intuition. We have therefore the question (§62):

> How, then, are numbers to be given to us, if we cannot have any ideas or intuitions of them?

Frege goes on immediately to reiterate the famous Context Principle, previously set down in the Introduction (p. x):

> Only in the context of a sentence does a word mean anything. Thus it will be a matter of defining the sense of a proposition in which a number word appears. (§62)

It is precisely at *this* point in the *Grundlagen* that Frege suggests (for discussion) *HP* as a means of following the Context Principle for the number terms and thus indirectly specifying the concept Number. Following Frege's analysis, 'number words' must stand for 'self-subsistent objects', and this means that a crucial type of sentence involving number words must have a sense, the 'recognition' or identity sentences. But this use of *HP* 'raises at once certain logical doubts and difficulties, which ought not to be passed over without examination' (§63). The central doubts are well known. The notion of identity *HP* uses is the *general* notion of identity which (for Frege) can be applied to *all* objects. This means that the object terms '*NxF*' are to be considered not only in pairs flanking an identity, as in '*NxF* = *NxG*', but also in combination with

[17] In §21, Frege states part of his task to be 'to assign to Number its proper place among our concepts!'

terms for objects otherwise given, as in '$NxF = a$'. And for these, *HP* assigns no truth-conditions. This gives rise to the famous 'Julius Caesar Problem'. But what this amounts to, as will become clear from the later discussion of definitions in the *Grundgesetze*, is that the *concept* of number is not properly fixed by *HP* since its extension is not fixed.

Frege's reaction in the *Grundlagen* confirms this. With respect to the analogy he develops for directions and straight lines, he says:

> Naturally no one will mistake England for the direction of the Earth's axis; but that is not due to our definition. That says nothing as to whether the sentence
>
> 'the direction of *a* is identical with *q*'
>
> is to be affirmed or denied if *q* is not given in the form 'the direction of *b*'. What we lack is the concept of direction; for if we had this, then we could determine that: if *q* is not a direction, our proposition is to be denied; if *q* is a direction, our original definition will decide. (§66.)

So, in the analogous case, *we lack the concept of number*, which means that *HP* fails to introduce that concept, to define it. The problem is not stated in quite this way in the *Grundlagen*, but it amounts to this, since the question of whether the (extension of the) concept Number is definite (has sharp boundaries) is just the question of whether we can assign a truth-value to any statement of the form $\exists F[a = NxFx]$ whatever object *a* is. Frege's statement of the Julius Caesar problem is just the claim that *HP* alone fails to assign these truth-values.

There is a second aspect of referential indeterminacy relevant here. The difficulty is hard to state, but Frege seems nevertheless moved by it. Even if the extension of the concept of number had sharp boundaries, it is not possible to tell from *HP* alone which among the objects the numbers are, i.e., what distinguishes, characterizes a number object from among all other objects. *HP* itself says nothing about this. This fact is brought out starkly by Boolos's reformulation[18] of *HP* as

$$\exists f \forall F,G[f(F) = f(G) \leftrightarrow F \approx G] \qquad (HP\text{ function})$$

[18] See G. Boolos, 'Is Hume's principle analytic?', in R. Heck (ed.), *Language, Thought and Logic: Essays in Honour of Michael Dummett* (New York and Oxford: Oxford University Press, 1997), pp. 245–61, here pp. 253–4.

where 'f' stands for a function from concepts to objects, or even Boolos's equivalent principle *Numbers*

$$\forall F \exists ! x \forall G [G \eta x \leftrightarrow F \approx G] \qquad (Numbers)$$

both of which involve an explicit pure existence assertion instead of a term-forming operator. It is clear from this that nothing is said about what kind of object satisfies the function or object quantifier.

The point is presaged at the very beginning of the *Grundlagen*. In the opening sentence, Frege raises the question of what the number one is, and supposes that we normally get as answer, 'Why, a thing'. He dismisses this answer, among other reasons, since

> it only assigns the number one to the class of things, but does not state which thing it is.

Having pointed this out, we will perhaps then 'be invited to select some thing or other that will be called one'. Frege goes on immediately:

> Yet if anyone had the right to understand by this name whatever he pleased, then the same proposition about one would mean different things for different people; there would be no common content for such propositions. (Introduction, p. i)

The same point surely holds on a much grander scale for *HP*; people could take different objects as the numbers and yet agree on all the fundamental truths. We therefore have a second sense in which *HP* generates a problem of indeterminacy. For Frege, of course, this is more serious than mere indeterminacy. If *HP* were taken as a primitive truth (axiom) and yet yields no direct knowledge of the kinds of things numbers are, then we have no guarantee that the numbers are 'logical' objects and arithmetic analytic, and not just a 'special science', like hydrodynamics or atomic theory. This is a difficult and murky point in Frege, and sometimes the difficulty seems to slide over into the Caesar problem. Nevertheless, there is a real question here. In short, from Frege's point of view, *HP* cannot be an adequate *definition* of numbers and hence not of the concept of Number. Frege's central objections to Hilbert's claim to have given *definitions* of the geometrical primitives via his axioms are in fact very similar to these objections to the use of his own putative axiom, *HP*, as a way of 'introducing' numbers.

1.3 Frege's criticisms of Hilbert on definitions and axioms

Hilbert's *Festschrift* begins by stating that '[w]e think of three different systems of things' which are designated 'points', 'lines', 'planes', and these 'things' are taken to be in certain relationships to one another designated by the terms 'lie on', 'between', 'parallel to' and 'congruent' (of line segments), etc. It is the axioms which give the 'complete and precise description' of these relationships (see p. 4). Here 'precise and complete description of these relations' is not meant to be a claim that the axiom system is complete in the sense in which we now understand that term, but rather it merely serves to state that, as far as operation with the theory goes, the properties of, and relations between, the primitives are just those stated in, or derived from, the axioms. The axioms come in five groups: incidence (I); order (II); (Euclidean) Parallels (III) (before and after the first edition of the *Festschrift*, IV); congruence (IV) (resp. III); and continuity (V). (In the first edition, there is just one continuity axiom, the Archimedean Axiom; in subsequent editions, Hilbert's Completeness Axiom is added to this.) Each group supposedly 'expresses certain basic and connected facts of our intuition' (p. 4). Hilbert goes further, claiming that the axioms of II (order) 'define the [primitive] concept "between"' (p. 6), and those of IV 'define' the notion of congruence (p. 10). Hilbert's use of the term 'define' in this context was a source of great agitation for Frege. But, before we come to this, let us try to make it a little clearer what is behind Hilbert's deliberate, if easily misunderstood, choice of term.

The *Festschrift* was immediately preceded by an extensive course of lectures on Euclidean geometry, given in Göttingen in 1898/9. Notes for these lectures exist both in Hilbert's own hand and also as an official record of the lectures of textbook quality prepared by Hans von Schaper.[19] In his own notes for the 1898/9 course, Hilbert

[19] Both are published in full in ch. 4 of Hallett and Majer (eds.), *Lectures*. The official record was 'published' in seventy copies in March 1899, and the *Festschrift* itself in June 1899. The former was prepared largely for the use of the students in the course, but it was widely distributed. Frege certainly saw it, indeed Heinrich Liebmann's copy, for in a letter to Liebmann of 29 July 1900, Frege thanks him for the loan of the work, and says that he is returning it. Although the material of the lectures and that of the book are very closely related, there are striking differences

says clearly that in his axiomatisation of geometry, the primitive terms 'point', 'line', 'plane', 'between', 'congruent', and so on at first have no meaning at all in themselves. As he puts it:

We take *points*, *lines* and *planes* as elements. Thus, there is a system of things which we call *points* and which we denote by *A*, *B*, *C*, … , and *another* and respectively a *third* system of things which we call *lines* (*a*, *b*, *c*, …) and *planes* (α, β, γ, …). Points, lines, planes are just terms for things; we associate with these no intuitions and no further properties. System given, i.e., one can distinguish each one from the others $A \neq B$.[20]

Thus, it is just postulated that there exist domains ('systems') for the different sorts of thing,[21] and beyond this Hilbert is clear that the axioms, and the axioms *alone*, assign whatever properties are to be assumed to be possessed by the things of the 'systems', and whatever relationships there are between the systems and their members. In the official record, Hilbert states:

We should not allow ourselves to be misled by the names ['points', 'lines', 'planes'] chosen, and ascribe to these things geometrical properties ordinarily associated with them. *At this juncture, all we know is that each thing of one system is different from every thing of the two other systems. These things obtain all their other properties from the axioms.*[22]

It is in just this context that Hilbert makes the following remark in the Introduction to his lectures in 1898:

I want to stress the main barrier to th[e] comprehension [of the lectures].

between the two, not least since the lectures are more discursive than the monograph, and a great deal more expansive philosophically. For further details, see my Introduction to the lectures in *ibid.*, ch. 4, and also that to ch. 5, which contains the 1899 edition of the *Festschrift*. It should be emphasized that, despite important differences, the lectures and the monograph represent essentially identical points of view as far as the foundations of geometry are concerned.

20 D. Hilbert, *Grundlagen der Euklidischen Geometrie*, lecture notes for a course held in the Wintersemester of 1898/9 at the Georg-August Universität, Göttingen (Göttingen Niedersächsische Staats- und Universitätsbiblithek, 1898/9), p. 9, in *Lectures*, p. 224.

21 See also, e.g., D. Hilbert and P. Bernays, *Grundlagen der Mathematik*, 1st edn (Berlin: Julius Springer, 1935), vol. I, pp. 1–2.

22 D. Hilbert, *Elemente der Euklidischen Geometrie*, *Ausarbeitung* by Hans von Schaper of Hilbert's lecture notes 1898/9 (see note 21 above) (Göttingen: Niedersächsische Staats- und Universitätsbibliothek and the Mathematisches Institut of the Georg-August Universität, 1899, pp. 1–2, in *Lectures*, p. 303.

It takes some effort and watchfulness to abstract constantly from things, ideas and intuitions with which one is familiar, and to set oneself back in a state of ignorance. To subject oneself to this effort is, however, easier, when one clearly recognizes the purpose.[23]

Thus, when using the word 'point', one must not associate with it the ordinary, informal connotations that this term carries: the *only* properties and relations which count are those specified by the axioms. This is the reason why Hilbert calls his axioms 'definitions' of the notions, because if anything is to be credited with assigning 'characteristics' to the primitives, it is these axioms.[24] As Hilbert says to Frege:

The definitions (i.e., explanations, definitions, axioms) must contain everything, but this said should contain only that which is required for the construction of the theory. With respect to my division into explanations, definitions, axioms, which together make up the definitions in your sense, these certainly contain much that is arbitrary. Nevertheless, I believe that, in general, my ordering is serviceable and perspicuous.[25]

[23] Hilbert, *Grundlagen der Euklidischen Geometrie* (1898/9), p. 7, in Hallett and Majer, *Lectures*, p. 223. See also M. Pasch, *Vorlesungen über neuere Geometrie* (Leipzig: Teubner, 1882), p. 3, where a similar statement appears.

[24] It is to be noted that there is another sense in which some of Hilbert's axioms are 'definitions'. Take Axiom I 5 of the first edition of the *Festschrift* (I 6 after). This says that if two points of a straight line lie in a given plane, then every point of that line lies in the plane. This follows directly one of the traditional 'definitions' of what a plane is, namely a plane is a surface such that any straight line joining two points of the surface lies wholly within the surface. (See Euclid's Definition 7 in Book 1 of the *Elements*, in T. L. Heath, *The Thirteen Books of Euclid's Elements*, 3 vols, 2nd edn (Cambridge: Cambridge University Press, 1925), vol. I, pp. 153, 171. Gauß was greatly occupied with this definition; see Hilbert's Remark [3] to his own copy of the official record of the 1898/9 lectures, and my editorial comments on it in Hallett and Majer (eds.), *Lectures*, p. 397. It is interesting to note that Poincaré states:

Sometimes one defines the plane in the following manner:

The plane is a surface such that all the points of the straight line joining any two of its points is always entirely on this surface.

This definition manifestly hides a new axiom (H. Poincaré, 'Les géométries non euclidiennes', Revue général des sciences pures et appliqués, 2 (1891), pp. 769–74, at p. 772).

[25] Hilbert to Frege, 29 December 1899 (II). The original of this letter is not extant, though there are two reports of its contents, a partial copy in Frege's hand, and what the editors of Frege's correspondence call a 'concept or partial excerpt' made by Hilbert. See Frege, *Wissentschaftlicher Briefwechsel*, XX–XXI, p. 65. The two reports differ slightly; they will be denoted here by '(I)' and '(II)' respectively.

In particular, it follows from this that proofs in the system can only draw on the axioms (or previously proved theorems) in constructing a new inference; indeed it would make no sense to do otherwise, since no other source of knowledge is allowed.

In the *Festschrift*, Hilbert only says that certain of the Axiom Groups 'define' certain of the primitives. But in his correspondence with Frege, Hilbert states clearly that the axioms 'define' all the primitives and moreover that *all* the axioms are involved in the 'definitions', and this is said while protesting about the impossibility of giving definitions of the kind that Frege recognizes:

> On the contrary, to wish to give a definition of point in 3 lines is, in my view, impossible, since the whole construction of the axioms gives the complete definition. Every axiom contributes something to the definition, and every new axiom thus alters the concept. 'Point' in Euclidean, non-Euclidean, Archimedean, non-Archimedean geometry is each time something distinct.[26]

Thus, the concepts cannot be treated individually; what can be asserted about points, say, is affected by what can be asserted about lines, planes, betweenness, congruence and continuity, in other words, by the totality of the axioms. In a subsequent letter, the holism Hilbert expresses here is even more explicit:

> My view is just this, that a concept can only be logically fixed through its relations to other concepts. These relations, formulated in definite statements, I call axioms, and thus I arrive at the view that these axioms (perhaps with the addition of names for the concepts) are the definitions of these concepts. I have not just dreamed up this view recently; rather I saw myself forced to it by the demands of rigour in logical argument, and by the logical construction of a theory. I came to the conviction that in mathematics and in natural science this is the only way to deal with subtler things with certainty, since otherwise one just goes round in circles.[27]

In short, Hilbert expresses the view that the geometrical axioms are what determine the extension of the concepts 'point', 'line', 'plane', the relations 'between', 'congruent to', and so on, in so far as anything determines these.

[26] Hilbert to Frege, 29 December 1899 (I).
[27] Hilbert to Frege, 22 September 1900.

It is important to bear in mind the comparison with Frege's *HP* and its treatment of the numbers, except that with Hilbert there are several concepts involved, not just one. Not surprisingly, Frege does not accept Hilbert's procedure.

First, the definitions given by Hilbert are bound to be *conditional*; one of the (several) definitions of 'point' can only be upheld under the condition that the axioms specified hold, and Hilbert makes it clear that, if different axioms are stated, different concepts of point will result. Hence, the concept of point must be relative to the axioms chosen. Secondly, it follows from this that Hilbert's definitions will be *piecemeal*; there will not be one overriding definition of 'point' which can be appealed to. The result is that there can never be a decision (in principle) of the question of whether an alleged point is the same as some other object; and no unambiguous answer to certain questions, for example, the question of whether, given a straight line and a point outside it, there is a unique straight line through that point parallel to the given line. Thirdly, the axioms *do not fix the reference* of the basic terms. Hence, they cannot be proper definitions for Frege; and this is the point which is most serious for him.

Let us look at Frege's objections in more detail.

Hilbert's 'definitional' procedure must have seemed to Frege hopelessly confused. To say that definitions are given by laying down axioms is clearly to get the semantic cart before the semantic horse. For Frege, as we have seen, axioms (like basic laws) must be truths. Thus, before one can declare something to express an axiom, all the terms in it must have a determinate meaning, and consequently one must already have performed any defining there is to be done. As Frege puts it with respect to the terms 'point' and 'between':

> If I were to set up your Axiom II 1[28] as an axiom, then I assume in doing so that the meanings of the expressions 'something is a point on a straight line' and '*B* lies between *A* and *C*' are completely and unambiguously known.[29]

Moreover, it is not formally clear how Hilbert's axioms *could* be definitions in the *Grundgesetze* sense, since the very terms to be defined ('point', 'line', 'plane' and so on) appear in the axioms, i.e.,

[28] II 1 says that if *A*,*B*,*C* are any three points of a straight line, and *B* is between *A* and *C*, then *B* is also between *C* and *A*.

[29] Frege to Hilbert, 27 December 1899.

the statements doing the defining. (Note that *HP* does not itself contain the concept term 'Number', which allows one to think of it as somehow 'creative', 'dividing up the content in a new way', as Frege says in the *Grundlagen*, §64.) Moreover, as Frege several times complains, assuming that Hilbert does succeed in defining 'point', then it simply is not clear whether he has defined a concept of the first-level, under which objects fall, or one of the second-level, under which first-level concepts fall.[30]

Frege's objection can be simply put. A proper definition of the concept 'point' ought to take the form

$$P(x) \equiv_{\mathrm{df}} \cdots x \cdots$$

where the concept being defined, '*P*', does not appear on the right. But then the axioms cannot define 'point', for the definition would be of the form:

$$P(x) \equiv_{\mathrm{df}} A_1 \wedge A_2 \wedge \cdots \wedge A_n$$

where $A_1, A_2, \cdots, A_n$ is a list of some or all of the axioms. This cannot be formally correct, since there is no free object variable on the right, and furthermore the procedure is clearly circular, since the axioms already employ the predicate 'point' (and the other primitives). The axioms on the other hand do contain what Frege would call second-level concepts, for some of them attempt to express relations between the (first-level) concept 'point' and the (first-level) concept 'line'. Thus, whatever is hereby defined cannot be a concept of the first level. The same holds for all the primitive concepts and relations, and thus also for all the terms which are (properly) defined by using them in explicit definition.[31] That Hilbert's 'definitions' are conditional ones is recognized by Frege in his letter of 6 January 1900. He thanks Hilbert for sending a copy of his lecture on the real numbers delivered in Munich in 1899.[32] He says:

From your Munich lecture, I believe I have recognized your plan still more clearly … It seems to me that you want to separate geometry completely from intuition of space, and make it a purely logical science like

[30] See Frege's letter of 6 January 1900 to Hilbert, and then GG 1903, p. 374. See also Frege's letter to Liebmann of 29 July 1900.

[31] Frege gives a full analysis of this in GG 1903.

[32] Published as Hilbert, 'Über den Zahlbegriff', *Jahresbericht der deutschen Mathematiker-Vereinigung*, 8 (1900a), pp. 180–5.

arithmetic. The axioms, which otherwise ought to be guaranteed through intuition of space, and laid at the foundation of the whole structure, are now, if I understand you aright, to be carried as conditions in every theorem, not indeed fully expressed, but rather contained in the words 'point', 'line', etc.[33]

And of course the axioms differ from system to system, even from presentation to presentation.

In sum, combining talk of definition with talk of axioms only introduces confusion.[34] To Liebmann, Frege was very forthright:

The axioms supposedly constitute the sole determination of the concept. But here we have the monstrosity that not one concept, but three (point, straight line, plane), are supposedly defined in this single definition, which stretches over a whole signature ... [The axioms] are supposed to help define, e.g., the concept of straight line, and at the same time, the term 'straight line' appears in those axioms, and not just this, but also 'point' and 'plane', which themselves are to be defined.[35]

Frege's main complaint, though, is that Hilbert's definitional procedures do not yield fixity of reference. For one thing, Frege says that in 'defining', Hilbert lays down no criteria, no characteristics which allow one to tell whether a given object or relation is of the right kind. For example, about the treatment of 'between', he quotes what was said to him about this by his colleague Thomae:

'That is no definition, since no characteristic is given through which it can be recognized whether the relation between holds or not'.[36]

He adds his own approval: 'I, too, cannot regard it as a definition'. To Frege's complaint about laying down 'characteristics', Hilbert replies that it is just a mater of 'taste',[37] and that one could easily modify his axioms to allow in the specification of characteristics: 'As far as I am concerned, you could say "characteristics" instead of "*axiom*".'[38] But this clearly does not resolve the problem; in a subsequent letter, Frege says:

[33] Frege to Hilbert, 6 January 1900.
[34] Frege to Hilbert, 27 December 1899.
[35] Frege to Liebmann, 29 July 1900.
[36] Frege to Hilbert, 27 December 1899.
[37] Hilbert to Frege, 29 December 1899 (I).
[38] Hilbert to Frege, 29 December 1899 (II).

I have no idea how, with your definitions, I could decide the matter of whether my pocket-watch is a point or not. Even the first axiom[39] deals with two points. Thus, if I wished to know whether it holds of my pocket-watch, I must first of all know of some other object that it is a point. However, even if I knew, for example, that my fountain-pen is a point, I still could not decide whether my pocket-watch and my fountain-pen together determine a straight line, since I do not know what a straight line is.[40]

As Frege says subsequently:

We get no further by means of this axiom, and so it is with all the axioms. When we arrive at the last, we still do not know whether these axioms hold of my pocket-watch in such a way that we are justified in calling it a point. (GG 1903, p. 370)

This amounts to saying that Hilbert's 'definitions' do not not lay down necessary and sufficient conditions for membership in the extension of the concept being defined; they do not specify 'sharp boundaries'. The pocket-watch example is nothing other than the Caesar problem: is Julius Caesar/my pocket-watch a point or not? Hilbert's definitional procedure is not in a position to tell us.

The other indeterminacy problem is also pointed out by Frege: it seems that not enough is said to fasten on one extension as the proper extension of the concept allegedly being defined. Frege takes again the example of Hilbert's treatment of the term 'point':

One is left in the dark about what you call a point. Initially, one imagines that point is meant in the sense of Euclidean geometry, a conviction which is strengthened through the statement that the axioms express basic facts of our intuition.[41]

But then, continues Frege, on p. 20 of the *Festschrift*, a point is taken to be a pair of numbers taken from a Pythagorean field Ω,[42] so the term 'point' now has a meaning different from its intuitive, Euclidean meaning. There is thus, from Frege's point of view, careless ambiguity; in particular, the axioms are no longer what they

[39] Hilbert's Axiom I 1 says that any two distinct points determine a straight line.

[40] Frege to Hilbert, 6 January 1900.

[41] Frege to Hilbert, 27 December 1899.

[42] Ω is a minimal Pythagorean sub-field of the reals, which Hilbert uses to construct a model of the plane part of his whole axiom system. Ω is countable, but recall that there is no completeness axiom in the first edition (1899) of Hilbert's *Festschrift*.

initially seemed to be, namely Euclidean truths, as is suggested (at least to Frege) by Hilbert's claim that the axiom groups express 'certain basic and connected facts of our intuition'. They do not fix unambiguously the reference of the primitive terms, and without fixity of reference, they are not even truths. Indeed, '[f]or mathematics, a word without fixed meaning has no meaning at all' (above, p. 420), and in the same essay, Frege remarks:

> When something expresses now this Thought, now that, then in truth it expresses no Thought at all. (GG 1906, p. 424)

Frege's evident frustration with Hilbert's *Festschrift* is more than understandable. The lack of a clear reference for the primitives means that Hilbert's axioms cannot be true Thoughts, and are thus quite the wrong thing to take as the fundamental propositions of geometry. Crucially, we see that at least some of Frege's dissatisfaction with Hilbert mirrors quite closely his own attempts in the *Grundlagen* to introduce the concept of Number using the 'axiom' *HP*.

Let us return to Frege and the failure of *HP* as a form of definition. As he says in the *Grundlagen* (§68):

> Since we cannot in this way [i.e., by using *HP*] achieve a sharply limited concept of direction, and, for the same reasons, not one of number, we shall attempt a different path.

What lies on Frege's 'other path' is explicit definition: the number term *NxF* is now defined as the *extension* of the higher-order concept '*X* is equinumerous with *F*'. This, of course, is provisional, for the *Grundlagen* has no theoretical treatment of the notion of extension. For this reason, among others, the *Grundlagen* (as Frege openly admits, §90) is incomplete, and its central philosophical goal consequently unachieved. In particular, without any attempt to introduce the concept of extension, there can be no definite solution to the problem of referential indeterminacy. The problem is taken up again in the *Grundgesetze*, part of whose purpose is to complete the project begun in the *Begriffsschrift*, and continued so brilliantly in the *Grundlagen* (see *Grundgesetze*, p. viii).

In *Grundgesetze*, instead of an explicit treatment of extensions, Frege adopts the the primitive notions 'function' and 'value-range of a function', which is meant to be something like the function in extension or the graph of the function. Value-ranges are treated theoretically in *Grundgesetze* (§20) through the infamous Law V which says

$$\forall f,g[\mathrm{vr}_f = \mathrm{vr}_g \leftrightarrow \forall y(fy = gy)] \qquad \text{(Law V)}$$

where 'f','g' stand for functions (including functions mapping to the objects *True* and *False*), and 'vr_f' stands for the 'value-range of the function f'. Concepts can be construed as functions which take objects to truth-values, and it is then easy to define extensions. If transposed to the language of extensions, Law V would read:

$$\forall F,G[\mathrm{ext}_F = \mathrm{ext}_G \leftrightarrow \forall y(Fy \leftrightarrow Gy)] \qquad \text{(Law V, extensions)}$$

Numbers can now be defined explicitly more or less as is done in the *Grundlagen*; i.e., NxF is defined as the extension of the higher-order concept $G \approx F$, everything now being translated into the language of value-ranges.

But does this explicit definition solve the problem of referential fixity?

1.4 Explicit definition and referential fixity

Clever explicit definition enables us to work with objects without expanding the realm of primitives and the list of primitive propositions; Frege's work, whatever its other merits, is a hymn to clever definition. But Frege was also perfectly aware that one cannot define everything, i.e., not the primitives:

> It will not always be possible to define everything properly, precisely because we must strive to get back to the logically simple, which, because it is such, is not strictly definable. I must therefore be satisfied with indicating through hints what I mean. Above all, I must endeavour to be understood.[43]

But the problem of referential fixity is surely no less serious for the primitives than for any other terms, for lack of referential fixity among the primitives would mean lack of referential fixity everywhere, no matter how clever the actual definitions. Hence the strictures about the meaning of axioms and basic laws so forcefully expressed to Hilbert.

[43] *Gg* I, p. 4.

Frege addresses this problem by introducing the notion of 'elucidation'. Frege writes to Hilbert:

[Elucidations] are similar to the definitions, in that they are also concerned with fixing the meaning of a sign (of a word). But in addition, they contain elements whose meaning cannot be assumed as known completely and beyond question, perhaps because they are used variously or ambiguously in the language of everyday life. In the cases where a meaning is to be given to a sign which is logically simple, then one cannot give a definition proper, but one must content oneself with fending off the unwanted meanings which crop up in the use of language, indicating the one intended. In doing this, certainly one must always count on a cooperative understanding trying to hit upon the meaning. Such statements of elucidation cannot be used in the same way that the definitions can, because they lack the necessary precision. For this reason, as I said, I confine them to the forecourt.[44]

Elucidations are thus meant primarily as hints to enable an interlocutor to 'catch on', as a means of achieving mutual understanding at the fundamental level. They certainly do nothing to reveal the 'logical linkage of truths'; if they could be rendered precise enough to serve as definitions, then the alleged primitives, to which they are addressed, would not actually be primitives. In GG 1906, Frege says that an individual researcher doing research 'just for himself' would not need elucidations. This is in contrast to definitions; even if mutual understanding were guaranteed, definitions would still be necessary, for these are fundamental in enabling formal proofs; the central role of these, after all, is to signpost the route back to the primitives and thereby to the primitive propositions.

The context of the Frege–Hilbert correspondence is geometry, and a glance at geometry's history shows that Frege is following in a venerable tradition.[45] But Frege's point is surely much more general. If we look at Frege's own work, it is clear that the initial explanatory sections of *Grundgesetze* are meant to ensure that the logical primitives have a clear meaning, and these sections thus amount to extended elucidations. Elucidations, then, are certainly meant to push in the direction of fixed reference, as the letter to Hilbert

[44] Frege to Hilbert, 27 December 1899.

[45] This, I think, emerges from the extensive discussion in Heath's edition of *Euclid's Elements*, vol. I, pp. 143–51, pp. 155ff.; see also Pasch, *Vorlesungen*, pp. 16–17.

strongly implies. The *Grundgesetze* itself goes further than this. §32 begins:

Thus, it is shown that our eight original names [for the logical primitives] have a meaning, and therefore so do all names properly built out of them.[46]

(The context of this passage is set in the discussions taken up in §2, below.) Surely, then, if meaning has been fixed at all, it is through a process of elucidation.

There are various problems here, important in the light of Frege's criticisms of Hilbert. Firstly, the confidence Frege expresses in *Grundgesetze* is not consonant with what he says elsewhere about the uncertainty surrounding elucidations. Consider this:

The purpose of elucidations is a practical one, and when this is achieved, one must be satisfied. In this, one must be able to bank on good will, on a cooperative understanding, on guessing; for without a figurativeness in the expression one can often not get anywhere. (GG 1906, p. 301)

Furthermore, as he says, what distinguishes elucidations from proper definitions is that the latter leave nothing to 'guesswork'. As he says:

[Definitions] also serve the purpose of mutual agreement, but they achieve this in a far more complete way than elucidations do, since they leave nothing to guesswork, and do not need to reckon on cooperative understanding, or on good will. (GG 1906, p. 302.)[47]

The implication is clear: the elucidations *do* leave room for 'guesswork'. Moreover, as Frege says in the letter to Hilbert, elucidations are (necessarily) given in, or imbued with, the 'language of everyday life'. However, one of the very purposes of the extended *Begriffsschrift* project is to circumvent the ambiguities, unclarities and misunderstandings that the 'language of everyday life' is heir to. It seems as if the reliance on elucidation, at least for *fixing* reference, is tantamount to an admission that this circumvention is ultimately impossible. How can proper definitions communicate

[46] *Gg*, vol. I, §32, p. 50.

[47] Incidentally, Frege says that Hilbert's 'definitions' are not elucidations, since their ambition is to be the 'foundation stone of the science', to 'serve as premises of inferences'; see GG 1906, p. 302.

meaning precisely if they ultimately rely on primitives, the fixity of whose meaning relies on 'cooperative understanding' and guesswork? This is not necessarily in itself a serious problem, but it becomes one in the face of Frege's insistence to Hilbert on the fixity of reference for all terms, even at the level of the primitives. Frege's view for geometry, just as for logic/arithmetic, is that the mere choice of primitives and the statement of axioms about them is not enough. That is surely one of the central problems with taking *HP* as primitive, and it is also at the heart of Frege's criticism of Hilbert with respect to 'point'.

Moreover, is not Frege simply wrong in thinking that the elucidations can be dispensed with? Even though they play no role in showing the 'logical linkage of truths', their role in fixing reference is surely crucial, even for the idiosyncratic practitioner. For what is to stop isolated practitioners from associating quite distinct meanings to the primitives? And is not this just the problem raised many times in the *Grundlagen*? Without successful elucidation, the whole basis for common understanding would thus crumble. (See the examples below, pp. 442ff.) Moreover, even *with* elucidation, Frege is in philosophical trouble of a different kind, for reliance on elucidation is surely reliance on knowledge extra to the system, and this might represent a severe danger for the logicist project.

The philosophical problems occasioned by reliance on elucidations are not my main concern here, so let us leave these aside and return to explicit definition.

2 PERMUTATIONS

2.1 The permutation argument

We have not finished with the Caesar difficulties, and it is important to consider in this context §10 of volume I of *Grundgesetze*, a section of vital importance in setting out Frege's elucidations of the logical primitives, an elucidation which is continued in §31 and which culminates in the strong declaration of §32 (previous page).

§10, short as it is, is not one of Frege's clearest. It also comes *before* the official adoption of Law V (in §20), though Frege in §9 does adopt an informal version of Law V by way of an initial explanation of how to deal with value-ranges. The arguments in §10 and in §31 are apparently designed to offer some explanation of the way value-ranges behave in the assignment of truth-values, how technically we could treat the truth-value objects, the *True* and the *False*, as themselves value-ranges, and what values the basic functions (identity, the judgement stroke, etc.) would take for what values as arguments. These convolutions make the passage hard to summarize, and I will not attempt a detailed analysis here.[48] Rather I want to focus on Frege's new presentation of the indeterminacy argument.

Frege begins §10 with a statement of the possibility of indeterminacy very similar to the statements of the Julius Caesar problem in the *Grundlagen*, i.e., the indeterminacy problem in its *first* sense, of determining whether an object not given as a value-range is identical to one which is so given. Following this statement, Frege then, without further preamble, introduces a second indeterminacy consideration, a very *general* statement of the *other* kind of indeterminacy that arises even when we assume fixed boundaries, i.e., the question of which class of things the concept term 'value-range' picks out. Frege's argument here was christened by Dummett the *permutation argument*,[49] although (in most cases) the important thing about the 'permutations' involved is that they are one–one transformations of the object domain *into*, and not necessarily *onto*, itself. Suppose, says Frege, that we have a function $X(x)$ from objects to objects which is one–one, and suppose we insist on the stipulation for functions f,g and their value-ranges which Law V (or its informal version) lays down, viz.,

$$\mathrm{vr}_f = \mathrm{vr}_g \leftrightarrow \forall y\,(fy = gy).$$

Then

$$X(\mathrm{vr}_f) = X(\mathrm{vr}_g) \leftrightarrow \forall y(fy = gy)$$

must also be true, since X is one–one. The effect of this is as follows. Suppose we think that the concept 'value-range' picks out a certain extension *D*; the argument shows that (for all Law V can judge), it

[48] For extensive summaries of §10, see M. Dummett, *Frege: Philosophy of Mathematics* (London: Duckworth, 1991), pp. 209–17, T. Ricketts, 'Truth-values and courses-of-values in Frege's *Grundgesetze*', in W. W. Tait (ed.), *Early Analytic Philosophy: Frege, Russell, Wittgenstein* (Chicago: Open Court Publishing Company, 1997), pp. 187–211, and R. Heck, 'Grundgesetze der Arithmetik I, §10', *Philosophia Mathematica*, 7 (1999), pp. 258–92.

[49] Dummett, *Frege: Philosophy of Mathematics*, p. 211.

could just as well be the extension $X[D] \neq D$ which is picked out. With this argument, then, Frege seems to recognize the second indeterminacy worry explicitly, or it seems that he should have. We will return to this point.

There is a sense in which the general argument Frege gives here generates the first worry, too. Suppose we have two value-ranges vr_f, vr_g and a mapping X which takes vr_g onto some other object t, but leaves vr_f unchanged. Then not only is

$$\mathrm{vr}_f = \mathrm{vr}_g \leftrightarrow \forall y(fy = gy)$$

correct according to Law V, but so is

$$\mathrm{vr}_f = X(\mathrm{vr}_g) \leftrightarrow \forall y(fy = gy).$$

In other words, from the correctness of $\forall y(fy = gy)$, which (we assume) is the only principle to which we have access governing the 'identification' of value-ranges (their 'recognition'), one cannot conclude anything final about the reference of vr_f, especially since there is no guarantee that the object $X(\mathrm{vr}_g)$ is a value-range at all, or, even if it is, which function generates it.[50]

How does Frege think the indeterminacy of reference which the argument threatens can be overcome?

In this way, that it is determined for each function when it is introduced which values it obtains for value-ranges as arguments, just as for all other arguments.[51]

It is clear at this point that he relies on a general version of the Context Principle for reference, as Dummett points out,[52] in the form: a singular term has a reference if every function which has that singular term as an argument has a reference, with the corresponding additional clauses for functions of more than one

[50] So-called abstraction principles such as *HP* and Law V merely assert the existence of separate object representatives for each equivalence class determined by the equivalence relation used on the right-hand side of the principle. This means that the theoretical work must be done by the selection of a representative from the equivalence class, in the case of Law V, some function from the class. In the case of, say, the very differently constructed *ZFC* and the von Neumann ordinals, these numbers are selected from *within* the equivalence class they represent, so no such subsequent choice is necessary.

[51] *Gg*, vol. I, §10.

[52] See Dummett, *Frege: Philosophy of Mathematics*, p. 212.

argument. (Frege's use of functions in *Grundgesetze* means that this generalizes the demand that every sentence in which the term appears has a truth-value.) He then turns his attention to the functions introduced hitherto in his system, i.e., the primitive functions, identity, the horizontal function (the judgement stroke), and the negation function. He correctly observes that what is at issue is really only the determination of the truth-values of identity statements. Law V governs this in the case where both objects referred to on either side of an identity are value-ranges, but not in the case where at most one of them is. Since, Frege argues, he has up to this point merely introduced two kinds of objects, the value-ranges and the truth-value objects, and we can simply list the identity conditions for identity statements involving just the latter, the problem reduces to determining the truth-conditions for statements of the form '$vr_f = V$' (where 'V' stands for one of the truth-value objects), in other words, the Julius Caesar problem for value-ranges. He then concludes that this problem would not arise were it the case that the truth-value objects are themselves value-ranges, for then Law V *would* give the identity conditions. In other words, for Frege, the problem apparently arises because there seem to be objects in the domain *other than* the value-ranges, namely (at least) the truth-value objects, things which are not *given* as value-ranges. The problem could therefore be solved, as far as Frege is concerned, by showing that the truth-value objects *are* value-ranges. But can they be value-ranges, and, if so, which value-ranges are they?

They are not value-ranges in any intuitive sense; as Ricketts makes clear,[53] the *Grundgesetze* simply takes over the view of 'On sense and reference' that the two truth-values are primitive objects denoted by (correctly formed) declarative sentences. But Frege now proceeds to argue that they could in fact be identified as value-ranges, and that doing so will not violate Law V. To show this, he exploits the very permutation argument used previously to state the indeterminacy problem.

The considerations here are somewhat roundabout. If the *True* is indeed the value-range of some function Φ, then the truth-value of the statement $vr_\Phi = vr_f$ (for any f) clearly follows from Law V; and if the *True* is *not* in fact a value-range, the truth-value is also

[53] Ricketts, 'Truth-values', pp. 187–8.

decided, since it is false, although this cannot be decided by Law V. (Frege at this point does not say this.) Suppose now, says Frege, that we have two functions Φ and Ψ, and that there are objects associated with Φ and Ψ, which we denote by $\tilde{\alpha}\Phi(\alpha)$ and $\tilde{\eta}\Psi(\eta)$ respectively.[54] We are now free to stipulate that $\tilde{\alpha}\Phi(\alpha) = \tilde{\eta}\Psi(\eta)$ means the same as $\forall\alpha[\Phi(\alpha) = \Psi(\alpha)]$, without thereby being in the least able to conclude that $vr_\Phi = \tilde{\alpha}\Phi(\alpha)$, even though we know that both the identities $vr_\Phi = vr_\Psi$ and $\tilde{\alpha}\Phi(\alpha) = \tilde{\eta}\Psi(\eta)$ mean the same as $\forall\alpha[\Phi(\alpha) = \Psi(\alpha)]$. As Frege says:

We would simply have a class of objects, which have names of the form '$\tilde{\eta}\Psi(\eta)$', and for which the criterion for differentiation and recognition is the same one which holds for value-ranges.[55]

Frege now invokes the permutation argument. Suppose the truth-value objects are denoted by *True* and *False*, and suppose we have functions Λ and M which do not always have the same values for the same arguments. We can now define a function $X(x)$ as follows:

$$X(x) = \begin{cases} \tilde{\eta}\Lambda(\eta), & \text{if } x = \mathit{True} \\ \mathit{True}, & \text{if } x = \tilde{\eta}\Lambda(\eta) \\ \tilde{\eta}M(\eta), & \text{if } x = \mathit{False} \\ \mathit{False}, & \text{if } x = \tilde{\eta}M(\eta) \\ x, & \text{otherwise.} \end{cases} \qquad (\mathit{Perm})$$

X is one-one, since Λ and M are not extensionally equivalent, which means that $X(\tilde{\alpha}\Phi(\alpha)) = X(\tilde{\eta}\Psi(\eta))$ must mean the same as $\forall\alpha[\Phi(\alpha) = \Psi(\alpha)]$ just as $\tilde{\alpha}\Phi(\alpha) = \tilde{\eta}\Psi(\eta)$ does. Frege goes on:

The objects whose names are of the form '$X(\tilde{\eta}\Phi(\eta))$' would then be recognized by the same means as the the value-ranges, and indeed $X(\tilde{\eta}\Lambda(\eta))$ would be the True and $X(\tilde{\eta}M(\eta))$ the False. (p. 17)

He concludes that, 'without falling into contradiction' with Law V for value-ranges,

[54] '$\tilde{\alpha}$' apparently represents some general term-forming operator associating objects with functions.

[55] *Gg*, vol. I, p. 17.

it is always possible to stipulate that an arbitrary value-range is the True and an arbitrarily chosen different value-range is the False. (p. 17)

For the sake of definiteness, Frege then chooses to identify the True and the False with what we would call their unit classes, i.e., *True* becomes {*True*}, more correctly in his system, the value-range of the horizontal function (the judgement stroke), whereas *False* becomes the value-range of the function f given by:

$$f(x) = \begin{cases} True, & \text{if } x = False \\ False, & \text{otherwise.} \end{cases}$$

Frege's whole line of argument seems less than perspicuous; but the point of repeating its core here is simply to make it clear that the example of (*Perm*) shows that Frege *recognizes* that there are functions X which permute the domain in such a way as to preserve the truth-value of the axioms, in this case Law V, and indeed consciously exploits the fact. The generality of the point is only a whisker away. As Dummett says:

a similar argument would defeat any claim to have fixed the reference of the primitive vocabulary of any formal language (provided, in the general case, that the extensions of the primitive predicates were also subjected to the permutation).[56]

And in a footnote, Dummett assimilates Frege's permutation argument to the one used by Hilary Putnam against metaphysical realism.[57]

Be that as it may; the point I wish to stress is that it seems clear that such permutations are really what lie behind many of the other arguments from the *Grundlagen* about failure to fix reference. Thus, what Frege consciously exploits here makes explicit what is *implicit* in many of the earlier considerations, and yet does so without recognition that the problems raised earlier are not solved. Let us look at the examples again.

[56] Dummett, *Frege: Philosophy of Mathematics*, p. 211.

[57] For a general account of this form of argument, see M. Hallett, 'Putnam and the Skolem paradox', in P. Clark and S. Read (eds.), *Reading Putnam* (Oxford: Blackwell, 1994), pp. 66–97.

2.2 *The pervasiveness of permutations*

1. Take the case Frege raises about the number one. Recall that Frege is worried about the possibility of there being two different people, A and B, who take the term '1' to denote distinct things, a and b respectively, and yet agree on the truth-value of all the usual statements involving the numeral '1'. Frege's point could be put by saying that there is a permutation X of the domain of objects M underlying A's interpretation I_M of the terms and sentences, and where $\mathrm{X}(a) = b$ (with $b \neq a$), and where the resulting $\mathrm{X}[M]$ (which we can assume $= M$) forms the basis of a new interpretation $I_{\mathrm{X}[M]}$, which is B's, and which satisfies exactly the same sentences involving reference to one as does I_M.

2. Now consider the second form of the indeterminacy objection stated with respect to *HP*. Assume that the domain M of objects is again given, and that $N \subseteq M$. We can now generalize the situation depicted above concerning the number one. Assume that there is an interpretation I_M on the basis of which N is picked out as the domain of numbers, and which satisfies *HP*; now let X be a mapping which permutes M in such a way that $\mathrm{X}[N] \neq N$; the danger now is that we can form a new interpretation $I_{\mathrm{X}[M]}$ where $\mathrm{X}[N]$ interprets the numbers, and where $I_{\mathrm{X}[M]}$ satisfies the same sentences about the numbers as does I_M, in particular *HP*. This would mean that *HP* cannot 'determine' the reference of 'Number'. One can dramatize this by postulating individuals A, B who base their understanding of the numbers on I_M and $I_{\mathrm{X}[M]}$ respectively. Hence, they will both take numbers to be objects which are associated with concepts and governed by *HP*, and will thus agree on the truths that follow from *HP*, but clearly will differ about what objects the numbers are.

3. Consider another, rather different, example, which shows how widespread the problem is. In §26 of the *Grundlagen*, Frege suggests a somewhat analogous situation in geometry. Suppose we take an interpretation $I = I_{P,\ L,Pl}$ of projective geometry based on three domains of objects, points, lines and planes satisfying all the central propositions of three-dimensional projective geometry. Now consider a permutation X on the domains which is such that $\mathrm{X}[P] = Pl$, $\mathrm{X}[Pl] = P$, but which is the identity map on the lines, i.e., $\mathrm{X}[L] = L$. This yields another interpretation $I' = I'_{\mathrm{X}[Pl],\ \mathrm{X}[L],\ \mathrm{X}[P]}$ which also satisfies all the propositions of the geometry. Now we can imagine, just

as Frege supposes, that there are two individuals, *A* and *B* such that the interpretation generated by *A*'s intuition is *I*, and that generated by *B*'s intuition is *I′*; the same sentences will be satisfied, but the references (unbeknownst to *A* and *B*) will be switched. In other words, the axioms do not determine the reference of the primitives. Indeed, Frege himself even suggests that it is agreement on the axioms and theorems which really matters and not the particular nature of the reference yielded by the different intuitions.

4. It might be thought that projective geometry is special, since the 'ambiguity of reference' which Frege relies on here is really at root the projective Duality Principle. But moving from projective geometry to the case of Euclidean geometry would not help Frege's case, even though the Duality Principle now fails, for the same will hold of any 'permutation' which preserves the truth of the axioms. Indeed, one example is surely the one which Frege himself is exercised about in his correspondence with Hilbert, i.e., the example where Euclidean points are mapped to pairs of elements in Hilbert's Ω (see note 43), and lines and planes to the right kind of linear equations.

In the case of cardinal numbers, it is not difficult to see that the adoption of explicit definition and the call on Law V do not really help matters. Frege's explicit definition fixes the numbers to certain value-ranges, and the further elimination of the truth-value objects in favour of certain value-ranges implies that it is consistent to assume that there are no objects present *other* than value-ranges. This guarantees that the numbers are *logical* objects, assuming that this correctly describes the nature of the value-ranges, so the numbers are thus bound to be the 'right' sort of thing. But if the underlying domain of objects is *VR*, even if the only things in *VR* are value-ranges, there can still be a non-trivial permutation Χ of *VR* where Χ[*N*] = *N* (*N* being the collection of Frege natural numbers in *VR*) where all the basic principles hold, Law V, of course, and the derived *HP*, the number operator now being taken to refer to the objects in Χ[*N*].

In short, the possibility of permutations seems to show that, no matter where we turn, the axioms (*HP*, Law V, the projective or Euclidean axioms) do not determine the underlying objects.

The examples are not fanciful. Consider the following, very similar case. Assume that the domain of quantification *M* for objects is fixed to Frege's satisfaction, and that M contains Frege's numbers and natural numbers; call these latter $0_M, 1_M, 2_M, \ldots$, collectively making up $\mathcal{N}_M$. Let us assume also that φ is the mapping on $\mathcal{N}_M$ which takes each number to its immediate succesor. Assume now that there is a permutation X of M such that $X(0_M) \neq 0_M$, and where $X(0_M)$ is some arbitrary object *not* a natural number (Julius Caesar?). Then it follows by Dedekind's Theorem 126 on the definition of functions by recursion[58] that there is a unique one–one mapping ψ on $\mathcal{N}_M$ satisfying the conditions:

(a) $\psi[\mathcal{N}M] \subseteq M = X[M]$
(b) $\psi(0_M) = X(0_M)$
(c) $\psi(\varphi(n_M)) = X(\psi(n_M))$.

This shows that the permutation X induces a map on the natural numbers which preserves their structure; it follows that the sequence

$$\mathcal{N}_{X[M]} = \{X(0_M), X(X(0_M)), X(X(X(0_M))), \ldots\}$$

appears just like the numbers. Frege incidentally was in a perfect position to prove this, for he actually proves a version of the theorem justifying definition of functions by recursion on the natural numbers in GG 1903, Theorem 256.[59]

How like the numbers is such a set as $\mathcal{N}_{X[M]}$? Dedekind gives a clear answer in article 134 of his *Was sind und was sollen die Zahlen?*:

[58] Dedekind, *Was sind und sollen die Zahlen?*.

[59] That Frege's system is in a position to echo Dedekind's main results should be no surprise. Dedekind's proof follows from the (second-order) principle justifying proof by induction; see Theorems 59, 60. In Frege's systems (in fact using just *HP*), this is an easy consequence of his definition of the ancestral, a definition which is in some ways very similar to Dedekind's Definition 44 in Dedekind, *Was sind und sollen die Zahlen?*, of the *chain* of a set $A \subseteq M$ under a one–one mapping of M onto itself. Dedekind denotes this by A_0, which is the smallest chain which includes A, i.e., the intersection of *all* chains including A. It is easy to transform Frege's definition of the ancestral relation xR^*y into Dedekind's definition of the set A_0. First, restrict the Frege definition to *functional* relations; this will then define the class of y ancestrally descended from x, i.e., if we look at the class $\{z : z \in F\}$ instead of the concept F. We then consider this for all $x \in A$. On the Dedekind side, then consider the condition for any y to be in A_0. See W. Demopoulos and P. Clark, 'The logicism of Frege, Dedekind, and Russell', in S. Shapiro (ed.), *Oxford Handbook of Philosophy of Mathematics and Logic* (New York: Oxford University Press, 2005), pp. 129–65, especially pp. 140–1.

every theorem about the numbers, i.e., about the elements n of the simply infinite system N [here $\mathcal{N}_M$] ordered by the mapping φ, and indeed every such proposition in which we leave entirely out of consideration the special character of the elements n and discuss only such concepts as arise from the arrangement φ, possesses perfectly general validity for every other simply infinite system Ω [here $\mathcal{N}_{X[M]}$] ordered by a mapping θ [here X restricted to $\mathcal{N}_{X[M]}$] and its elements ν, and that the passage from N to Ω (e.g., also the translation of an arithmetical theorem from one language into another) is effected by the mapping ψ.[60]

(ψ is the map in effect given by $n_M \mapsto X \ldots X\ (0_M)$.) Dedekind points out that natural number system*s* in effect form an equivalence class, and that *any* member of the class will satisfy all the same *arithmetical* theorems. One thing that therefore follows from Dedekind's isomorphism theorem is that the genuinely true sentences about natural numbers are the ones which are true whichever representative is chosen from the equivalence class of simply infinite systems, and not those which allude to 'the special character of the elements *n*'. Thus, a sentence like '$0 = a$', where 'a' stands for the particular value-range which Frege gives as the definition of 0 (*Grundgesetze*, §40) will only be made true by some of the simply infinite systems.[61] Thus, if $\{0_M, 1_M, 2_M, \ldots, n_M, \ldots\}$ are the Frege numbers in a domain of value-ranges ruled over by Law V, then, according to Dedekind's characterization of simply infinite systems, $\{1_M, 2_M, \ldots, (n+1)_M, \ldots\}$ must be a simply infinite system of objects (satisfying Law V, of course) satisfying all the right theorems and where Frege's '1_M' now plays the role of the least element, i.e., zero. But note that *this* zero will *not* satisfy the sentence '$0 = a$', since Frege's 1_M *is not* that element.[62]

To summarize, Frege saw the possibility of these transformations X as giving rise to a kind of indeterminacy, to which both his *Grundlagen* and *Grundgesetze* attempt responses. The *Grundgesetze* is to some extent reconciled to 'permutations', and indeed exploits

[60] Dedekind, *Was sind und sollen die Zahlen?*, §134.

[61] Cf. R. Heck, 'The Julius Caesar objection', in R. Heck (ed.), *Language, Thought and Logic: Essays in Honour of Michael Dummett* (Oxford and New York: Oxford University Press, 1997), pp. 273–308, at p. 290.

[62] The generality of Dedekind's point does not quite come through in Frege's setting. If one insists that all objects are in fact value-ranges, then 'All natural numbers

them. Yet it is clear from Dedekind's analysis that not all of the indeterminacy which Frege saw as undesirable is removed in the *Grundgesetze* framework. Dedekind, on the other hand, saw the possibility of 'permutations' as simply a fact about central mathematical theories, even ones which (as we would put it now) are categorical. Thus, he takes it that a condition on a correct account of mathematics is that it be based on this fact, and should not seek to avoid it. *Any* representative of the appropriate isomorphism class will do, and Dedekind saw it as a *mistake* to attempt to fix reference over and above the satisfaction of the basic axioms (one is tempted to say *characteristic* axioms) in the way that Frege appears to want. In a letter to Weber of 23 January 1888, Dedekind says the following:

This is precisely the same question that you raise at the end of your letter in connection with my theory of irrationals, where you say that the irrational number is nothing other than the cut itself, while I prefer to create something *new* (different from the cut) that corresponds to the cut and of which I say that it brings forth, creates the cut. We have the right to ascribe such a creative power to ourselves; and moreover, because of the similarity of all numbers, it is more expedient to proceed in this way. The rational numbers also produce cuts, but I would certainly not call the rational number identical to the cut it produces; and after the introduction of the irrational numbers one will often speak of cut-phenomena with such expressions, and ascribe to them such attributes, as would sound in the highest degree peculiar were they to be applied to the numbers themselves. Something very similar holds for the definition of cardinal number as a *class*; one will say many things about the class (e.g., that it is a system of *infinitely many* elements, namely, of all similar systems) that one would attach to the number (as a deadweight) only with the greatest reluctance. Does anybody think, or will he not gladly forget, that the number four is a system of infinitely many elements? (But that the number four is the child of the number three and the mother of the number five is something that nobody will forget.)[63]

are value-ranges' will come out to be true in all the simply infinite systems, even though it will clearly not be an arithmetical truth in Dedekind's sense.

[63] R. Dedekind, *Gesammelte mathematische Werke*, vol. III, ed. Robert Fricke *et al.* (Braunschweig: Friedrich Vieweg und Sohn, 1932), pp. 489–90; for an English translation, see Ewald, *From Kant to Hilbert*, p. 835.

It is of some importance to point out that Dedekind certainly *was* concerned with fixity of reference in Frege's 'fixed boundaries' sense; something like the Caesar problem is what lies behind his concern with 'intruders' ('non-standard elements'), and he uses the results mentioned above to show that his account of natural number solves it. This is made strikingly clear by a long passage (Article 6) in Dedekind's famous letter to Keferstein of 2 February 1890:[64] Dedekind emphasizes that the very notion of the chain of a set *A* under the mapping φ (i.e., A_0) is designed to show that there cannot be things among the natural numbers which ought not to be there. Indeed, it is just this definition which leads easily to the second-order Principle of Induction, and it is this Principle which in turn leads to Theorems 126 and 132 showing that any two systems satisfying the number axioms must be isomorphic. But Dedekind solves the intruder/Caesar problem in a relative way: 'intruders' are not present only because there can be nothing there which is not related in the right way to the distinguished element, not because there is something inappropriate about them as objects. Frege, of course, also proves corresponding results,[65] but it is not so clear why he proves them, in particular it is not clear that his purpose was to contribute to a solution of Caesar indeterminacy.[66]

Dedekind's view seems to have been more that it is matter of expediency and good mathematical practice not to give direct definitions of the Frege kind, consequently that a correct account of mathematics ought to take this into account. What is new with

[64] See R. Dedekind, 'Brief an Keferstein', in M.-A. Sinaceur, 'L'Infini et les nombres', *Revue d'Histoire des Sciences*, 27 (1974), pp. 251–78, at pp. 271–8, translation in J. van Heijenoort (ed.), *From Frege to Gödel: A Source Book in Mathematical Logic* (Cambridge, Mass.: Harvard University Press, 1967), pp. 99–103.

[65] See R. Heck, 'Definition by induction in Frege's *Grundgesetze der Arithmetik*', in W. Demopoulos (ed.), *Frege's Philosophy of Mathematics* (Cambridge, Mass.: Harvard University Press, 1995), pp. 295–333.

[66] In this context, the central result of *Gg* II is Theorem 263, which says that the cardinal number *Endlos*, which is Frege's name for the number of natural numbers, is also the cardinal number of any system which is isomorphic to the natural numbers. Part of showing this involves showing that any appropriate injection taking $\mathcal{N}$ into $\mathcal{M}$ must be a bijection. It is tempting to read this result as stating (*a*) an isomorphism theorem like Dedekind's, and (*b*) as showing, echoing Dedekind, that there can be no 'intruders' in the extension of the concept 'Natural number'. There is, however, nothing in Frege's text to support this, although Heck has put forward an argument for Frege's having intended the former. See Heck, 'Definition by induction', pp. 324–5.

Hilbert, though, is of the view that it is mathematically fruitful not to tie down reference in the way that Frege wanted. Let us now return to Hilbert and his work on geometry.

3 HILBERT'S PROJECT AND HIS MAIN REPLY TO FREGE

Hilbert deliberately took the position (see §1.3) that the primitives of an axiomatized mathematical theory should come with no fixed reference, and that it is in virtue of this that the axioms take on a 'defining role' in assigning them their basic 'characteristics'. *That* Hilbert takes this position on reference and definitions is clear. His correspondence with Frege also makes it clear both that this met with severe, sustained and articulate opposition from Frege, and that Hilbert nevertheless made no concessions to Frege's position. There is good reason for this: Hilbert was pursuing a distinctive mathematical programme of which this position was an essential part, a programme with a very different conception of foundational investigation from Frege's. The aim of this last section is to sketch briefly this different conception of mathematics, and to outline how it relates to Frege's difficulties with referential indeterminacy.

3.1 Hilbert's foundational investigations and reference

In his first and principal reply to Frege, Hilbert outlines his foundational project for geometry as follows:

If we wish to understand one another, then we must not forget the quite different nature of the intentions which guide us. I was forced to set up my system of axioms by necessity. I wanted to make it possible to understand those geometrical theorems which I regard as the most important results of geometrical research, that the Parallel Axiom is not a consequence of the other axioms, likewise not the Archimedean Axiom, etc. I wanted to answer the question whether the theorem that in two equal rectangles with the same base the sides are also equal* can be proved, or whether it has to be a new postulate, as it is in Euclid. I wanted to create the possibility of understanding and answering such questions as why the angle sum in triangles is 2 right angles and how this fact is related to the Parallel Axiom. I believe that my *Festschrift* shows that my system of axioms was shaped to answer such questions in a quite definite way and that in many cases these questions have very surprising and quite unexpected answers. This is also shown by the work of several of my students which pursue the methods of the *Festschrift*; here I mention only the forthcoming dissertation of Herr Dehn, soon to be published in the *Mathematische Annalen*. Such anyway was my main intention. In carrying this out, I certainly think I have set up a system of geometry which satisfies the most rigorous demands of logic, and with this I come to the answer to your letter proper. [* This theorem is after all the foundation of the whole theory of surface measurement.][67]

Three of the geometrical questions Hilbert mentions here have to do with logical independence. This is obvious in the case of the Parallel Axiom and the Archimedean Axiom. The third concerns the investigation of the relationship between the Euclidean Parallel Axiom and one of its first consequences, the Euclidean angle sum theorem (*AST*). Certainly *AST* fails in non-Euclidean geometry, and it was often thought to be equivalent to the Parallel Axiom, sometimes even taken as a possible substitute. However, Dehn's work showed that *AST* only implies the Parallel Axiom in the presence of the Archimedean Axiom. This involves showing that the Parallel Axiom is *independent* of *AST* if the Archimedean Axiom is not present.[68]

The other result mentioned by Hilbert is of a rather different kind. In his lectures on Euclidean geometry, Hilbert set out to show that his version of the Euclidean theory of 'surface content' can be founded on the Euclidean theory of linear proportion, relying only on the incidence, order and congruence axioms, with no assumption of any continuity principle. Important in this reconstruction of the Euclidean theory is the result (*Festschrift*, Theorem 27) that any two triangles with the same base and height have the same surface content. But then the question arises of whether the definition of

[67] Hilbert to Frege, 29 December 1899 (I).

[68] The result was in Dehn's dissertation, essentially M. Dehn, 'Die Legendre'sche Sätze über die Winkelsumme im Dreieck', *Mathematische Annalen*, 53 (1900b), pp. 404–39; this is the work mentioned by Hilbert, though this result was only a small part. It was also reported on extensively by Hilbert himself in a new section written for the French translation of the *Festschrift* (D. Hilbert, 'Les Principes fondamentaux de la géométrie', *Annales scientifiques de l'École Normale Supérieur* (3)17 (1900), pp. 103–209), a section which appeared also in the first English translation (Hilbert, *The Foundations of Geometry* (LaSalle, Ill.: Open Court Publishing Company, 1902). This section did not appear in subsequent editions of Hilbert's *Festschrift*.

surface content is empty or not, i.e., whether all triangles have the same surface content, rendering the notion trivial. To show that it is not empty, it is essential to show (*Festschrift*, Theorem 28) that any two triangles with the same content and on the same base have the same height; this result is just what is stated by Hilbert in his letter to Frege for rectangles. Hilbert, unlike those of his immediate predecessors who had also dealt with this question, proves the central theorems *without* use of the Archimedean Axiom or any other continuity principle.[69]

The contrast between these two different kinds of result is important. The result concerning surface content is of the kind which sets out to show that some proposition P *can* be proved from a given set of assumptions Σ (or more generally $\Sigma - \Gamma$), and the only way to do this, of course, is actually to exhibit a derivation from $\Sigma - \Gamma$. This kind of result was very important to foundational research in the later nineteenth century, and very widespread. Frege's work was of this kind, so was Pasch's work on the empirical nature of projective geometry, and Dedekind's on the reconstruction of the theory of real numbers avoiding reliance on geometrical intuition. *Some* of Hilbert's work in the reconstruction of Euclidean geometry is of exactly this kind, as the example mentioned illustrates. But the kind of work represented in the other questions he mentions in his letter to Frege is very different, for the aim in these cases is to show that P *cannot* be derived from Σ. This kind of work, with its focus on questions of independence, was what was genuinely novel in Hilbert's analysis of geometry. Both kinds of work are stressed in the Conclusion to Hilbert's *Festschrift*.[70] Hilbert says the following:

The present memoir is a critical examination of the principles of geometry. In this investigation, we have been guided by a fundamental tenet, namely to elucidate every question which presents itself in such a way that we examined whether or not the question can be answered in a prescribed way with certain restricted means. This basic tenet seems to me to contain a

[69] The complicated theoretical development is to be found on pp. 40–9 of the *Festschrift*; the connections are also set out, more clearly, in Hilbert, *Grundlagen der Euklidischen Geometrie* (1898/9), pp. 122–38. See respectively pp. 368–76 and 475–85 in Hallett and Majer (eds.), *Lectures*, p. 475.

[70] See Hilbert, 'Grundlagen der Geometrie', p. 89, or Hallett and Majer (eds.), *Lectures*, p. 525.

general and natural directive. In fact, whenever in our mathematical considerations we encounter a problem or conjecture a theorem, our passion for knowledge is only then satisfied when we have succeeded in giving the complete solution of the problem and the rigorous proof of the theorem, or when we recognise clearly the grounds for the impossibility of doing this and therefore the necessity of the failure.[71]

Hilbert goes on to stress the importance and fruitfulness of impossibility proofs in mathematics in general, and also to state that there is a close connection between his 'directive' and the demand for the 'purity of method', presumably since a full investigation of a 'purity' question will involve a thorough examination of what is, and what is not, deductively possible in a given context.[72]

Hilbert's remark in his letter to Frege that he 'wanted to make it possible to understand' these central results is key. The basic technique which Hilbert adopted for the investigation of *un*provability questions is that of modelling. For this, it is essential that the primitive concepts employed are not tied to their usual fixed meanings, but must rather be free for *reinterpretation*; no *one* interpretation of an axiom system is privileged above others, despite what might seem like the overwhelming weight of the interpretation underlying the 'facts' as originally given, for example, the weight of the 'intuitive' or 'empirical' origins. This is the point of Hilbert's insistence (see §1.3) that the geometrical primitives be divorced from any standard or intuitive meanings they might carry, and that only the axioms 'define' them. In this there is a radical departure from the kind of enterprise Frege was engaged in (not to mention Pasch and later Russell), an enterprise part of whose very point was to explain

[71] *Ibid.* In his 1898/9 lecture notes, Hilbert writes: 'However, we wish to set this as a modern principle: One should not stand aside when something mathematical does not succeed; one should only be satisfied when we have gained insight into its unprovability. Most fruitful and deepest principle in mathematics ('Grundlagen der Euklidischen Geometrie', p. 106, in Hallett and Majer (eds.), *Lectures*, p. 284). This is clearly the origin of the remark just cited from the *Festschrift*. See also D. Hilbert, 'Mathematische Probleme', in *Nachrichten von der königlichen Gesellschaft der Wissenschaften zu Göttingen, mathematisch-physikalische Klasse* (1900), pp. 253–6, at p. 261.

[72] For a thorough examination of Hilbert's treatment of 'purity of method' in his work on geometry, see M. Hallett, 'The "purity of method" in Hilbert's *Grundlagen der Geometrie*', in P. Mancosu (ed.), *The Philosophy of Mathematical Practice* (Oxford: Clarendon Press, 2008), pp. 198–255.

(thereby to delimit) the meaning of the primitives and thus of the axioms and theorems. For Hilbert, the primitives cease to have fixed reference, and the axioms thus cease to be, for Frege, genuine axioms at all. Thus, Hilbert's axiomatic method abandons the direct concern with the *kind* of knowledge represented in a given mathematical theory, i.e., with showing the kind of knowledge the individual propositions represent because they are ultimately about the specified primitives. Hilbert's method concentrates instead on what he calls 'the logical relationships' between the propositions in a theory.[73]

One of Frege's objections to this procedure was that, if we remove the usual Euclidean meanings from the axioms of Euclidean geometry, then they cease to be axioms of *Euclidean* geometry, or indeed Thoughts, at all.[74] It follows that, for Frege, Hilbert's method of establishing independence cannot work; for if *P* is shown to fail in a model of the Euclidean axioms which is not 'Euclidean', what can this tell us about *Euclidean* geometry? But this is to ignore the deep-seatedness of what Hilbert was proposing. Along with this approach to independence proofs goes a new picture of mature mathematics, a picture which was constant across Hilbert's concern with foundational investigation. In lectures on the foundations of geometry from as early as 1893/4, Hilbert says:

In general one must say: Our theory furnishes only the schema of concepts, which are connected to one another through the unalterable laws of logic. It is left to the human understanding how it applies this to appearances, how it fills it with material. This can happen in a great many ways. But always when the axioms are fulfilled, then the theorems hold precisely, too. The easier and more multi-faceted the application, the better* the theory. [* Every system of units and axioms which describes experience

[73] This is made clear in D. Hilbert, *Grundlagen der Mathematik*, lecture notes for a course held in the Wintersemester of 1921/2 at the Georg-August Universität, Göttingen (Göttingen: Library of the Mathematisches Institut). See also P. Bernays, 'Die Bedeutung Hilberts für die Philosophie der Mathematik', *Die Naturwissenschaften*, 10 (1922), pp. 93–9, at pp. 95–6.

[74] See, e.g., GG 1906, pp. 402, 424. For discussion, and many other citations, see J. Tappenden, 'Frege on axioms, indirect proof, and independence arguments in geometry: Did Frege reject independence arguments?' *Notre Dame Journal of Formal Logic*, 41 (2000), pp. 271–315.

completely is justified. Show that nevertheless the axiom system specified here is in a certain sense the only possible one.][75]

It is the mathematical theory *itself* which is 'only a schema of concepts' and which can be differently interpreted, both in the various (predictable and unpredictable) applications, sometimes to the physical world and sometimes in other mathematical theories, and also in meta-mathematical study.[76] Hilbert's 'way of understanding' the independence results therefore introduces, and is based on, the distinction between the axiomatized theory on the one hand and the various models on the other. The investigation of the (complicated) relationship between the two was to become a fixture of twentieth-century foundational work.

The same points are made again by Hilbert in his principal letter to Frege:

I have only one objection left on which to touch. You say my concepts, e.g., 'point', 'between', are not unambiguously determined; e.g., on p. 20, 'between' is taken differently and there a point is a number-pair. – Yes, it is obvious that any theory is actually only a framework or a schema of concepts, together with the necessary relations of these concepts to each other, and the base elements can be thought of in an arbitrary way. If, as my points, I think of some system of things, e.g., the system: love, law, chimney sweep … and then assume my axioms as relations between these things, then my theorems, too, hold of these things, e.g., the Pythagorean Theorem. In other words, any theory can always be applied to infinitely many systems of basic elements. One is required only to apply a one-to-one transformation and to ascertain that the axioms are the same also for the things transformed. In fact, one frequently applies this circumstance, e.g., the Duality Principle, etc., and I do in my independence proofs. All the statements of a theory of electricity naturally hold also of any other system of things which one substitutes for the concepts magnetism, electricity, assuming only that the axioms in question are fulfilled. This circumstance is not a defect* of a theory, and is in any case inescapable. Certainly, though, in my opinion the application of the theory to the world

[75] D. Hilbert, *Grundlagen der Geometrie*, lecture notes for a course to have been held in the Wintersemester of 1893/4 at the University of Königsberg (Göttingen: Niedersächsische Staats- und Universitätsbibliothek), p. 60, or Hallett and Majer (eds.), *Lectures*, p. 104.

[76] This is emphasized again in lectures in 1921/2, though with a slightly different stress; see Hilbert, *Grundlagen der Mathematik* (1921/2), p. 3.

of appearances always requires a certain measure of goodwill and tact: e.g., that one substitutes for points bodies as small as possible, for straight lines, things as long as possible, perhaps light rays, etc. Also one should not be all too precise in the examination of the propositions, since these are only propositions of the theory. Moreover, the further a theory is worked out, and the more finely branched its structure, then the form of its application to the world of appearances will become all the more obvious. It would require a very great measure of malice if one were to apply the more refined propositions of the theory of surfaces or of Maxwell's Theory of Electricity to appearances other than those for which they were intended. [* Rather, a powerful advantage.][77]

3.2 The mathematical fruits of Hilbert's position

The possibility of manifold interpretation is, for Hilbert, 'a powerful advantage'. The reason is clear: the more interpretations of a theory one can find, the greater the possibility of demonstrating independences of the most remarkable kind. Three remarkable independence investigations conducted by Hilbert are discussed in detail elsewhere;[78] they concern analysis of the famous Desargues Theorem (DT) of projective geometry, the Three Chord Theorem (TCT) and the Isoceles Triangle Theorem (ITT). What stands out particularly in these investigations is the idea that, although Hilbert sees the origin of elementary geometry in intuitive and even (perhaps especially) empirical investigation, it is higher mathematics which in the end informs the intuitive quite as much as the other way around, usually through the use of highly sophisticated analytic constructions. This kind of examination is what Hilbert calls 'analysis of intuition'.[79] This resolves into two separate investigations, one at the intuitive level, and one at the abstract level, levels which frequently interact and instruct each other. Furthermore, extracting the requisite information often itself involves a detour

[77] Hilbert to Frege, 29 December 1899 (I).

[78] Hallett, 'Purity of method', §8.4.

[79] In his *Festschrift*, Hilbert describes the 'logical analysis of our spatial intuition' as the prime task of geometry (see p. 3, or Hallett and Majer (eds.), *Lectures*, p. 436). In the official record of his 1898/9 lectures, Hilbert says that we could call the task he undertakes in his course the 'logical analysis of our capacity to intuit' (Hilbert, *Elemente der Euklidischen Geometrie*, p. 2, or Hallett and Majer (eds.), *Lectures*, p. 303).

into the abstract. One of the reasons why Hilbert thinks that intuition requires analysis is that it is not, for him (unlike for Frege), a certain source of geometrical knowledge, and certainly not a final source. Hence the need for the analysis, designed to throw light on the question: what is one committed to exactly when one adopts certain principles, among them principles suggested by intuition? In particular, what is shown is that, in interpreting and reinterpreting the geometrical primitives and thereby the geometrical propositions, one is not forced to abide by the intuitive; hence the often strange and contorted models (and thus the reliance on analytic techniques for full control over these interpretations) which one sees at work in the examples, e.g., in the analysis of DT and ITT. Although the investigation might start from questions raised by geometrical intuition, the final analysis produces results which inform or educate (perhaps even *challenge*) our intuition. In short, one has to leave the Euclidean view if one is to achieve more complete and precise information about what this view commits us to.

Thus, full investigation of geometry in this sense requires, first, its axiomatization, and proper examination of this axiomatization then requires, furthermore, that it be cut loose from its natural epistemological roots, or, at the very least, no longer immovably tied to them or to any other fixed interpretations of the primitives. According to Hilbert's new conception of mathematics, an important part of geometrical knowledge is knowledge which is quite independent of interpretation, knowledge of the logical relationships between the various parts of the theory, the way the axioms combine to prove theorems, the reverse relationships between the theorems and the axioms, and so on. And in garnering this sort of geometrical knowledge, there is not the restriction to the 'appropriate' which we see in the 'Euclidean' part of Hilbert's concerns. What is invoked in pursuing this knowledge might be some highly elaborate theory (a complex non-Archimedean field, as it is in the analysis of the ITT), a theory far removed from the 'appropriate' intuitive roots of geometry. Even in the cases of the fairly simple models of the analytic plane used to demonstrate the failure of the Planar DT, the models are far from straightforwardly 'intuitive'.

One might be tempted to say that the knowledge so achieved is not *geometrical* knowledge, but rather purely formal logical knowledge or (as it would be usually put now) *meta*-geometrical knowledge.

But although this designation is convenient in some respects, it is undoubtedly misleading. The 'meta'-geometrical results have a direct bearing on what is taken to be geometrical knowledge of the most basic intuitive kind; in particular it can reveal a great deal about the content of intuitive geometrical knowledge. In short, it effects an *alteration* in geometrical knowledge, and must therefore be considered to be a *source* of geometrical knowledge. To repeat: for Hilbert, meta-mathematical investigation of a theory is as much a part of the study of a theory as is working out its consequences, or examining its foundations in the way that Frege, for instance, does. In particular, and to repeat, one cannot fully understand the Euclidean (Fregean) framework unless one does this.

The examples mentioned above are striking. It is possible that Frege did not know of Hilbert's work on these, for none is fully represented in the original 1899 version of Hilbert's *Festschrift*. The central result on DT *is* represented, but what leads up to this result, namely, the philosophical reflection and analysis undertaken in the 1898/9 lecture notes, is suppressed; the analysis of the TCT is an important part of the 1898/9 lectures, but only the abstract algebraic mathematical result, and not the analysis itself, appears in the *Festschrift*; and the analysis of the ITT makes no appearance until the lectures of 1902 and a subsequent paper.[80] In any case, these investigations are all based on independence results, and (as was pointed out) Frege was sceptical about these. Nevertheless, Hilbert's view of the axiomatized theory (the 'framework of concepts') as itself the central object of mathematical study should have been clear also from Hilbert's important paper from 1900 'Über den Zahlbegriff', a paper which Frege certainly knew.[81] And here the view comes out clearly as quite separate from independence questions.

[80] See D. Hilbert, *Grundlagen der Geometrie, Ausarbeitung* by August Adler for lectures in the Sommersemester of 1902 at the Georg-August Universität, Göttingen (Göttingen: Library of the Mathematisches Institut, published as ch. 6 in Hallett and Majer (eds.), *Lectures*); D. Hilbert, 'Über den Satz von der Gleichheit der Basiswinkel im gleichschenklichen Dreieck', *Proceedings of the London Mathematical Society*, 35 (1902/3), pp. 50–67. Of course, it should be recalled that Frege *did* see the protocol of Hilbert's notes for the 1898/9 lectures. See note 20 above.

[81] D. Hilbert, 'Über den Zahlbegriff', *Jahresbericht der deutschen Mathematiker-Vereinigung*, 8 (1900), pp. 180– 5. See p. 429 above.

Hilbert's paper presents an axiom system for complete ordered fields, a system first presented (minus the Completeness Axiom) in the 1898/9 lectures on Euclidean geometry.[82] The (unstated) background to Hilbert's paper is the following question, which is pursued in the 1898/9 lectures and to some extent in the *Festschrift*. If a synthetically presented geometry such as Hilbert's is to match analytic geometry, then there must be a discernible 'magnitude structure' among, say, segment 'lengths' measured off from an arbitrarily chosen point of 'origin', since the numbers at the basis of an analytic structure would measure off just such segment lengths. The first step in showing this is to isolate the relevant 'magnitude properties', by capturing what is essential, and therefore jettisoning what is irrelevant, in the structure of the real number system as usually given, and what Hilbert identifies here at the core is an ordered field structure. The second step is showing that the operations of segment addition and multiplication in the synthetic geometry, when appropriately defined, possess these field properties. Clearly one of the things which is irrelevant, when viewed from the perspective of geometry, is that the real numbers are, for example, constructed as sets of a certain kind from the natural numbers. In other words, the axiom system for ordered fields has several, quite different, but equally natural, interpretations, one through the Dedekind Cuts in the rationals, another using equivalence classes of Cauchy sequences of rationals (Cantor's analysis) and a third, which is to the point in Hilbert's work, the appropriate geometrical segment calculus. Thus, there are many interpretations which arise quite naturally out of perfectly legitimate mathematical questions, not strange interpretations dreamed up for the purpose of an independence proof. The mathematical importance of *all* these ways of interpreting is surely one main reason why Hilbert expresses himself in favour of the 'axiomatic method' of presenting theories, and against what he calls the 'genetic method' of describing and generating mathematical objects,[83] the method exemplified in the 'generation' of the number systems through successive 'Fregean' definition.

Hilbert's presentation of the theory of real numbers underlines the shift away from a privileged interpretation of the theory.

[82] See also §13 of Hilbert's *Festschrift*, beginning on p. 26.

[83] Hilbert, 'Über den Zahlbegriff', pp. 180–1.

Moreover, it makes it clear that Hilbert's foundational work, as opposed to Frege's, is not concerned with defining central mathematical objects like numbers as far as it is possible to take such definitions, because it is not concerned with Frege's project, namely finding *the* basic foundational axiom systems and a definitional route to it. As a consequence, the Hilbert view consciously rejects the foundational use of what Frege terms 'analytic definitions', at least in presenting the final stage of a theory. What underlies the Hilbert view, taken over and expanded from Dedekind, is that the method of explicit definition is simultaneously both too precise and yet not on its own enough. By fixing the reference of the central terms, even if only relatively, explicit definitions rule out other possibilities, apparently arbitrarily. Moreover, and connected to this, the objects as defined often have too much structure, a point made clearly by Dedekind. (See the letter to Weber referred to above, p. 446.) But while the definitional approach is, in a sense, too precise, it is also not by itself enough. The definitions are subordinate to what, in effect, is an axiom system. That is, one has to be able to tell whether a given definition is a good one, and this requires showing that certain necessary requirements are satisfied, requirements which form, in effect, an axiom system, or part of one.[84] The axiomatic method then turns these necessary conditions into necessary and *sufficient* conditions; in short, the extra structure is jettisoned fully. Thus, it is the axioms alone which give the 'characteristics' of the primitive notions, which means that the characteristics will only to be able to circumscribe the reference of the primitives at best up to isomorphism. As Hilbert says to Frege, to repeat a passage from the correspondence cited earlier (on p. 426):

> The definitions (i.e., explanations, definitions, axioms) must contain everything, but this said should contain only that which is required for the construction of the theory. With respect to my division into explanations, definitions, axioms, which together make up the definitions in your sense, these certainly contain much that is arbitrary. Nevertheless, I believe that, in general, my ordering is serviceable and perspicuous.[85]

[84] To reinforce this point: what is it that makes two different, putative definitions of some notion (say Dedekind Cuts in the rationals and equivalence classes of Cauchy sequences of rationals) both good definitions of that notion?

[85] Hilbert to Frege, 29 December 1899 (II).

This statement is all the more remarkable when viewed in the light of what Dedekind says to Weber against explicit definition (see p. 446), and also in *Was sind und was sollen die Zahlen?* about the 'genuine' propositions of number theory (see above, p. 445).

All this means that, to a large extent, Frege-style definitions can be dispensed with. They have a very definite purpose, though, for they reappear as assignments of objects in the construction of an interpretation to assess the principles under investigation. We will come to this point in the next section.

3.3 Hilbert's method

Let us now tie this to Frege's difficulties with referential indeterminacy.

What exactly is Hilbert's method of giving models of his geometry and therefore independence proofs?

There are two ways to read this.

In the first way, a domain is given, (characteristically, in Hilbert's case, a domain of real numbers). Next, from within this domain, or a domain readily built out of it, parts are singled out to be the interpretation of 'point', 'line', 'plane'; the relation 'between' which holds among points is defined, i.e., an appropriate three-place relation over the point-objects is picked out; line segments are defined, and then the relation of congruence for line-segments, and so on. So, we might think of the specification of the domain then supplemented by the *definition* for this domain of the concepts, relations, etc. which are to satisfy the primitives; thus reference for them is *fixed* (through Frege-style definitions) once the domain has been specified. Given this, the propositions in the 'formal' language of geometry are reinterpreted through this new way of reading them. As Demopoulos has pointed out,[86] this is close to the modern way of interpreting a formal language. A domain is specified (characteristically, an unstructured set), and relative to this domain one fixes interpretations for what are called the *non-logical constants* of the language, characteristically the primitive concepts and primitive

[86] See W. Demopoulos, 'Frege, Hilbert and the conceptual structure of model theory', *History and Philosophy of Logic*, 15 (1994), pp. 211–25.

relations, considered as non-logical constants. There are thus two levels of variation: (1) The first variation is effected by varying domains. Once the domain is fixed precisely, the interpretation of the primitives (non-logical constants) is fixed to that domain, which is done by Frege-style explicit definition. (2) Of course, given the domain, other definitions of the primitives could be given, and this leads to a second level of indefinite variation; different definitions for the primitives would lead to a different interpretation. The point, though, is that the primitives are not treated as variables varying over the domain of interpretation.[87]

Let us use the example of Hilbert's which Frege refers to, involving the minimal Pythagorean field Ω. (See note 43.) Given ordinary analysis together with some set theory, the domain can be precisely specified as $\Omega \times \Omega$; 'point' is defined as $\{(x,y) : x,y \in \Omega\}$, and a straight line is now to be thought of as a collection of points determined by a linear equation $ux + vy + w = 0$, where u,v,w are parameters taken from Ω, i.e., the collection of pairs (x,y) for which $ux + vy + w = 0$. So the straight lines are sets of sets of points. The same will be true (for triples) when we define the interpretation of 'plane'. This shows us that the domain cannot straightforwardly be Ω or even $\Omega \times \Omega$, for the set of lines will be a countable set in $P(P(\Omega \times \Omega))$ (or $P(P(\Omega \times \Omega \times \Omega))$ when we consider planes); but the simplest way to think of it is as a many-sorted system, with different domains for the different sorts. The point, though, is that the definitions are fairly straightforward, given analysis and a modicum of set theory, and that they look like Frege-style definitions.

There is a second way of interpreting what Hilbert does. Suppose one considers the geometry he describes as written in a language $\mathcal{L}_E$. Suppose, as before, the interpretation is to be built via the Pythagorean field Ω.

This is described in the language of analysis, augmented perhaps by the modicum of set-theoretic language adequate to the definitional means. Call this $\mathcal{L}_A$. What we get now is a 'translational' scheme something like the following:

[87] Note that, conceived in this way, the traditional Euclidean view (perhaps based on intuition) is not available as an interpretation.What exactly is the domain? And even once the domain is taken as given, how is one to *define* 'point', 'straight line', 'congruent', etc. to interpret the non-logical constants?

$\mathcal{L}_E$	mapping τ	$\mathcal{L}_A$
point	$\mapsto$	the set of pairs of numbers in Ω, i.e., Ω×Ω;
straight line	$\mapsto$	the set of sets of elements in Ω × Ω satisfying any two parameter linear equation defined over Ω;
plane	$\mapsto$	the set of sets of elements in Ω × Ω × Ω satisfying any three parameter linear equation defined over Ω;
⋮	⋮	⋮

There are now several interesting things about this. For one thing, we get the following form of consistency proof. Suppose $\phi_1,\phi_2, \ldots, \phi_n$ is a proper proof of Ψ in $\mathcal{L}_E$ from certain axioms, say $\phi_1,\phi_2, \ldots, \phi_k$ ($k \leq n$), then, provided we insist that the translation τ preserves logical form, $\tau(\phi_1),\tau(\phi_2), \ldots, \tau(\phi_n)$ will be a proper proof in $\mathcal{L}_A$ of τ(Ψ) from the premises $\tau(\phi_1),\tau(\phi_2), \ldots, \tau(\phi_k)$. From this it follows that if Ψ were a contradiction of the form $\gamma\wedge\neg\gamma$ derivable in *E*, then $\tau(\Psi) = \tau(\gamma\wedge\neg\gamma) = \tau(\gamma)\wedge\neg\tau(\gamma)$ is a contradiction derivable in $\mathcal{L}_A$ via the proof $\tau(\phi_1), \tau(\phi_2), \ldots, \tau(\phi_n)$. If we choose the translation τ in such a way that it takes the theorems of *E* in $\mathcal{L}_E$ to theorems of *A* in $\mathcal{L}_A$, then τ will show that *A* must be inconsistent (prove a contradiction) if *E* is, which means that *E* is consistent relative to *A*. As Hilbert puts it immediately after indicating how to give the straightforward interpretation based on Ω:

> We conclude from this that any contradiction in the consequences drawn for our axioms [I–V] must also be recognisable in the arithmetic of the domain Ω.[88]

Secondly, note that the definitions of 'point', 'straight line', etc. in $\mathcal{L}_A$ are not used just to fix the interpretations of the primitives; they are now fully analogous to Frege-style definitions, for they are used here *à la* Frege to show that the (translates of) the axioms I–V can be proved 'within the arithmetic of the domain Ω'. The role of the Dedekind Cut or Cauchy sequence constructions could be

[88] *Festschrift*, p. 20, or Hallett and Majer (eds.), *Lectures*, p. 455.

construed analogously in an interpretation of the axioms for the theory of real numbers, and likewise with the 'definitions' of the integers as equivalence classes of ordered pairs of natural numbers, complex numbers as ordered pairs of reals, and so on. And much the same could be done for an interpretation of the notion of cardinal number within a theory of value-ranges governed by Law V. And, of course, the same procedure governs all of Hilbert's models in the course of his investigation of geometries.[89]

In Hilbert's view, the languages, the axioms and the different interpretations or translations are all part of what the mathematician considers in examining, and working with, a theory.

CONCLUSION: REFERENTIAL INDETERMINACY AGAIN

The connection to Frege's concern with referential indeterminacy as it arises in the consideration of 'permutations' is now this. Suppose we have a domain of objects which are taken to be the referents of the primitives in the axioms, say Euclidean points, lines and planes. Suppose we now move to another interpretation, say in a field of numbers. We can, first, regard this as a 'permutation' of objects which leaves the underlying truth-values of the axioms unchanged, just as we had a 'permutation' on the value-ranges before. (See §2, above.) Generalizing this, we can look, not just for isomorphisms, but for maps of one domain into another, not necessarily injective, and where some axioms continue to hold, but in which certain central propositions fail. Hilbert's method is clearly a further generalization, where the underlying language is not taken to be common, and the objects are not assumed to be part of a common domain. Indeed, the concentration on languages and only indirectly on objects makes sense, since it is part of Hilbert's position that there is no theory-independent access to the objects. This was a fundamental platform for the 'syntactic turn' in the foundations of mathematics which Hilbert's work brought about.

The insights on which this rest are by no means new with Hilbert. Translations/permutations of these various sorts were employed in fruitful ways by Pasch and Dedekind, both strong influences on

[89] To get a sample of the inventiveness of the translations Hilbert uses, see Hallett, 'Purity of method', §8.4.

Hilbert. Both used isomorphic interpretations, the latter to point out that the 'special character of the elements' must be left 'entirely out of consideration' in a theory (see p. 445, above), and the former as part of an argument justifying the Duality Principle in projective geometry, in the course of which Pasch presents a strengthened view of his argument that proofs are to be carried through independently of the meaning of the central terms involved in them.[90] Poincaré, too, made important heuristic use of linguistic translations, as part of an explanation of the consistency of non-Euclidean geometry relative to Euclidean geometry, in effect by using a transformation between languages which preserves proof-structure (a 'translation', as he calls it) very similar to both Pasch's and Hilbert's.[91] But while the considerations are not new with Hilbert, he both generalizes various elements of the way they were applied, and makes it the cornerstone of a new, and highly productive, approach to mathematical theories generally.

The novelty of Hilbert's approach is, I think, profound. It led to different views from Frege's, not just on reference and meaning, but also on matters very much related, for example, concerning the conception and place of logic, the uniformity of mathematics, consistency and existence, ideal elements, the relation of number to geometry, and a quite different view of application. Many of these differences were, of course, seen, and objected to, by Frege; nevertheless, they had a profound influence on the direction of foundational research in the twentieth century. It was only in the aftermath of Gödel's Incompleteness Theorems and, later, the independence results in set theory, that there was anything like a revival of the Fregean view.

I have concentrated here on the issue of reference, but there are, of course, other important aspects of the interaction between Frege and Hilbert. Not the least important are Frege's attempts to cast Hilbert's axioms as general Thoughts according to his own lights, and also his attempt to come to terms with the geometrical independence results, an attempt which, both in spirit and detail, is close to the account of Poincaré's which Hilbert generalizes.[92] An important part of Frege's work here is the beginnings of an attempt

[90] See Pasch, *Vorlesungen*, pp. 98–9.

[91] See Poincaré, 'Les géométries non euclidiennes'.

[92] For discussion, see A. Antonelli and R. May, 'Frege's new science', *Notre Dame Journal of Formal Logic*, 41 (2000), pp. 242–70; Tappenden, 'Frege on axioms'.

to articulate foundations for logic itself, in which the formal aspects of logic are both brought out (something which is intrinsic to the considerations of Pasch, Poincaré and Hilbert just mentioned), and circumscribed. Logic was studied intensively in the decades following the Frege–Hilbert debate, particularly in the Hilbert school, where it was treated axiomatically, in an un-Fregean way, on a par with, and open to the same kind of axiomatic analysis as, say, geometry. But this, like many others touched on here, is a topic for further study.[93]

[93] This paper is a much shortened version of a longer piece. Much of the material was first presented at a conference on the philosophy of science at the Inter-University Centre in Dubrovnik in April 1998, and then in May 1998 at a conference on philosophy of mathematics organized jointly by Pittsburgh and Carnegie Mellon Universities. Subsequent outings were given at a workshop on Hilbert in Göttingen in June 1999, to the Philosophisches Seminar at the University of Dortmund in July 1999 and at a conference at Notre Dame University in March 2001. Some new material was added for presentation to a conference in honour of Bill Demopoulos at the University of Western Ontario in May 2008. I am grateful to participants at all these meetings for discussion and comments, especially to Bill Demopoulos, who recently worked through the extended version. Earlier versions were read by Emily Carson and Stephen Menn, to whom I wish to express thanks. I also wish to express my gratitude to the editors of this Companion for their tolerance, advice and extraordinary patience, and to my fellow general editors of Hilbert's unpublished lectures on foundational subjects,. William Ewald, Ulrich Majer and Wilfried Sieg, for discussions over many years on these and other issues. I would also like to acknowledge the generous support of the Social Sciences and Humanities Research Council of Canada over many years, as well as the FQRSC of Québec, formely FCAR. This paper is dedicated to the memory of George Boolos, Frege scholar extraordinaire.

PETER MILNE

12 Frege's folly: bearerless names and Basic Law V

1 FREGE ON TRUTH

Frege tells us surprisingly little about truth. And some of what little he does say, he repeats:

> One can, indeed, say: 'The thought that 5 is a prime number is true'. But closer examination shows that nothing more has been said than in the simple sentence '5 is a prime number'. The truth claim arises in each case from the assertoric sentence, and when the latter lacks the usual force, e.g., in the mouth of an actor upon the stage, even the sentence 'The thought that 5 is a prime number is true' contains only a thought, and indeed the same thought as the simple '5 is a prime number'.[1]

Whereas in the much later 'Thoughts' from 1918 it is sameness of thought/content that is emphasized:

> It is also worth noticing that the sentence 'I smell the scent of violets' has just the same content as the sentence 'It is true that I smell the scent of violets'.[2]

in the 1897(?) 'Logic' we find the emphasis is on assertion:

> If I assert that the sum of 2 and 3 is 5, then I thereby assert that it is true that 2 and 3 make 5.[3]

Strictly, the claim about assertion is distinct from that about content (sameness of thought). It could be that it is a fact about assertion, e.g., that it aims at truth, as people say, that nothing different is accomplished in asserting that the sum of 2 and 3

[1] 'On sense and reference', *CP*, p. 164.
[2] 'Thoughts', *CP*, p. 354.
[3] 'Logic', *PW*, p. 129.

is 5 and that it is true that the sum of 2 and 3 is 5, rather than a fact about the contents asserted. But this is not Frege's view. In both 'On sense and reference' and 'Thoughts' Frege points to a certain redundancy where truth is concerned. For any assertoric sentence *p*, *p* and 'It's true that *p*' have 'just the same content' ('Thoughts'), they 'contain the same thought' ('On sense and reference'). Nevertheless, in contemporary terms, Frege is no deflationist. He asks rhetorically, 'And yet is it not a great result when the scientist after much hesitation and laborious researches can finally say, "My conjecture is true"?'

A true thought refers to the True; properly speaking, truth is not a property of true thoughts ('On sense and reference', p. 164; 'Thoughts', pp. 354–5). Frege is, to modern eyes, surprisingly lax regarding the difference between a truth-operator and a truth-predicate. While truth may not be a property of true thoughts – they refer to the True rather than falling under some concept – it is hard to see a truth-predicate as doing anything other than picking out some property common to all true sentences: to be sure, a language-relative property, and one possessed derivative upon the sentences expressing (in the context of use) a thought that refers to the True.

As is well known, Frege rejects the correspondence theory of truth. The content of the word 'true' is, he says, *sui generis* and indefinable; the meaning of the word 'true' seems to be altogether *sui generis* ('Thoughts', pp. 353, 354). Here's the best gloss I can put in his remarks about truth.[4]

There should be no difference at all between asking, for suitable *p*, whether *p* and whether it's true that *p* (e.g., there is no difference between asking whether Jena is a city on the Saale and asking whether it is true that Jena is a city on the Saale). This gets us to Frege's fundamental thought: that knowledge of what it takes for a particular sentence to be true cannot be something added on after an understanding of the sentence itself. An understanding of what it is to be true ('in general') cannot come after an understanding of the rest of the language, as would be possible if it were possible to offer a proper *definition* of truth.

[4] I owe some of this to Luis Fernández Moreno, 'Die Undefinierbarkeit der Wahrheit bei Frege', *Dialectica*, 50 (1996), pp. 25–35.

Frege says: 'If I do not know that a picture is meant to represent Cologne Cathedral then I do not know what to compare the picture with in order to decide on its truth.' In the case of a declarative sentence there is *never* an equivalent problem, *if one understands the sentence*, i.e. *if one knows the thought it expresses*. If one understands a sentence then one knows how the world must be for (the thought expressed by) the sentence to be true.

A thought is something for which the question of truth can properly be raised. The *only* thing for which the question of truth can properly be raised is the sense expressed by an assertoric or interrogative sentence, so this is a thought. 'Truth does not consist in correspondence of the sense with something else' ('Thoughts', p. 353). Although it is not at first obvious, this connects with something Frege says much later on in 'Thoughts': 'What is a fact? A fact is a thought that is true' ('Thoughts', p. 368). This is, at first glance, a very odd conception of facts. Thought through carefully, it may not be compelling but it does tie in with how we should think about Fregean thoughts.

Think of declarative sentences in a language (or perhaps some of the stuff inside your head) as 'representations'. One uses them to represent 'ways the world might be', to speak very loosely. In asserting a declarative sentence one is claiming things *are* that way. What is the *content* expressed by a sentence, the content of what does the representing? It is what is represented: a way things might be. When is the representation correct? When is the sentence true? – When the way it represents things as being is the way (or, perhaps better, a way) things are. Put another way, a Fregean thought is *not* a 'picture' or representation of a way things might be; rather, it is the way they are represented as being.

Now, if one ascribes truth to sentences, then, as said above, a sentence is true if the way it represents things as being is a way things are. But if, like Frege, one ascribes truth primarily to thoughts, then it is what is represented, not the representation, that is true (or false). And the various ways things are are just some among the way things might be. They're the true ones, the facts. There is no separation between truths, true thoughts, and facts. Facts and truths are the same things, namely – staying with the way I have been speaking – ways things are. Thus Frege ends up with what by his own lights is a perfectly good correspondence theory: thoughts are true if they are identical with, i.e. correspond perfectly to, facts. Frege is an identity theorist of sorts.[5]

This argument turns on Frege (i) taking, if we may speak loosely, truth to be a property of what is represented, not of what does the representing, and (ii) treating what I have called ways things might be fully on a par with what I have called ways things are. Now, one may reasonably object to an identity theory of truth of Frege's sort, a theory that identifies facts with true propositions/thoughts, that the role of the world in determining which propositions are true has been lost. It's all very well identifying facts and true propositions but what is it that fixes the facts? We may resolve that matter if the world is, to borrow a phrase, the totality of facts, but only by granting an explanatory priority to facts that is not Frege's. Truth is, for Frege, unanalysable. Jena falls under the concept *city on the Saale* because the thought *That Jena is a city on the Saale* refers to the True. That is the order of explanation, not, if we take Frege at his word, the reverse. We may limn the laws of truth – the laws of logic – but we cannot give a general account of what it takes for a thought to be true.[6]

What I have just elaborated is the picture that emerges most clearly from 'Thoughts', a late post-paradox piece that contains Frege's only extended discussion of truth – compressed as it may be – in an article that covers an awful lot of ground. How much or how little can be projected back to the 1890s is far from clear. On the one hand, there is the complete absence of any mention of concepts in 'Thoughts'. On the other, in the close to contemporaneous 'Notes

[5] Cf. Julian Dodd and Jennifer Hornsby, 'The identity theory of truth: Reply to Baldwin', *Mind*, 101 (1992), pp. 319–22.

[6] Russell objected to coherence theories of truth that there could be more than one maximal consistent set of beliefs. Likewise, there can be more than one maximal consistent set of Fregean thoughts and it would seem that the world must somehow play a role in determining the set which is the set of true thoughts (facts). Perhaps because he was concerned as much with mathematical truths as empirical ones, Frege says little regarding contingency and modality generally.

Crispin Wright has objected persuasively that coherence theories have a hard job in accommodating contingency say in the sense of genuine chance events ('Truth: A traditional debate reviewed', in S. Blackburn and K. Simmons (eds.), *Truth* (Oxford: Oxford University Press, 1999), pp. 203–38. at pp. 221–2). That is not the objection here. The objection here to a Fregean identity theory is more fundamental: it is that we are given no clue as to what fixes the set of truths. The coherence-theorist has a story to tell about that; Wright's complaint is that that story is defective.

for Ludwig Darmstaedter' (as the editors of the *Nachlass* call it), we find doctrines of long standing reiterated and others portrayed as being of long standing.

What is distinctive about my conception of logic is that I begin by giving pride of place to the content of the word 'true', and then immediately go on to introduce a thought as that to which the question 'Is it true?' is in principle applicable. So I do not begin with concepts and put them together to form a thought or judgement; I come by the parts of a thought by analysing the thought.

… The first thing that strikes us here is that a thought is made up out of parts that are not themselves thoughts. The simplest case of this kind is where one of the two parts is in need of supplementation and is completed by the other part, which is saturated: that is to say, it is not in need of supplementation. The former part then corresponds to a concept, the latter to an object (subsumption of an object under a concept).[7]

What is clear is that Frege's *practice* in *Die Grundgesetze der Arithmetik* is quite different. In §31 he expounds what purports to be a proof that every proper name (including sentences) and first-level function name formed according to the procedures he lays down in §30 has a reference. To carry out the proof he proceeds 'recursively', starting with the primitive signs, using the explanations previously given of them, and moving on to more complex expressions. This 'proof' immediately precedes a section with the title 'Every proposition of *Begriffsschrift* expresses a thought', the first sentence of which summarizes what has, supposedly, been accomplished in §31: 'In this way it is shown that our eight primitive names have denotation, and thereby that the same holds good for all names correctly compounded out of them.' As the system of *Grundgesetze* is inconsistent, the proof must fail. The nub is the

[7] 'Notes for Ludwig Darmstaedter', *PW*, pp. 253–4. Missing from the Darmstaedter notes is any mention of extensions of concepts. The notes end with brief remarks on numerical quantifiers – expressions of the form 'There are *n* …' – which on Frege's reckoning are second-level concepts. The closing lines are:

But still we do not have in them the numbers of arithmetic; we do not yet have objects, but concepts. How can we get from these concepts to the numbers of arithmetic in a way that cannot be faulted? Or are there simply no numbers in arithmetic? Could the numerals help us to form signs for these second level concepts and yet not be signs in their own right? (*ibid.*, p. 257) use of second-order quantifiers in forming predicates in which only first-order variables are free.[8]

Little as he says about truth, Frege says far less about the nature of falsity. A false thought is one that refers to the False. Introducing his symbol for the negation function in *Grundgesetze*, he tells us:

We need no special sign to declare a truth-value to be the False, so long as we possess a sign by which either truth-value is changed into the other.[9]

The thought here is that a sentence (or the thought it expresses) is false if, and only if, its negation is true.[10]

2 ABOUTNESS

Frege's claim that sentences containing non-referring singular terms are neither true nor false follows, in his eyes, from what I think of as Frege's thesis about *aboutness*:

If words are used in the ordinary way, what one intends to speak of is their reference. ('On sense and reference', p. 159)[11]

There is an unambiguous statement of this thesis in the notes Carnap took at Frege's lectures:

A proper name has

1) a meaning [reference]: the thing about which something is said;
2) a sense that is part of the thought.[12]

[8] See Michael Dummett, *Frege: Philosophy of Mathematics* (London: Duckworth, 1991), pp. 214–22, for a detailed analysis.

[9] *Gg*, vol. I, §6.

[10] In §5 Frege has introduced the judgement stroke which indicates assertion, the acknowledgement of a thought as true. To say that no special sign is needed in the case of falsity is to say that acknowledgement of a thought as false can be effected by acknowledging its negation as true (and likewise the work done by denial can be accomplished by assertion of the negation).

[11] I have silently restored 'reference' and its cognates for the translation of '*Bedeutung*' and the like here and in subsequent quotations.

[12] E. Reck and S. Awodey (eds.), *Frege's Lectures on Logic: Carnap's Student Notes, 1910–1914* (Chicago: Open Court, 2004), p. 148.

Given that, there's nothing a sentence containing a bearerless name is about, hence nothing for what it predicates to be true or false of.

Take a sentence containing a Fregean proper name, i.e. either a proper name or a definite description. Replace all occurrences of said name or description by some place-holding marker, say 'ξ'. We have what Frege calls a (first-level) concept-word, an expression that refers to a concept. First-level concepts are functions from objects to truth-values. Going back to the original sentence, the Fregean proper name refers to an object, if it refers at all. If it refers, its reference is an argument of the function. If it refers, the sentence is true or false as the value of the function for that argument is the True or the False. If it does not refer, no argument is selected, hence the function takes no value: no input, no output.[13]

This is what leads Frege to maintain that sentences containing bearerless proper names are neither true nor false. In a discussion of the name 'Odysseus', which he holds may be bearerless, he says,

> Is it possible that a sentence as a whole has only a sense, but no reference? At any rate, one might expect that such sentences occur, just as there are parts of sentences having sense but no reference. And sentences which contain proper names without reference will be of this kind. The sentence 'Odysseus was set ashore at Ithaca while sound asleep' obviously has a sense. But since it is doubtful whether the name 'Odysseus', occurring therein, has reference, it is also doubtful whether the whole sentence has one. Yet it is certain, nevertheless, that anyone who seriously took the sentence to be true or false would ascribe to the name 'Odysseus' a reference, not merely a sense; for it is of the reference of the name that the predicate is affirmed or denied. Whoever does not admit the name has reference can neither apply nor withhold the predicate. ('On sense and reference', p. 162)

In one published article from 1897, and in posthumously published writings dated 1897 and 1914 by the editors (although neither date may be reliable), the connection made between 'aboutness' and lack of truth-value is rendered quite transparently, this time with the name 'Scylla':

> In poetry and legend … there occur sentences which, although they have a sense, have no reference – like, e.g., 'Scylla has six heads'. This sentence

[13] Cf. Susan Haack, *Philosophy of Logics* (Cambridge: Cambridge University Press, 1978), p. 212; and Scott Lehmann, 'Strict Fregean free logic', *Journal of Philosophical Logic*, 23 (1994), pp. 307–36.

is neither true nor false since, for it to be one or the other, it would have to have a reference; but no such reference is available, because the proper name 'Scylla' designates nothing.[14]

The sentence 'Scylla has six heads' is not true, but the sentence 'Scylla does not have six heads' is not true either; for it to be true the proper name 'Scylla' would have to designate something. ('Logic', pp. 129–30)

And when we say 'Scylla has 6 heads', what are we talking about? In this case nothing whatsoever; for the word 'Scylla' designates nothing. Nevertheless we can find a thought expressed by the sentence, and concede a sense to the word 'Scylla'.[15]

Likewise, back with Odysseus, in the 1906 diary entries that form 'Introduction to Logic':

Proper names are meant to designate objects, and we call the object designated by a proper name its reference. On the other hand, a proper name is a constituent of a sentence, which expresses a thought. Now, what has the object got to do with the thought? We have seen from the sentence 'Mont Blanc is over 4000 m high' that it is not part of the thought. Is then the object necessary at all for the sentence to express a thought? People certainly say that Odysseus is not an historical person, and mean by this contradictory expression that the name 'Odysseus' designates nothing, has no reference. But if we accept this, we do not on that account deny a thought-content to all the sentences of the *Odyssey* in which the name 'Odysseus' occurs. Let us just imagine that we have convinced ourselves, contrary to our former opinion, that the name 'Odysseus', as it occurs in the *Odyssey*, does designate a man after all. Would this mean that the sentences containing the name 'Odysseus' expressed different thoughts? I think not. The thoughts would strictly remain the same; they would only be transposed from the realm of fiction to that of truth. So the object designated by a proper name seems quite inessential to the thought-content of a sentence which contains it. To the thought-content! For the rest it goes without saying that it is by no means a matter of indifference to us whether we are operating in the realm of fiction or of truth.[16]

[14] 'On Mr Peano's conceptual notation and my own', *CP*, p. 241.

[15] 'Logic in mathematics', *PW*, p. 225.

[16] 'Introduction to logic', *PW*, p. 191. I cite this lengthy passage in full because it is so very much at odds with the reading given by Gareth Evans, *The Varieties of Reference* (Oxford: Oxford University Press, 1982), pp. 29–30) and by John McDowell ('Truth-value gaps', in McDowell, *Meaning, Knowledge, and Reality*

For Frege the truth-functional logical connectives are literally functions (or, if you think of connectives as linguistic, refer to functions). They are functions whose values are truth-values. For reasons we shall briefly examine below, Frege demands that functions defined for any objects be defined for all, so we cannot say that the connectives are functions whose arguments are truth-values; we can, however, say that it is only for arguments that are truth-values that we need to take note of the values they assign. Negation maps the True to the False and the False to the True; conjunction maps the pair <the True, the True> to the True, the pair <the True, the False> to the False, and so on. As a consequence of this understanding of the connectives, a sentence containing any sentential clause in a direct/non-oblique context that contains a bearerless proper name must lack a truth-value (must fail to refer).

3 THE LOGICAL PROBLEM OF BEARERLESS NAMES

Frege holds that *any* sentence containing a bearerless name in a direct/non-oblique context is neither true nor false. That is the completely general thesis advanced in the quotation from 'On sense and reference'. He terms the thought expressed by such a sentence 'fictitious' and a 'mock thought' ('Logic', p. 130); they are such exactly and only in that they fail to be about actually existing objects. In particular, he says

'Scylla has six heads' is not true

and

'Scylla does not have six heads' is not true.

Lack of a bearer for a singular term spreads lack of truth-value pervasively to logically complex sentences. What holds for negation applies equally to the other familiar connectives. We can set out the Fregean picture in what look like truth-tables for three-valued logic:[17]

(Cambridge, Mass.: Harvard University Press, 1999), pp. 212–13) to Frege's talk of 'mock proper names' and 'mock thoughts' in the 1897 piece 'Logic' as to encourage me in the belief that they have simply misread Frege. Cf. David Bell, 'How "Russellian" was Frege?', *Mind*, 99 (1990), pp. 267–77, §4.

[17] Cf. Timothy Smiley, 'Sense without denotation', *Analysis*, 20 (1960), pp. 125–35.

FREGEAN TRUTH TABLES

A	**¬A**
T	F
–	–
F	T

			B	
	A&B	T	–	F
	T	T	–	F
A	–	–	–	–
	F	F	–	F

			B	
	A∨B	T	–	F
	T	T	–	T
A	–	–	–	–
	F	T	–	F

			B	
	A→B	T	–	F
	T	T	–	F
A	–	–	–	–
	F	T	–	T

			B	
	A↔B	T	–	F
	T	T	–	F
A	–	–	–	–
	F	F	–	T

Beware! The bar is not a third truth-value; it signifies the absence of a truth-value. Where both **A** and **B** have truth-values, the connectives behave classically.

Lack of truth-value bothered Frege, his reason being that it subverts classical logic.[18] Going by the truth-tables above and taking for granted that a valid inference transmits truth from premises to conclusion, Frege was right to be bothered. Some familiar natural deduction rules fail:

> ∨-introduction, →-introduction (conditional proof), *reductio ad absurdum*, *ex falso quodlibet*, the law of excluded middle.

On the other hand, enough of classical logic survives for Frege to be in deep trouble, very deep trouble. Various classical equivalences still hold:

> $P \leftrightarrow Q$ and $\neg P \leftrightarrow \neg Q$; $\neg\neg P$ and P;

and we still have this rule:

> from $P \leftrightarrow Q$ and $Q \leftrightarrow R$ infer $P \leftrightarrow R$.

Albeit that numerous familiar rules fail to be uniformly truth-preserving, there is a simple criterion of validity in this setting:

the inference from premises Σ to conclusion φ is valid iff (i) it is classically valid and (ii) no proper name occurs in φ that does not occur in at least one member of Σ.

[18] See, e.g., 'Function and concept', *CP*, p. 148; *Gg*, vol. II, §165.

There are two routes to trouble. The first adopts (and adapts) an argument due to Herbert Heidelberger.[19]

The indirect argument

With the valid equivalences and the rule noted above in play, we can do this:

> It's true that *P* if, and only if, *P*
> So, it's not true that *P* if, and only if, not-*P*
> But it's true that not-*P* if, and only if, not-*P*
> And it's true that not-*P* if, and only if, it's false that *P*
> Hence it's not true that *P* (if, and) only if it is false that *P*.

And:

> It's false that *P* if, and only if, it's true that not-*P*
> And it's true that not-*P* if, and only if, not-*P*
> So, it's not false that *P* if, and only if, not-not-*P*
> Thus, it's not false that *P* if, and only if, *P*
> But it's true that *P* if, and only if, *P*
> Hence it's not false that *P* (if, and) only if it is true that *P*.

Putting that all together we get,

> It's not true that *P* and it's not false that *P* only if it's both true that *P* and false that *P*.

In short, everywhere we think there's a truth-value gap, there's also a 'glut'! (And vice versa!)

In reaching this conclusion we have used a little logic and Frege's claim about the sameness of thought expressed by *P* and 'It's true that *P*'. Is the little logic used sound with respect to the Fregean truth-tables? Well, if *P* is neither true nor false the biconditionals above are all neither true nor false. But that's not really germane. What matters is that in asserting that a sentence containing a bearerless name is neither true nor false Frege surely intends to say

[19] Herbert Heidelberger, 'The indispensability of truth', *American Philosophical Quarterly*, 5 (1968), pp. 212–17.

something true: he asserts it, so, by his own lights has judged it to be true, not truth-value-less. Now, as the transitions licensed by the biconditionals above are truth-preserving (even if the biconditionals themselves are neither true nor false), we can indeed claim that there is a truth-preserving inference from the supposedly true

> It's not true that *P* and it's not false that *P*

to the contradictory

> It is both true that *P* and false that *P*.

The latter is certainly contradictory for it expresses the same thought as

> *P* and not-*P*.

Heidelberger's argument is perhaps not as well known as it should be. It's not a knock-down argument that any theory that acknowledges truth-value gaps must acknowledge all instances of gaps as being simultaneously instances of truth-value gluts. One needs to know what logical principles are in play.[20] In Frege's case enough is in play to use at least a variant of the argument: the claim that a sentence containing a bearerless name is neither true nor false is contradictory, provably so even in a logic that allows for gaps as profligate as those of the Fregean truth-tables.

The direct argument

A step taken in the course of the indirect argument suffices to establish the incoherence of Frege's claims about sentences containing bearerless names and the thoughts they express. It is a step that

[20] There are non-standard logics in which biconditionals do not contrapose (see, e.g., Richard Holton, 'Minimalism and truth-value gaps', *Philosophical Studies*, 97 (2000), pp. 137–68, at pp. 154–5, for an application to present subject matter) and logics in which the negations of logically equivalent formulas need not be logically equivalent (such as Nelson's Logic of Constructive Falsity and Priest's Logic of Paradox).

The observation regarding biconditionals is something of a red herring. When we reconstruct the argument in terms of inferences, what matters is the rule of proof inversion (a weak form of *reductio ad absurdum*): if *A* entails *B* then not-*B* entails not-*A*. Now, true enough, this rule will not hold in general when bearerless names give rise to sentences that are neither true nor false, but the instances that we need do preserve truth-preservation.

Frege ought to have considered, for it turns on the answer to the simple question, what is the difference between it's not being true that *P* and not-*P*'s being true? For the Fregean there can be none.

(1) By the truth-equivalence, *P* and 'It's true that *P*' express the same thought.
(2) By the functional understanding of negation, not-*P* and 'It's not true that *P*' must therefore express the same thought.
(3) By the truth-equivalence, not-*P* and 'It's true that not-*P*' express the same thought.
(4) Therefore, 'It's not true that *P*' and 'It's true that not-*P*' express the same thought.

Crispin Wright says, '[T]he equivalence schema entails, given only the most basic assumptions about its scope and about the logic of negation, that truth and negation commute as prefixes'.[21] More narrowly, we have used only claims about sameness of meaning (thought expressed) to obtain the same conclusion. Frege wants 'It's not true that *P*' to be TRUE when *P* contains a bearerless name; and, at the same time, he wants 'It's true that *P*' to say the same as *P* (and 'It's true that not-*P*' to say the same as not-*P*) even though, in virtue of containing a bearerless name, *P* is, he wants to say, NOT TRUE and NOT FALSE. Now, the very fact that Frege tells us so little about falsity, and what he does tell us is exactly that judgement of a thought as false is accommodated by judging its negation to be true, shows us that he takes 'It is false that *P*' and 'It's true that not-*P*' as ways of expressing the same thought. But if he is right about this then he cannot coherently maintain of any thought that it is neither true nor false, for 'It's true that not-*P*' is entailed by, indeed *says the same as*, 'It's not true that *P*'.

The following constitute an inconsistent triad (which we may call 'Frege's trilemma'):

(i) The truth-equivalence
(ii) The functional reading of negation
(iii) The truth-value gap thesis concerning the thoughts expressed by sentences containing bearerless names.

[21] Wright, 'Truth', p. 213.

In the *Notes Dictated to G.E. Moore in Norway, April 1914* Wittgenstein states as a *definition*:

p is false = ~(*p* is true) Def.[22]

In a Tarskian, recursive definition of truth we standardly have the clause

~*p* is true if, and only if, *p* is not true.

Such stipulations threaten not just the functional understanding of negation but bring pressure to bear on the very foundation of Frege's function/argument analysis of propositions. For suppose *P* has the form *F*(*a*) where the name *a* does not refer. Then *P*'s negation has the form ¬*F*(*a*) and is true, since *P* is not true. But the function denoted by ¬*F*(ξ) does not, on its own, name either the True or the False, and yet *a* supplies no argument for it. To avoid this consequence, it would seem that one must give up the functional understanding of the logical constants.

Otherwise, one must give up either the truth-equivalence or deny the existence of truth-value gaps. Michael Dummett gives up the former.

A Dummettian interlude

We have seen that (i), (ii) and (iii) are inconsistent. Dummett has argued that (i) and (iii) are inconsistent. The argument is given originally in his article 'Truth'. It has been endorsed by many. Simon Blackburn and Keith Simmons, in the Introduction to their collection, *Truth*, rehearse it and wield it fiercely without further ado. Richard Holton has said of it that it is as damaging as it is simple. It's certainly simple. Here's the argument:

Suppose that *P* contains a singular term which has a sense but no reference: then, according to Frege, *P* expresses a proposition which has no truth-value. This proposition is therefore not true, and hence the statement

22 Ludwig Wittgenstein, *Notebooks 1914–1916*, ed. G. H. von Wright and G. E. M. Anscombe, trans. G. E. M. Anscombe, 2nd edn (Oxford: Basil Blackwell, 1979), p. 116.

'It is true that *P*' will be false. *P* will therefore not have the same sense as 'It is true that *P*', since the latter is false while the former is not.[23]

As it stands this argument is hardly compelling. It is an argument in the logician's sense: it has premises; it has a conclusion. What connects them is a premise that Dummett has endorsed time and again: that 'It's true that …' is an *oratio obliqua* context,[24] an oblique, opaque or indirect context. 'It's true that *P*' is to be read as predicating truth of the thought that is the reference of 'That *P*'. There being no failures of reference in indirect contexts, 'It's true that *P*' cannot be neither true nor false.

Dummett admits that the context governed by 'It's true that …', unlike, say, propositional attitude contexts, fails the standard substitution test for opacity. That test, however, he takes as being only a sufficient criterion. What is at stake here is the way we should read 'It's true that *P*'. As I have said, Dummett reads it as predicating truth of the thought referred to by the name 'That *P*'. This is to be contrasted with how we read a sentence of the form 'It is not the case that *P*'. Here 'It's not the case that …' attaches, as an operator to the sentence *P*.[25] Why can we not read 'It's true that *P*' analogously (for surely 'It is the case that …' should be like *both* 'It is true that …' *and* like 'It is not the case that …')? Before we come to Dummett's response to that question let's ask another. Does Frege concur with Dummett's reading?

Nothing Frege says encourages the thought that he does. There are substantial reasons to think that he does not. Here's one. The passages quoted from 'On sense and reference' that give us theses (i) and (iii) occur before any mention of oblique contexts (more properly, of the customary/indirect distinction for sense and reference). It would be disingenuous in the extreme, not to say outright dishonest, of Frege to use a locution that requires that distinction for its

[23] Michael Dummett, 'Truth', in his *Truth and Other Enigmas* (London: Duckworth, 1978), pp. 1–24. at p. 4.

[24] E.g. *ibid.*, Michael Dummett, *The Interpretation of Frege's Philosophy* (London: Duckworth, 1981), ch. 6, and 'Of what kind of a thing is truth a property?', in S. Blackburn and K. Simmons (eds.), *Truth* (Oxford: Oxford University Press, 1999), pp. 264–81.

[25] Cf. A. N. Prior, 'Oratio Obliqua', *Aristotelian Society Supplementary Volume*, 37 (1963), pp. 115–26, at p. 116, and 'Is the concept of referential opacity really necessary?', *Acta Philosophica Fennica*, 16 (1963), pp. 189–99, at pp. 193–4.

proper interpretation prior to advancing it, the more so as *nowhere* does he ever so much as mention it as an example giving rise to an indirect context. Here's another reason for thinking Frege didn't adopt Dummett's reading. Take a sentence such as 'Jena is a city on the Saale'. The reference of the name 'Jena' is a particular German city. If the sentence is true then so too is the sentence 'Jena exists'. Jena's existence is an existential commitment of that sentence's being true. On Dummett's reading of 'It's true that Jena is a city on the Saale', this second sentence has no such existential commitments, or at least has none such directly, because in this sentence the name 'Jena' now refers to the customary sense expressed by the name 'Jena', the sense expressed in the original sentence, its indirect reference. There is no obvious explanation why the truth of 'It's true that Jena is a city on the Saale' has the existential commitments of 'Jena is a city on the Saale'. Now, surely, Frege, had he intended Dummett's reading, would have realized this and, having realized it, balked at the sameness of thought claim.

Frege does not, I contend, concur with Dummett's reading. Should we? As I read him Dummett presents only one argument that is intended to clinch the claim that we should.[26] It is this:

[I]f there are meaningful sentences which say nothing which is true or false, then there must be *a* use of the word 'true' which applies to propositions; for

[26] In 'Of what kind of a thing is truth a property?', he offers another which he describes as providing a strong reason in favour of the *oratio obliqua* thesis. The argument form

> *X* believes that *P*
> It's true that *P*
> Therefore, *X* has a true belief

is, as Dummett puts it, unquestionably valid. It is also, as he says, unproblematic if we read 'It's true that *P*' as predicating truth of the proposition 'That *P*', i.e., the *oratio obliqua* reading. It may, Dummett concedes, be objected that the form

> *X* believes that *P*
> *P*
> Therefore, *X* has a true belief

is equally valid, but problematic on the view Dummett maintains. The difficulty can be 'localized' and validity explained by allowing inference of 'It's true that *P*' from *P*. But if this fact 'provides strong reason for construing the phrase "it is true that" as inducing an opaque context' (p. 271), it does so at the cost of rendering the validity of the latter inference wholly unexplained: to allow the inference is not to explain it.

if we say 'It is neither true nor false that *P*' the clause 'that *P*' must here be in *oratio obliqua*, otherwise the whole sentence would lack a truth-value.[27]

What reason could there be to believe this conclusion? I suspect that Dummett is making an assumption that we have already seen to be false on pain of contradiction *when the truth-equivalence is accepted*: that if a sentence or thought *P* is neither true nor false so too is its negation. The conclusion of the direct argument shows us that, if *P* is neither true nor false, then, since in particular it is not true, its negation *is* true. Without an assumption to the contrary in play, I cannot see how Dummett's conclusion follows. To show that it does not we must elaborate a coherent position that admits that 'It is neither true nor false that *P*' is true for some sentences *P* consistently with the truth equivalence.

Before that, notice that, if Dummett is to avoid the same morass Frege got himself into, he must deny one of the following (what we might call 'Dummett's trilemma'):

(i) That *A* entails 'It's true that *A*'
(ii) That not-*B* entails not-*A* when *A* entails *B*
(iii) That 'It's true that not-*A*' entails 'It's false that *A*'.

Here I take entailment to be necessary truth-preservation. (i) and (ii) suffice to get from 'It's not true that *P*' to 'It's true than not-*P*'.

A semantic conception of falsity

The Fregean wants to say that an assertoric sentence may be neither true nor false. For this to be possible while endorsing the truth-equivalence, he must *not*, as we have seen, equate being false and not being true. If a sentence is not true that is either because it is false or because it is neither true nor false. Frege's account of falsity, such as it is, fails to allow for this second possibility.

Dummett has argued that the truth-value gap thesis is incompatible with Frege's claim that *P* and 'It's true that *P*' express the same thought/content. I have suggested that Dummett's grounds for this are not compelling. I do, nonetheless, hold that the truth-equivalence and the truth-value gap thesis are incompatible with Frege's conception of how sentences come to be false, i.e., how they come to name

[27] Dummett, 'Truth', p. 5.

the False. What we have been led to this far, via the direct argument, is an account of negation that assigns to it this truth-table:

A	**¬A**
T	F
–	T
F	T

This is a three-valued truth-table in a purely formal sense: as before, '–' stands for neither true nor false, not for some distinct, third *value*.

This is an unorthodox truth-table, one that follows from our Fregean theses (together with the commonplace that the negation of a truth is false), but one which there can be little doubt Frege failed to consider. Why so? Because *P*'s negation has a truth-value even when *P* doesn't. In Fregean terms, even when the assertoric sentence *P* fails to refer either to the True or the False, because containing a non-referring singular term, its negation, which contains exactly the same non-referring term or terms, succeeds in referring to the True. But the negation operator, for Frege, stands for a function, a function that maps the True to the False and everything else to the True (see, e.g., 'Function and concept', pp. 149–50, *Gg*, vol. I, §6). It must, as all functions must according to Frege, be defined for all objects, but the whole point of the passage about Odysseus is that the sentence containing a non-referring term fails to refer, hence can supply no argument for the function for which negation stands to act upon. There is, as we have seen, a fundamental incoherence in Frege's *use* of negation in 'On sense and reference'.

It is important to appreciate exactly which of Frege's semantic theses poses the problem. It is *not* the thought that sentences are proper names referring to the True and the False if to anything, nor the thought, however problematic elsewhere, that the True and the False are objects on a par with tables, chairs and extensions of concepts. Nor is it the thought that logically unstructured predicates refer to concepts and the latter are functions from objects to truth-values. Nor yet is it a mere commitment to compositionality, at least as that is understood in contemporary terms, for there is no failure of compositionality in having a negation satisfying the truth-table above. Where the problem lies is in the supposition that *any* sentence containing a non-referring singular term must itself express a thought that fails to have a reference, i.e. fails to be either true or false. That thesis, entirely plausible taken on its own, no doubt, is incompatible with the conjunction of the truth equivalence and the truth-value gap thesis. Frege holds to this thesis because he has a narrow reading of compositionality in functional terms, from which follows the principle 'no input, no output'.[28]

Holding to the truth-equivalence, we must give up the claim that there are sentences expressing thoughts that are neither true nor false, or the thesis that *any* assertoric sentence containing a non-referring singular term must express a thought that fails to be either true or false. The latter, of course entails the former, so that we cannot give the former up without rejecting the latter. The entailment does not reverse. What *is* the case is that if we give up on the specifically Fregean thesis that any sentence containing a non-referring singular term must itself express a thought that fails to be either true or false, it may seem that we have little reason to continue endorsing the claim that there are sentences expressing thoughts that are neither true nor false. Why should the Fregean continue to suppose some sentences containing non-referring terms are neither true nor false when having to give up on the claim that all are? The best answer, it seems to me, is that anyone with Fregean sympathies would have to give up on the stronger thesis anyway, irrespective of any problems occasioned by his/her treatment of negation.[29]

[28] Evans makes the weaker point that Frege had no means to rule out a 'wide scope' negation with our truth-table. But Evans makes a mistake when he goes on to say:

> I said this was essentially the same point, because it rests upon the incomprehensibility of the idea that the thought that *p* and the thought that *it is not true that p* can both fail to be true. Surely the thought that it is not true that *p* is true just when the thought that *p* is not true. So resistance to the idea that both thoughts may fail to be true is, once again, resistance to the idea of a gap between a determinate thought's failing to have the value True and its having the value False. (Evans, *The Varieties of Reference*, p. 25)

We agree that the thought that it is not true that *p* is true just when the thought that *p* is not true. But that, as we shall show, does not preclude the possibility of *p*'s being neither true nor false.

[29] We should note that the weaker claim is compatible with what Frege lays down in *Grundgesetze* in a section headed 'When does a name denote something?' There he says,

Our Fregean truth-tables for conjunction and disjunction have a quite extravagantly damaging effect on what Christine Tappolet calls truisms about truth.[30] To be fair, her one example – that a conjunction is true if, and only if, its conjuncts are true – survives; but the parallel truism for (inclusive) disjunction, that a disjunction is true if, and only if, at least one of its disjuncts is true, fails, as does the truism that a conjunction is false if, and only if, at least one of its conjuncts is false. Tappolet also proposes it as a truism that truth is what is conserved in valid inference. If so, with these truth-tables the natural deduction rule of ∨-introduction fails, as we noted above. If one supposes that, likewise, that it is a truism that if the conclusion of a valid inference is false so too is at least one premise – and is this any less of a truism? Perhaps it is – then the natural deduction rule of &-elimination (simplification) also fails.

Now, of course, Frege did think that in order to save logic, by which he meant classical logic, a logically perfect language must satisfy the requirement that there be no non-referring singular terms, and that ordinary language is sadly deficient in this respect. Anyone less sanguine than Frege about a wholesale revision of everyday conceptions in favour of the logically perfect, will, I suspect, feel moved to hold on to Tappolet's truisms and so reject the Fregean truth-tables for the sentences of everyday language. Rather, she will endorse these truth-tables:

A proper name has a denotation if the proper name that results from that proper name's filling the argument places of a denoting name of a first-level function of one argument *always* has a denotation, and if the name of a first-level function of one argument that results from the proper name in question's filling the ξ-argument places of a denoting name of a first-level function of two arguments *always* has a denotation, and if the same holds also for the ζ-argument-places. (*Gg*, vol. I, §29, emphasis added)

He has preceded this by saying:

A name of a first-level function of one argument has a denotation (denotes something, succeeds in denoting) if the proper name that results from this function-name by its argument-places being filled by a proper name always has a denotation if the name substituted names something.

[30] Christine Tappolet, 'Truth pluralism and many-valued logics: A reply to Beall', *Philosophical Quarterly*, 50 (2000), pp. 382–5.

NON-FREGEAN TRUISTIC TRUTH TABLES

A	**¬A**
T	F
–	T
F	T

	A&B	**B:** T	–	F
	T	T	–	F
A	–	–	–	F
	F	F	F	F

	A∨B	**B:** T	–	F
	T	T	T	T
A	–	T	–	–
	F	T	–	F

	A→B	**B:** T	–	F
	T	T	–	F
A	–	T	?	–
	F	T	T	T

	A↔B	**B:** T	–	F
	T	T	–	F
A	–	–	?	–
	F	F	–	T

I put a '?' rather than '–' because one may well maintain that *any* instance of $A \rightarrow A$ should be TRUE, but equally clearly not every sentence of the form $A \rightarrow B$ in which both A and B contain non-referring proper names should be TRUE. Arguably it's a truism that if A entails B then 'If A then B' is true; and even more arguably, it's a truism that A entails A. Likewise, if A and B say the same thing, 'If A then B' should be true.[31]

At the same time as endorsing these truth-tables (which merely enshrine truisms),[32] we don't have to give up entirely on the original Fregean perception that leads to the thesis that sentences containing non-referring singular terms are neither true nor false.

Think of simple predications. A predicate refers to a concept; some objects fall under the concept, some don't. A predicate therefore just refers to something that maps objects to the semantic values of sentences, which, for Frege, are truth-values (as Dummett has emphasized: e.g., chapter 6 of *The Interpretation of Frege's Philosophy*). And we need feel no obligation to say that a sentence comprising

[31] Regarding truisms, it has to be admitted that in twentieth-century logic nothing is sacred. Quantum logic has been seen as admitting true disjunctions neither of whose disjuncts need be true. Even the rule of &-elimination (simplification) has been denied: a conjunction may not entail its conjuncts (see, e.g., Robert Gahringer, 'Intensional conjunction', *Mind*, 79 (1970), pp. 259–60, and Bruce Thompson, 'Why is conjunctive simplification invalid?', *Notre Dame Journal of Formal Logic*, 32 (1991), pp. 248–54).

[32] In the case of the conditional, conditionally so: if there's any truth in the material implication account of the conditional.

a simple predication and a non-referring singular term is anything other than neither true nor false.

We must divorce falsity and non-truth. We do this for atomic sentences: for an atomic sentence to be false it must be not true *and* all the singular terms it contains refer. Otherwise it is neither true nor false. Rather than an analogue of the simple truth equivalence, we then proceed to give a recursive definition of falsity, guided by the platitudes enshrined in the non-Fregean truth-tables. We must also give an account of the quantifiers. The reader interested in seeing how this goes may consult the Appendix.

Now, it is a fact to be celebrated that, if we take Tappolet's truism that truth is what is preserved in valid inferences to heart, we find that the (formal) truth-tables above for negation, conjunction and disjunction deliver that all and only classically valid propositional logic inference patterns involving those connectives preserve truth. Employing classically valid inference patterns in conjunction with the truth equivalence, we can then *derive* the standard recursive clauses for truth concerning these connectives – accepting that 'if *A* then *B*' is true when *A* entails *B*. What is more, we also have that an atomic sentence is neither true nor false when, and only when, at least one of its terms fails to refer. And, returning to Dummett's discussion of 'It's true that …' contexts, we find that not only may the thought expressed by such a sentence be neither true nor false, it is then true that it is neither true nor false.

4 STATEMENTS OF NON-EXISTENCE

To paraphrase Leonard Linsky, Frege does not address the analysis of negative existential sentences involving proper names, but we can construct a Fregean account of them.[33] Existence is a higher-level concept, under which concepts of lower level fall, or not: to say that there is at least one φ is to say that at least one object falls under the concept that is the reference of 'φ'. To say that Jena exists is to say that at least one object falls under the concept *identical to Jena*; formally, $\exists x(x = a)$. To say that unicorns don't exist is to say that nothing falls under the concept *unicorn*; formally $\neg\exists x\varphi x$. Likewise, one

[33] Leonard Linsky, *Names and Descriptions* (Chicago: University of Chicago Press, 1977), p. 5.

might expect, to say that Pegasus does not exist is to say that nothing falls under the concept *identical to Pegasus*. But this suggestion falls foul of Frege's general thesis concerning sentences containing non-referring terms: the thought expressed, that Pegasus does not exist, is neither true nor false on Frege's account.[34]

It has struck some as intolerable that this should be a consequence of Frege's theses on non-referring singular terms. Linsky, for example, says that 'Pegasus does not exist' is true, a truth, indeed, that we must insist upon.[35] Now, while common sense balks at Meinong's 'There are objects of which it is true to say that there are no such objects', it happily countenances assertions such as 'There are lots of things that don't exist: Atlantis, El Dorado, the planet Vulcan, Santa Claus, the Abominable Snowman, the Loch Ness Monster, the Big Grey Man of Ben Macdui …'[36] But be that as it may, it hardly sanctions a reflective insistence on the correctness of saying 'Pegasus does not exist'. Of what is one denying existence? Not of Pegasus, for, as one would like to say, the point is that there is no Pegasus of which existence may be denied. There's a real and pressing sense in which the sentence cannot be about Pegasus (even if, in some more attenuated sense of 'about', it is about Pegasus). The name 'Pegasus' does not refer; it expresses a sense but nothing answers to this mode of presentation.

To say that Pegasus does not exist may, in some indirect way, be a claim about the name 'Pegasus' or the sense it expresses. The former is perhaps Frege's view. The latter is Linsky's contention. In negative existential sentences 'exists' 'induces an oblique context in which the proper name denotes its customary sense'.[37] This is, he says, 'a rather satisfying result since it exploits the intuition that existence-contexts are indeed special, and that what prevents

[34] We have found Frege's own account to be thoroughly unsatisfactory. We could with good conscience argue thus: as 'Pegasus' does not refer, 'There is at least one object identical to Pegasus' is not true; therefore its negation is true. For the time being I ask the reader to suspend her dissatisfaction. Dissatisfaction suspended, we note that the Fregean must say that the predicate 'identical to Pegasus' expresses a sense but fails to refer: no concept answers to the sense expressed.

[35] Leonard Linsky, 'Frege and Russell on vacuous singular terms', in M. Schirn (ed.), *Studien zu Frege/Studies on Frege*, vol. III (Stuttgart: Frommann, 1976), pp. 97–115, at p. 112.

[36] If to be is to be the value of a bound variable, existence comes cheap in ordinary usage.

[37] Linsky, *Names and Descriptions*, p. 6.

"Pegasus does not exist" from being meaningless is the fact that the denotationless name "Pegasus" is not devoid of sense'.[38]

Whatever the merits of Linsky's proposal, it has two significant demerits. The first is that in classical logic the sentence φa containing the singular term a is logically equivalent to $\neg\exists x(x = a \mathbin{\&} \neg\varphi x)$. So either we junk classical quantification theory or thoughts about objects have the same truth-conditions (which may, for the Fregean, mean that they express the same thought)[39] as sentences about the modes of presentation of those objects. Neither is a happy place to end up.

The second demerit is that it fails to take account explicitly of what Frege says about existence presuppositions. In 'On sense and reference' we find,

> If anything is asserted there is always an obvious presupposition that the simple or compound proper names used have a reference. If therefore one asserts 'Kepler died in misery', there is a presupposition that the name 'Kepler' designates something; but it does not follow that the sense of

[38] Linsky, 'Frege and Russell', p. 112. Linsky does not restrict the induction of oblique contexts by 'exists' to negative existentials. Some support for Linsky's position might be drawn from Anthony Kenny's observations:

> If a man uses a proper name, then he implies that it has a bearer, that is to say, that the object which he means exists. If someone says 'Satan exists' or 'Satan does not exist' then he does not imply, but respectively asserts or denies, that Satan exists. It follows that he is not using 'Satan' in these sentences as a proper name … [W]hether Satan exists or not, 'Satan' is not used as a proper name either in 'Satan exists' or in 'Satan does not exist'. Neither of these sentences, moreover, is about Satan, whether or not he exists. (Anthony Kenny, : 'Oratio Obliqua', *Aristotelian Society Supplementary Volume*, 37 (1963), pp. 127–146, at p. 141.

But why suppose it is existence statements that are special, why not the embedded identity? It seemed to Bas van Fraassen that 'we cannot plausibly reject that '$t = t'$' is false when t has a referent and t' does not.' He offers the example: that Santa Claus does not exist is sufficient reason to conclude that the president of the US is not Santa Claus ('Singular terms, truth-value gaps, and free logic', *Journal of Philosophy*, 63 (1966), pp. 481–95).

[39] See *Gg*, §32; 'Compound thoughts', *CP*, pp. 393 and 405; Letter to Husserl, 9 December, 1906, *PMC*, pp. 70–1; 'A brief survey of my logical doctrines', *PW*, p. 70; but see Charles Parsons, 'Review article: Gottlob Frege *Wissenschaftlicher Briefswechsel*', *Synthese*, 52 (1982), pp. 325–43, at pp. 328–9, and Jean van Heijenoort, 'Frege on sense identity', *Journal of Philosophical Logic*, 6 (1977), pp. 103–8.

the sentence 'Kepler died in misery' contains the thought that the name 'Kepler' designates something. If this were the case the negation would have to run not

Kepler did not die in misery

but

Kepler did not die in misery, or the name 'Kepler' has no reference.

That the name 'Kepler' designates something is just as much a presupposition for the assertion

Kepler died in misery

as for the contrary assertion. Now languages have the fault of containing expressions which fail to designate an object (although the grammatical form seems to qualify them for that purpose) because the truth of some sentence is a prerequisite. ('On sense and reference', pp. 168–9)

It is, then, a presupposition of the sentence 'Pegasus does not exist' that the name 'Pegasus' designate something, and hence that 'Pegasus exists' is true. Thus the sentence 'Pegasus does not exist' cannot be (truthfully) asserted: one cannot acknowledge it as true for in the very attempt to do so one must accept its contrary. Frege's observation concerning negations of sentences containing singular terms is correct to the extent that we do not normally make explicit existence assumptions when we negate a sentence. (Russell's theory of descriptions is a case in which the distinction between wide and narrow scope negations is in this respect abnormal.)

Taking the observation at face value, Frege has a strong motive for insisting, as he does on numerous occasions, that in a perfected scientific language every properly constructed (closed) singular term refers, for only then is it the case that the language is used in accordance with the presuppositions of (its) correct usage. We can see from the way we say what the negation (or, more generally, the contrary) of a given sentence says that we take for granted that the proper names of our language refer. It is, then, no surprise if we run into difficulties when those presuppositions fail. Frege, never being one to cast doubt on the correctness of classical logic, takes the presuppositions as well founded. A language rid of imperfections must accord with those presuppositions: well-formed proper names of the language must – as a matter of logic! – refer.

5 THE PERFECTED LANGAUGE OF A DEMONSTRATIVE SCIENCE

If we are to use a language in accord with the presuppositions for its use, all singular terms of the langauge must refer:

> A logically perfect language (*Begriffsschrift*) should satisfy the conditions, that every expression grammatically well constructed as a proper name out of signs already introduced shall in fact designate an object, and that no new sign shall be introduced as a proper name without being secured a reference. ('On sense and reference', p. 169)

As Frege indicates here, there are general methods for forming proper names; all such names must be assured of a reference. The best known *variable-binding term-forming operators* (vbtos) generate definite descriptions and set abstracts. Let α be a vbto. Then, for any predicate $\varphi(x)$, $\alpha x\varphi(x)$ is a singular term (in Frege's terminology: a proper name).

In *Grundgesetze*, terms for courses-of-values are introduced this way (in the notation ἐΦε) (*Gg*, §9).[40] What Frege calls his 'substitute for the definite article' is introduced a little differently, as a function mapping objects to objects:

$$\text{if } \xi = \grave{\varepsilon}(\varepsilon = \Delta)\text{, for some object } \Delta \text{ then } \backslash\xi = \Delta;$$
$$\text{otherwise } \backslash\xi = \xi.$$ (*Gg*, §§11 and 31)[41]

The only objects Frege has previously introduced are the two truth-values, the True and the False, and courses-of-values. In §10 he has argued that the True and the False may be identified with courses-of-values. The net effect is therefore the same as introducing a vbto meeting these constraints:

[40] It is tempting to put '$\{x: \varphi(x)\}$', in a more modern notation, for Frege's 'ἐΦε' but that, while largely harmless, would be misleading in that Frege's extension of a concept is closer to the graph of a set's characteristic function than to the set itself.

[41] Frege's Basic Law VI says only: $\forall x(x = \backslash\grave{\varepsilon}(\varepsilon = x))$. That is, the Law says *nothing* about how $\backslash\xi$ is to be interpreted when ξ is not of the form $\grave{\varepsilon}(\varepsilon = \Delta)$ for some properly formed name Δ.

if a unique object falls under the concept $\varphi(\xi)$ then '$\backslash x\varphi(x)$' denotes that object;

otherwise '$\backslash x\varphi(x)$' denotes $\grave{\varepsilon}\varphi\varepsilon$.

From the Fregean perspective, vbtos are second-level functions mapping concepts to objects. Like all functions, they must be defined for all possible arguments – all concepts – and they must be well defined, that is, the function must assign the same object to co-extensive concepts, for co-extension is the analogue for concepts of identity between objects. In the Fregean scheme, the vbto α satisfies these two axioms, where we read the second-order quantifiers as quantifying over concepts:

α-Existence: $\forall X \exists z(z = \alpha y X y)$;

α-Extensionality: $\forall X \forall Y(\forall x(Xx \leftrightarrow Yx) \rightarrow \alpha y X y = \alpha y Y y)$.

We should note that these are trivially consistent with respect to standard second-order semantics, for we may take a one-element domain, $D = \{0\}$, and a function $a{:}\wp(D) \rightarrow D$ which assigns 0 to both subsets of D. We should note too that, as George Boolos observed, α-Extensionality may be considered a logical truth: given extensional semantics, α is interpreted as a function from subsets of the domain to the domain and $\forall x(Xx \leftrightarrow Yx)$ is satisfied under an assignment of values to the second-order variables if, and only if, X and Y are assigned the same subset of the domain.[42] Lastly, we should note that in standard second-order logic α-Existence is a consequence of α-Extensionality as $\forall X \forall x(Xx \leftrightarrow Xx)$ is a logical truth.

6 BASIC LAW V

Introducing courses-of-values, in §3 of *Grundgesetze*, Frege says

I use the words

'the function $\Phi(\xi)$ has the same *course-of-values* as the function $\Psi(\xi)$'

generally to denote the same as the words

'the functions $\Phi(\xi)$ and $\Psi(\xi)$ have always the same value for the same argument'.

[42] George Boolos, 'Frege's theorem and the Peano postulates', *The Bulletin of Symbolic Logic*, 1 (1995), pp. 317–26, at p. 322.

This is Basic Law V informally stated. There is further discussion of courses-of-values in §§9 and 10, and a formal statement in §20. Later on he splits the law into Va and Vb. Va is α-Extensionality for the extension-of-concept vbto ὲ. It is unexceptionable. As Montgomery Furth says, 'This is no news to us; it merely follows from the extensionality of concepts'.[43] Vb is the converse of Va.

Converse of α-Extensionality: $\forall X \forall Y(\alpha y X y = \alpha y Y y \rightarrow \forall x(Xx \leftrightarrow Yx))$.

This says that the vbto α stands for a one–one function from concepts to objects.

7. CANTOR'S THEOREM

In the 1890–1 volume, the first volume, of the *Jahresbericht der deutschen Mathematiker-Vereinigung* Cantor published a proof of what is now widely known[44] as Cantor's Theorem: every set is of lower cardinality than the set of its subsets. Emphasizing its pertinence to Fregean concerns, we note that Cantor's original proof was phased not directly in terms of subsets but in terms of functions defined on a given set and taking only two values (0 and 1).[45]

Cantor's theorem can be phrased in two equivalent ways:

(1) There is no function from a set X onto the set of all its subsets.
(2) There is no one–one function from the set of all subsets of X into the set X.

The orthodox textbook proof of Cantor's Theorem, following Cantor's original, is a proof of (1).

[43] Montgomery Furth, 'Editor's Introduction', in Frege, *The Basic Laws of Arithmetic; Exposition of the System*, ed. and trans. M. Furth (Berkeley and Los Angeles: University of California Press, 1964), pp. xlv–xlvi.

[44] Ernst Zermelo, 'Untersuchungen über die Grundlagen der Mengenlehre I', *Mathematische Annalen*, 65 (1908), pp. 261–81, translated as 'Investigations in the foundations of set theory I' in J. van Heijenoort (ed.), *From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931* (Cambridge, Mass.: Harvard University Press, 1967), pp. 199–215.

[45] Cantor's proof is reproduced in Michael Hallett, *Cantorian Set Theory and Limitation of Size* (Oxford: Oxford University Press, 1984), p. 77.

> Let f be a function from X into the set of all subsets of X. The set $Y = \{x \in X: x \notin f(x)\}$ is a well-defined subset of X. But no element of X is mapped to Y by f. For suppose, to the contrary, that $f(y) = Y$; we ask, Does y belong to $f(y)$? If $y \in f(y)$, *i.e.* if $y \in Y$, then, by the definition of Y, $y \notin f(y)$; conversely, however, if $y \notin f(y)$ then y meets the defining condition for membership of Y and so $y \in Y$, i.e. $y \in f(y)$. Thus $y \in f(y)$ if, and only if, $y \notin f(y)$ – a contradiction.

With a proof of (1) in hand, (2) may quickly be derived via a proof by contradiction.

> Suppose that there is a one–one function h from $\wp(X)$ into X. As it is one–one it has an inverse: for any x in X, there is at most one subset of X mapped to x by h. The inverse maps some not necessarily proper subset Y of X *onto* $\wp(X)$. Pick an arbitrary subset of X and map any remaining members of X – those not in Y – to that subset. We then have a function from X onto $\wp(X)$, in contradiction to (1).

For present purposes it is of more interest to prove (2) directly. We do this twice over.

> Let g be a function from $\wp(X)$ into X. Let Y be the subset $\{x \in X: \exists Z \subseteq X\ [x = g(Z) \text{ and } x \notin Z\ ]\}$, a well-defined subset of X. Let $y = g(Y)$.
>
> I_A If $y \notin Y$ then y satisfies the condition for membership of Y, *i.e.* $y \in Y$. This suffices to establish that $y \in Y$.
>
> I_B But now, think what this, *i.e.* $y \in Y$, says. On the one hand, $y = g(Y)$ and $y \in Y$. On the other, for some subset Z of X, $y = g(Z)$ and $y \notin Z$. Thus g is not one-one as two distinct subsets of X are mapped to y.
>
> II_A Suppose that g is one–one and that $y \in Y$. For some Z, $y = g(Z)$ and $y \notin Z$. But as g is one-one, $Z = Y$. So, if $y \in Y$ then $y \notin Y$. This establishes that $y \notin Y$.
>
> II_B But think what this says: for any subset Z of X that gets mapped to y by g, y is a member of Z. But Y is one such and y, as just demonstrated, is not a member of it. Contradiction. – g cannot be one–one.

We have here *three* proofs: $I_A + I_B$; $II_A + II_B$; and $I_A + II_A$, for the latter two give us, respectively, $y \in Y$ and $y \notin Y$, a contradiction, on the assumption that g is one–one.

To bring out the role of the clause '*Y* is a well-defined subset of *X*', here's a neat little exercise in set theory:

> The co-finite subsets of $\mathbb{N}$ are those subsets of $\mathbb{N}$ whose complements with respect to $\mathbb{N}$ are finite. The set of all finite and co-finite subsets of $\mathbb{N}$ is countably infinite. Let $X_0, X_1, \ldots, X_n, \ldots$ be some enumeration of this set. Show that the set
>
> $Y = \{n \in \mathbb{N}: n \in X_n\}$
>
> is neither finite nor co-finite.
>
> PROOF Consider the complement of Y, $\mathbb{N} - Y = \{n \in \mathbb{N}: n \notin X_n\}$, which is finite or co-finite as Y is co-finite or finite. If it is either finite or co-finite then $\mathbb{N} - \mathrm{Y} = X_m$, for some $m \in \mathbb{N}$. But then, $m \in X_m$ if and only if $m \in \mathbb{N} - \mathrm{Y}$ if and only if $\mathrm{m} \notin X_m$. So $\mathbb{N} - \mathrm{Y}$, and hence Y itself, is neither finite nor co-finite.

The point is that here we can have a one–one correspondence between $\mathbb{N}$ and the family of finite and co-finite subsets of $\mathbb{N}$, exactly because the 'diagonalizing set' $\mathbb{N} - Y$ doesn't belong to that family.

Cantor's Theorem, published at around the time Frege was finishing the writing of the first volume of *Grundgesetze*, provides a stark warning. Sadly, it was a warning to which Frege was blind.

8 THE PARADOX

With a vbto α in our (second-order) language, we may form the predicate containing one free variable

$$\exists \mathrm{X}(x = \alpha y X y \mathbin{\&} \neg X x),$$

which we shall abbreviate as $\Psi(x)$. On the assumption that it is a suitable substituend for the second-order quantifiers, i.e. on the assumption that this predicate does refer to a concept, we can show, as a matter of logic, that $\Psi(\alpha y \Psi(y))$. We shall conduct the proof in a weak second-order free logic, so that we explicitly mark assumptions that a formula is an appropriate substitution instance for a quantifier – semantically, that a formula does refer to an entity falling within the range of the quantifiers, be that at first- or second-order.[46] We take the natural deduction rules for quantifiers and identity in free logic from Tennant.[47] The proof looks like this:

$$\dfrac{\forall X \exists z(z = \alpha y X y) \quad \exists!\psi}{\exists z(z = \alpha y \psi(y))}\ \forall_2\text{-E}$$

$$\dfrac{\overline{t = \alpha y \psi(y)}^{\,1} \quad \overline{\exists! t}^{\,1}}{\exists! \alpha y \psi(y)}\ = \text{-E}$$

$$\dfrac{\exists z(z = \alpha y \psi(y)) \qquad \exists! \alpha y \psi(y)}{\exists! \alpha y \psi(y)}\ 1\ \exists_1\text{-E}$$

$$\dfrac{\exists! \alpha y \psi(y)}{\alpha y \psi(y) = \alpha y \psi(y)}\ = \text{-I} \qquad \overline{\neg\psi(\alpha y \psi(y))}^{\,2}$$

$$\dfrac{\alpha y \psi(y) = \alpha y \psi(y) \qquad \neg\psi(\alpha y \psi(y))}{\alpha y \psi(y) = \alpha y \psi(y)\ \&\ \neg\psi(\alpha y \psi(y))}\ \&\text{-I}$$

$$\dfrac{\alpha y \psi(y) = \alpha y \psi(y)\ \&\ \neg\psi(\alpha y \psi(y)) \qquad \exists\psi}{\exists X(\alpha y \psi(y) = \alpha y X y\ \&\ \neg X(\alpha y \psi(y)))}\ \exists_2\text{-I}$$

$$\exists X(\alpha y \psi(y) = \alpha y X y\ \&\ \neg X(\alpha y \psi(y)))$$
$$=========================\ \text{Definition}$$
$$\dfrac{\psi(\alpha y \psi(y))}{\psi(\alpha y \psi(y))}\ 2\ \text{CRA}$$

From α-Existence and the referential assumption ∃!Ψ – that the predicate $\exists X(x = \alpha y X y\ \&\ \neg X x)$ refers to a concept – we have derived $\Psi(\alpha x \Psi(x))$. Only the last step, an application of a weak form of classical *reductio ad absurdum*, is essentially classical.

This proof is a natural deduction free-logic variation on the proof Frege himself gives in the Appendix to Volume II of *Grundgesetze*. He then goes on to parallel step I_B above. He summarizes the result:

[46] Stewart Shapiro and Alan Weir, '"Neo-Logicist" logic is not innocent', *Philosophia Mathematica*, 8 (2000), §§IV and V, pp. 160–89, use free logic at first order but not second.

[47] Neil Tennant, *Natural Logic* (Edinburgh: Edinburgh University Press, 1978), pp. 167–8. Tennant takes (the first-order) $\exists! a$ as an abbreviation for $\exists x(x = a)$. One need not, for the rules for the quantifiers and identity allow one to prove their logical equivalence. In the present setting, in which we have taken over the rules at second order too, and in which the significance of '∃!φ' has yet to be fully worked out, it is best that we take ∃! as primitive.

In other words: for every second-level function of one argument of type 2 there are two concepts such that, taken as arguments of this function, they determine the same value, but also such that this value does fall under the first concept and does not fall under the second. (*Gg* II, Appendix)

As Frege realizes only too well, the result is quite general: it holds for any vbto α in the extensional/Fregean framework.[48]

We can give a formal version of the proof II_A. It's more involved – see next page – but, as we shall see shortly, there is a point to considering the $I_A + II_A$ proof of Cantor's Theorem.

Making explicit the assumption that the predicate $\exists X(x = \alpha y Xy \ \& \ \neg Xx)$, which we have abbreviated $\Psi(x)$, refers to a concept, what the formal proofs give us is:

(i) α-Existence + $\exists!\Psi \vdash \Psi(\alpha x \Psi(x))$;

(ii) α-Existence + Converse of α-Extensionality + $\exists!\Psi \vdash$ $\neg\Psi(\alpha x \Psi(x))$.

Hence the combination of α-Existence, Converse of α-Extensionality, and the assumption that the predicate $\exists X(x = \alpha y Xy \ \& \ \neg Xx)$ refers to a concept is inconsistent.[49] Moreover, in view of the publication of Cantor's Theorem in 1890, this inconsistency was foreseeable.[50]

[48] Famously, in 'On concept and object', Frege tells us that 'the concept *horse*', being complete or saturated, names an object, not a concept. The expression 'The concept' is therefore a vbto. It requires rather strong Fregean nerves then to countenance that there must be two concepts, $\varphi(\xi)$ and $\chi(\xi)$, such that

the concept φ = the concept χ

but the object that we may variously refer to as 'the concept φ' or 'the concept χ' falls under one of these concepts but not the other. If, as seems plausible, we take it that 'the concept φ' names the extension of the concept φ – see Tyler Burge, 'Frege on extensions of concepts, from 1884 to 1903', in Burge, *Truth, Thought, Reason: Essays on Frege* (Oxford: Clarendon Press, 2005), pp. 273–98, at pp. 283–4 – this is, of course, merely a restatement of the application that led Frege to rethink extensions in the Appendix to volume II of *Grundgesetze*.

[49] The use of second-order quantification in obtaining the contradiction is essential. It is known that the (standard, hence also the free) first-order fragment of the system of *Grundgesetze* is consistent. See Terence Parsons, 'On the consistency of the first-order portion of Frege's logical system', *Notre Dame Journal of Formal Logic*, 28 (1987), pp. 161–8.

[50] As J. N. Crossley points out ('A note on Cantor's Theorem and Russell's paradox', *Australasian Journal of Philosophy*, 51 (1973), pp. 70–1, at p. 71), a derivation

$$\frac{\overline{\exists!\phi}^{\,2} \qquad \dfrac{\forall X\,\forall Y(\alpha yXy = \alpha yYy \rightarrow \forall x(Xx \leftrightarrow Yx)) \quad \exists!\psi}{\forall Y(\alpha y\psi(y) = \alpha yYy \rightarrow \forall x(\psi(x) \leftrightarrow Yx))}}{\alpha y\psi(y) = \alpha y\phi(y) \rightarrow \forall x(\psi(x) \leftrightarrow \phi(x))}$$

$$\frac{\overline{\alpha y\psi(y) = \alpha y\phi(y)\ \&\ \neg\phi(\alpha y\psi(y))}^{\,2}}{\alpha y\psi(y) = \alpha y\phi(y)}$$

$$\frac{\alpha y\psi(y) = \alpha y\phi(y) \rightarrow \forall x(\psi(x) \leftrightarrow \phi(x)) \qquad \alpha y\psi(y) = \alpha y\phi(y)}{\forall x(\psi(x) \leftrightarrow \phi(x))}$$

$$\frac{\forall X\exists z(z = \alpha yXy) \quad \exists!\psi}{\exists z(z = \alpha y\psi(y))} \qquad \frac{\overline{t = \alpha y\psi(y)}^{\,1} \quad \overline{\exists!t}^{\,1}}{\exists!\alpha y\psi(y)}$$

$$\frac{\exists z(z = \alpha y\psi(y)) \qquad \exists!\alpha y\psi(y)}{\exists!\alpha y\psi(y)}\,1$$

$$\frac{\forall x(\psi(x) \leftrightarrow \phi(x)) \qquad \exists!\alpha y\psi(y)}{\psi(\alpha y\psi(y)) \leftrightarrow \phi(\alpha y\psi(y))}$$

$$\frac{\overline{\alpha y\psi(y) = \alpha y\phi(y)\ \&\ \neg\phi(\alpha y\psi(y))}^{\,2}}{\neg\phi(\alpha y\psi(y))}$$

$$\frac{\psi(\alpha y\psi(y)) \leftrightarrow \phi(\alpha y\psi(y)) \qquad \neg\phi(\alpha y\psi(y))}{\neg\psi(\alpha y\psi(y))}$$

$$\overline{\psi(\alpha y\psi(y))}^{\,3} \quad =\!=\!=\!=\!=\!=\ \text{Def'n} \quad \exists X(\alpha x\psi(x) = \alpha xXx\ \&\ \neg X(\alpha x\psi(x)))$$

$$\frac{\exists X(\alpha x\psi(x) = \alpha xXx\ \&\ \neg X(\alpha x\psi(x))) \qquad \neg\psi(\alpha y\psi(y))}{\neg\psi(\alpha y\psi(y))}\,2$$

$$\frac{\neg\psi(\alpha y\psi(y))}{\neg\psi(\alpha y\psi(y))}\,3$$

Proof that $\forall X\exists z(z = \alpha xXx)$, $\forall X\,\forall Y(\alpha yXy = \alpha yYy \rightarrow \forall x(Xx \leftrightarrow Yx))$, $\exists!\psi \ \vdash \neg\psi(\alpha y\psi(y))$,
where $\psi(x)$ abbreviates $\exists X(x = \alpha yXy\ \&\ \neg Xx)$.

Foreseeable, not foreseen. John Burgess offers this account:

> The explanation is *not* that Frege rejected Cantor's results. A sufficient explanation is that Frege (like so many others) was largely *unaware of the bearing* of Cantor's cardinality theorems on the issues that concerned him. If he had pondered that bearing, he would surely have begun by translating Cantorian jargon into Fregean jargon. He would then immediately have seen that the Cantorian greater cardinality theorem says that there are more Fregean 'concepts' than Fregean 'objects'. He would then immediately have seen that this contradicts an axiom of the Fregean system, according to which there is a distinct 'object' associated with each 'concept,' namely, the 'class' that is its 'extension.' He would then surely have gone on to ponder whether or not the Cantorian proof can be reproduced within the Fregean system. He would then surely have seen that it can, and would thus have seen that his system is inconsistent.[51]

What adds pathos is that in his review of Frege's *Grundlagen* in 1885, Cantor had warned against taking extensions of concepts as the building blocks. Cantor already held then that there could be no set of all sets. Opinions divide on whether Cantor's warning was obscurely put or Frege simply negligent in, apparently, failing to understand it.[52]

of Russell's paradox is easily obtained from Cantor's proof of his theorem: if we take the domain of all sets to be, itself, a set, then, for that set V, we must have that $\wp(V)$ is V itself, for, on the one hand, all sets are contained in V, and, on the other, every set is a subset of V (as all its members are sets), and hence that the identity function is, *per impossibile*, a function from V onto the set of all its subsets. The set one then constructs in the course of Cantor's proof that there is no such function is the set $\{x \in V: x \notin x\}$, the Russell set. But by 1890 Cantor knew that the collection of all sets was, in the terminology he would later use, an absolutely infinite and inconsistent multiplicity, so he would not have carried out this application of his proof. (See Michael Hallett, *Cantorian Set Theory and Limitation of Size* (Oxford: Oxford University Press, 1984), chs. 1 and 4.

[51] John P. Burgess, 'Frege and arbitrary functions', in W. Demopoulos (ed.), *Frege's Philosophy of Mathematics* (Cambridge, Mass.: Harvard University Press, 1995), pp. 89–107, at pp. 101–2.

[52] Contrast Hallett, *Cantorian Set Theory*, pp. 126–7, and W. W. Tait, 'Frege versus Cantor and Dedekind: On the concept of number', in Tait (ed.), *Early Analytic Philosophy: Frege, Russell, Wittgenstein* (Chicago: Open Court, 1997), pp. 213–48, esp. §12.

9. WAYS OUT, EXPLORED AND UNEXPLORED

In recent years much highly productive effort has been spent in exploring weakened versions of Frege's theory.[53] What concerns me here is what responses were open to Frege in the light of his philosophy in the years from 1890 to Russell's delivery of his bombshell. It seems to me that there are two responses open to Frege, neither of which is there any evidence he considered.

The first is prompted by our having made the concept-existence assumption explicit in using a free-logic framework in deriving inconsistency. So one can read the proof as showing that α-Existence and Converse of α-Extensionality jointly entail that the predicate $\exists X(x = \alpha y Xy \; \& \; \neg Xx)$ does not refer to a concept. Formally, it is quite consistent to take that line, for in free logic, where the existence presuppositions of classical logic figure as *refutable* assumptions, α-Existence, α-Extensionality, and Converse of α-Extensionality are consistent. The little set-theoretic exercise with finite and co-finite sets shows this: take the domain to be the set of natural numbers, take concepts to be finite and co-finite subsets of that domain and, under some enumeration of the finite and co-finite subsets, take $\alpha x\varphi(x)$ to be the index of the set to which φ is mapped (and similarly for assignments of concepts to the second-order variables).

Under what conditions does a predicate not refer to a concept? In *Grundgesetze* Frege is quite explicit on this: φ(ξ) denotes a function (concept) if, whenever 'ξ' is replaced by a name that denotes, the resulting sentence denotes (§§29 and 31). In volume II of *Grundgesetze* we get something perhaps a little different:

> Any object Δ that you choose to take either falls under the concept Ψ or does not fall under it; *tertium non datur.* (*Gg* II, §56)[54]

[53] See the papers collected in §§II and III of Demopoulos, *Frege's Philosophy of Mathematics*, and John P. Burgess, *Fixing Frege* (Princeton: Princeton University Press, 2005.)

[54] Frege seems to be skittering between objectual and substitutional readings of his quantifiers. On Frege's reading of quantifiers see Leslie Stevenson, 'Frege's two definitions of quantification', *Philosophical Quarterly*, 23 (1973), pp. 207–23, which goes some way to explaining why objectual and substitutional readings may not be so far apart for Frege. At this point it is first-order quantifiers that concern us. Dummett remarks that at second order Frege's 'formulations make it likely that he thought of his function-variables as ranging only over those

Immediately before this Frege says,

The law of excluded middle is really just another form of the requirement that the concept have a sharp boundary.

A concept without sharp boundary is, he says, 'wrongly termed a concept'. There are other places in his writings where Frege makes similar stipulations.[55] We now see how to read our marker for the second-order existence assumption in our proof of contradiction:

'$\exists!\Psi$' means $\forall x(\Psi(x) \vee \neg\Psi(x))$.

Of course, if the Law of Excluded Middle is part of our free logic, this does no good. More precisely, we would have that α-Existence and the Converse of α-Extensionality are inconsistent in a second-order classical free logic. But to have the Law of Excluded Middle as part of our logic wouldn't be to play the game, if we wish to turn existence presuppositions into explicitly formulated existence assumptions.[56]

Our proof of $\neg\Psi(\alpha y\Psi(y))$ from α-Existence, Converse of α-Extensionality and $\exists!\Psi$ uses no essentially classical rule. Not so our proof of $\Psi(\alpha x\Psi(x))$ from α-Existence and $\exists!\Psi$; it uses classical *reductio ad absurdum*. But it uses it only once, as the last step in the proof. Instead we could use *reductio ad absurdum* to obtain a proof of $\neg\neg\Psi(\alpha x\Psi(x))$, which suffices for obtaining a contradiction from α-Existence, Converse of α-Extensionality and $\exists!\Psi$, the latter now construed as $\forall x(\Psi(x) \vee \neg\Psi(x))$.

In this setting we read our proof as a proof of

$$\neg\forall x(\exists X(x = \alpha yXy \mathbin{\&} \neg Xx) \vee \neg\exists X(x = \alpha yXy \mathbin{\&} \neg Xx)).$$

The predicate $\exists X(x = \alpha yXy \mathbin{\&} \neg Xx)$ cannot, on pain of contradiction, denote a concept. As our α-Existence claim is confined to

functions that could be referred to by functional expressions of his symbolism' (Dummett, *Frege: Philosophy of Mathematics*, p. 220).

55 'Function and concept', p. 148; 'The argument for my stricter canons of definition', *PW*, p. 152; 'Logic in mathematics', pp. 229, 241, 243.

56 Frege considers, but rejects, failure of the Law of Excluded Middle. He does so because he sees its failure as indicating that extensions of concepts would not be proper objects. He does not consider that the fault could lie with the predicate used. (See the Appendix to *Gg*, vol. II.)

concepts, it does not apply to this predicate – this predicate is not an allowed substituend – and the known route to paradox is blocked.

Can anything more general be said? In first-order intuitionist logic, $\neg\forall x(\varphi x \vee \neg\varphi x)$ is formally consistent (although $\exists x\, \neg(\varphi x \vee \neg\varphi x)$ is not). Now, our current reading of ∃! limits the range of the second-order quantifiers to what, in intuitionist terms, are decidable properties. *If* the predicate $\exists X(x = \alpha yXy\ \&\ \neg Xx)$ denotes a decidable property, paradox ensues. But even though the variable X ranges over decidable properties, it is not immediately evident that $\exists X(x = \alpha yXy\ \&\ \neg Xx)$ is itself decidable. Providing a semantic model in which the denotation of $\exists X(x = \alpha yXy\ \&\ \neg Xx)$ falls outside the range of second-order quantifiers appropriately limited in range so as to secure truth of $\forall X\forall x(Xx \vee \neg Xx)$ turns upon fine points in the interpretation of second-order quantifiers in the model-theory of second-order intuitionist logic. One thought against decidability of $\exists X(x = \alpha yXy\ \&\ \neg Xx)$ would be that, just as the domain of individuals can increase from lesser to greater states of information (earlier to later nodes) in Kripke models for first-order intuitionist logic, so too can the range of the second-order quantifiers when we make the move to second order: as information increases one learns of new decidable properties (or learns of old ones that they are decidable).

All of this may seem, even if feasible, desperately ad hoc. All I wish to claim for it is that it has its roots in Frege's pre-paradox writings.

A second route is also licensed by those writings, and perhaps more so than Frege realized. Consider, for a moment, Frege's stipulations regarding his surrogate for definite descriptions. The surrogate behaves as it should when exactly one object falls under the concept used in constructing the description: it denotes that object. When less than or more than one object falls under the concept, it denotes the extension of the concept. *This* isn't a matter of getting anything right: it's just a stipulation that ensures description-terms always have a reference. The same attitude is to the fore when, in 'Function and concept', Frege says that we must 'lay down rules from which it follows, e.g., what "⊙ + 1" is to mean, if "⊙" refers to the Sun' ('Function and concept', p. 148). He follows this injunction with the comment, 'What rules we lay down is a matter of comparative indifference.' In principle it is open to Frege to behave just as cavalierly in the case when it is determined that a predicate does not pick out an extension, cases of the kind Cantor was well aware of, cases of the kind that emerged from Cantor's, Russell's and Burali-Forti's paradoxes. In a logically perfect language, set-abstract terms that *seem* to pick out those 'impossible sets' must be assigned some reference, but need not denote extensions of the predicates occurring in them. Basic Law Vb then needs to be qualified. It prescribes 'normal behaviour', when set-abstracts do refer to, so to say, the right extensions. 'Abnormal' predicates may be assigned the same extension even though not co-extensive – we know there must be some predicates for which this happens. The problem with such an approach is in determining the range of the 'abnormal'. (This is not to say that the 'abnormal' cases must be explicitly taken care of in the revised basic law: compare Basic Law VI, which describes only the well-behaved cases for definite descriptions.)

This too was not the way Frege chose to go. Because of the constructive role played by extensions of concepts in both *Grundlagen* and *Grundgesetze*, Frege took the route of rethinking the very notion of *extension of a concept* in the light of the very general result he obtained in the Appendix to the second volume of *Grundgesetze*: some concepts must have the same extension even though not being co-extensive. It may be true that 'the function $\Phi(\xi)$ has the same course-of-values as the function $\Psi(\xi)$' even though it is not the case that 'the functions $\Phi(\xi)$ and $\Psi(\xi)$ have always the same value for the same argument'. Frege made the minimal change possible in the light of how he came to that discovery. The only problematic examples he knew of being obtained from $\exists X(x = \alpha y Xy \mathbin{\&} \neg Xx)$ and $\forall X(x = \alpha y Xy \rightarrow \neg Xx)$, he proposed that two concepts $\Phi(\xi)$ and $\Psi(\xi)$ have the same extension if, and only if, the functions $\Phi(\xi)$ and $\Psi(\xi)$ have always the same value for the same argument save with the possible exceptions of the object that is their common extension.

It is known that Frege's specific proposal fails to avoid paradox (as recognized by Lesniewski, Geach and Quine).[57] Dummett says,

[57] See Gregory Landini, 'The ins and outs of Frege's way out', *Philosophia Mathematica*, 14 (2006), pp. 1–25 for a recent, and somewhat wayward, discussion. One enterprise that has attracted a small following away from the mainstream of Frege scholarship is the investigation of something akin to Frege's system in weak logics. The naive comprehension principle (roughly, α-Existence for set abstracts) is known to be consistent in certain weak logics. It is also known

The inconsistency of Frege's *Grundgesetze* system was not a mere accident (though a disastrous one) due to carelessness of formulation. He discovered, by August 1906, that it could not be put right within the framework of the theory, that is, with the abstraction operator as primitive and an axiom governing the condition for the identity of value-ranges: but the underlying error lay much deeper than a misconception concerning the foundations of set theory. It was an error affecting his entire philosophy.[58]

Exactly what Frege realized in late spring or the summer of 1906 is not quite clear. Surmise is aided by the unfinished manuscript of a response to an article of Arthur Schönflies's. A list of headers includes, for parts unwritten,

Concepts which coincide in extension, although this extension falls under the one but not the other.

Remedy from extensions of second level concepts impossible.

Set theory in ruins.[59]

Clearly the hopes of the Appendix to volume II of *Grundgesetze* had been dashed.[60]

One *methodological* error is Frege's belief that in a logically perfect language all properly formed singular terms must refer. As indicated above, there are ways to dilute the consequences of that principle, but it is, nevertheless, ill founded.

It is true that we usually do not use names that we know do not refer (save perhaps ones like 'Santa Claus' that have a recognized social context for their use). Standard logic codifies usage with referential assumptions built in. Frege himself says that in a logically perfect language 'no new sign shall be introduced as a proper name without being secured a reference'. He attempts to secure reference by stipulation. This way of proceeding is very much at odds with not just ordinary but also mathematical practice.

that this need not be the comfort it may at first seem: the naive comprehension principle is consistent in what Petr Hájek calls Basic Fuzzy Logic but the theory is not consistent with the existence of a set of natural numbers obeying a certain, moderately strong schema of mathematical induction (Petr Hájek, 'On arithmetic in the Cantor–Lukasiewicz fuzzy set theory', *Archive for Mathematical Logic*, 44 (2005), pp. 763–82).

[58] Dummett, *Frege: Philosophy of Mathematics*, p. 223.

[59] 'On Schoenflies: *Die logischen Paradoxen der Mengenlehre*', *PW*, p. 176.

[60] See further Dummett, *Frege: Philosophy of Mathematics*, pp. 4–6.

Our language contains general means for producing singular terms, e.g., definite descriptions. In *some* sense they are part of our language: they are products of its generative capacity. But by and large they are not part of our *language-in-use*. To take one of Frege's own examples, exactly because there is no least rapidly convergent series, the mathematician has no use for the expression 'the least rapidly convergent series'. True, false beliefs can lead one to use non-referring singular terms, but in use, in conversation say, matters will not run along normal lines if some parties are appraised of the facts that deny the term a reference.

Supported by theoretical claims, notably the aboutness thesis and the functional conception of concepts, connectives and quantifiers, which lead to the unassertibility of singular negative existential claims, Frege mistakes a *defeasible presumption* of reference in ordinary usage for a *presupposition*. He then takes it as given that a properly systematic, logically perfected language must respect that supposed presupposition for all singular terms generable in the language, not just those that have found a use to date. This leads to oddity but is not itself responsible for error. Error comes in the contrast between the treatment of definite descriptions and of terms for extensions of concepts: the wholesale attribution of references with a particular characteristic – satisfaction of Basic Law V – to the latter, the more relaxed who-cares-as-long-as-there-is-a-reference? attitude to description-terms when not exactly one item falls under the concept involved.

10 FREGEAN SET-THEORY: RETAINING A SEMBLANCE OF FREGEAN PREOCCUPATIONS

What were called above Frege's and Dummett's trilemmata show that we cannot maintain all that Frege says about truth. But we can keep a fair amount and a surprisingly large simulacrum of the whole Fregean project, once we reject the functional account of the connectives. To be more exact, we may maintain, with Frege:

I) Truth is unanalysable and *sui generis*.
II) For any assertoric sentence *P*, *P* and 'It is true that *P*' express the same thought.

III) If singular terms are used in the ordinary way in sentences involving simple predications, what one intends to speak of is their reference.
IV) Concepts map objects to truth-values.
V) Simple, i.e. logically unstructured, predicates refer to concepts (if they refer at all).
VI) Our propositional logic is classical (at least for negation, conjunction and disjunction).
VII) Sentences comprising a simple predication and one or more non-referring singular terms are neither true nor false.
VIII) For reasoning within the scope of presumptions of reference, our first-order logic is standard first-order logic.
IX) Basic Law V applies to extensions of concepts (or set abstracts) with second-order quantifiers ranging over sharply defined concepts. I.e., we have α-Existence, α-Extensionality, and the Converse of α-Extensionality for the extension-of-a-concept vbto.

What is unFregean is that

X) We adopt the semantic conception of falsity.
XI) We accept various truisms incompatible with the functional understanding of the logical connectives.
XII) Our general, first- and second-order logic is classical but free.
XIII) For some sentences *P*, it is true that not-*P* even though it is not false that *P*. (This is how Dummett's trilemma is evaded.)

In this setting, '$\neg\exists x(x = a)$' is true when *a* does not refer, which is to the good. Furthermore, the logic being free, there is room to give an inferentialist account of vbtos, as does Tennant.[61]

We have Basic Law V in form. It is consistent provided appropriate constraints are placed on the range of the second-order variables, i.e. on what count as concepts. (The finite–co-finite subsets-of-$\mathbb{N}$ interpretation shows that consistency is attainable.) The hard work

[61] Neil Tennant, 'A general theory of abstraction operators', *Philosophical Quarterly*, 54 (2004), pp. 105–33 (on which see further Peter Milne, 'Existence, freedom, identity, and the logic of abstractionist realism', *Mind*, 116 (2007), pp. 23–53).

goes into what we might call 'the theory of second-order ∃!', which remains to be elaborated, i.e. in specifying concepts. What are the closure conditions of the domain of (extensions of) concepts? Which predicates with one free first-order variable (and no free second-order variable) refer to concepts? Investigation of these topics puts a new spin on the old Quinean saw that second-order logic is set theory.

APPENDIX

We need several clauses to take care of falsity, clauses providing a recursive account:

(i) For any atomic sentence $Rt_1t_2\ldots t_n$
It's false that $Rt_1t_2\ldots t_n$ if, and only if,
$\neg Rt_1t_2\ldots t_n$ and $\exists! t_1$ and $\exists! t_2$ … and $\exists! t_n$.

(This applies as much to identity statements as any other atomic formulas.)

(ii) For any assertoric sentence *P*,
it's false that $\neg P$ if, and only if, *P*.

(iii) For any assertoric sentences *P* and *Q*,
it's false that $P \vee Q$ if, and only if, it's false that *P* or it's false that *Q*.

(iv) For any assertoric sentences *P* and *Q*,
it's false that $P \vee$ Q if, and only if, it's false that P and it's false that Q.

[(v) For any assertoric sentences P and Q,
it's false that P → Q if it's true that P and it's false that Q.]

One feature of this account is to be noted. We have included a minimalist account of reference:

't' refers if, and only if, $\exists! t$

or, equivalently,

'*t*' refers if, and only if, $t = t$.

This, on the face of it, is a rather unFregean thing to do. On the face of it, Frege would want to *explain* the failure of '$t = t$' to be true by saying that '*t*' fails to refer. But, on the other hand, he might be thought to come close to equating the two when he says, People certainly say that Odysseus is not an historical person, and mean by this contradictory expression that the name 'Odysseus' designates nothing, has no reference.[62]

It all rather depends on what he means by 'mean'. He might, after all, just be saying that what people who say that Odysseus is not an historical person really mean to say, what they are trying to express by that – as he calls it, but he's on dodgy ground in his own terms doing so – contradictory formulation, is that the name 'Odysseus' fails to refer.

It remains to extend the definition of falsity to the first-order quantifiers. Here we make matters easy for ourselves by assuming that every object has a name. There may, of course, be singular terms that do not refer.

(vi) For any sentence $\forall x\varphi$

it is false that $\forall x\varphi$ just in case, for some singular term t, $\exists!t$ and it is false that $\varphi[t/x]$.

(vii) For any sentence $\exists x\varphi$

it is false that $\exists x\varphi$ just in case, for every singular term t, if $\exists!t$ then it is false that $\varphi[t/x]$.

Our definition of truth being given by the equivalence scheme, we do not have, as yet, *any* constraints on how either connectives or quantifiers behave with respect to truth. But that is as it should be. If one holds that the equivalence thesis says all there is to say, fundamentally, about truth, one does not look to it to justify one's logic. Rather, one looks to the logic to draw out consequences of the equivalence thesis. If the propositional logic is classical we find that $\neg p$ is true if, and only if, p is not true.

What we do find is that the negation–conjunction–disjunction fragment of classical propositional logic is sound and complete should we aim to obtain Tappolet's truisms. In similar truistic spirit, bearing in mind that some names may not refer, what we'd expect to hold for truth is this:

(vi°) For any sentence $\forall x\varphi$

[62] 'Introduction to logic', p. 191.

it is true that $\forall x\varphi$ just in case, for all singular terms 't', if $\exists!t$ then it is true that $\varphi[t/x]$.

(vii°) For any sentence $\exists x\varphi$

it is true that $\exists x\varphi$ just in case, for some singular term 't', $\exists!t$, and it is true that $\varphi[t/x]$.

The logic we want will fail to be classical precisely because we are not granting that all singular terms refer, equivalently, we are not granting $\exists!t$, equivalently, $\exists x(x = t)$, for all singular terms t. So formulas of that form play a special role. We may adopt, for example, the natural deduction rules as laid out in Tennant's *Natural Logic*, including now his 'denotation rule', or we might take the axiomatic system of Tyler Burge's 'Truth and singular terms'.[63] The denotation rule (or Burge's axiom (A9)) has it that an atomic formula entails $\exists x(x = t)$ for any name occurring in it, and that's what we want: an atomic sentence is true only if all the terms it contains refer. Ignoring the conditional and biconditional, this logic, however formulated, is sound and complete with respect to truth-preservation as determined by our truisms (suppressing any worries, which are certainly not special to this context, issued by the appeal to a substitutional reading of the quantifiers).

(Is it a truism, once we allow truth-value gaps, that a false conclusion may only follow from premises at least one of which is false? Or that a false conclusion cannot follow from true premises? *If* the former, our propositional logic is weakened, for the rules of disjunctive syllogism, double negation introduction, and *ex falso quodlibet* cease to hold.)

[63] Tyler Burge, 'Truth and singular terms', *Noûs*, 8 (1974), pp. 167–81; reprinted in M. Platts (ed.), *Reference, Truth and Reality* (London: Routledge and Kegan Paul, 1980), pp. 309–25. Both treat definite descriptions, which I have been studiously ignoring. Burge treats function symbols.

PETER HYLTON

13 Frege and Russell

I

Frege and Russell are often linked, as the founders of twentieth-century analytic philosophy. Besides this historical, retrospective, connection, there are also important similarities in doctrine between them.[1] Each was a logician, whose work in logic was closely integrated with his work in philosophy; each held that philosophical problems can be clarified and, in some cases, solved, by means of logic. (This view that the technical and the philosophical are not distinct is characteristic of one clear line of thought in twentieth-century analytic philosophy.) Each argued for, and tried to prove, logicism, the thesis that arithmetic can be reduced to logic, and is thus no more than logic in disguise.[2] Each was strongly opposed to psychologism; each believed in a 'third realm', neither physical nor mental, which provides the subject matter for objective judgements

[1] I speak, here and throughout this essay, of Russell's views after his break with Idealism, around 1900, and before his shift towards pragmatism and behaviourism, around 1920. All of his works which played a foundational role for twentieth-century analytic philosophy were written in these two decades. Frege's views change much less markedly. I do attribute logicism to Frege, although he abandoned that view towards the end of his life. I also attribute to him a view of functions as non-linguistic entities, in spite of some remarks to the contrary in the early sections of *Begriffsschrift*. Finally, I attribute to him some version of the distinction between sense and reference that the he puts forward in the 1892 essay 'On sense and reference'; although not articulated clearly until that essay, the distinction seems to me present, although in nascent form, as early as *Begriffsschrift*.

[2] Russell accepted, as Frege did not, that geometry can be reduced to arithmetic, and thus, via logicism, to logic. Frege's view here reveals something important about his inchoate epistemological views; I shall not go further into this matter here, however.

about abstract matters. (In Frege's case, however, it is perhaps unclear just what this belief comes to.) In particular, each believed that our declarative sentences have an objective content, independent of human action – that, as Frege puts it, there is not *my* Pythagorean theorem and *your* Pythagorean theorem but *the* Pythagorean theorem, independent of both of us, and timelessly true.[3] (Russell to some extent backs away from this view after 1906, as we shall see; the shift, however, has relatively little effect on the issues I shall be discussing in this essay. See p. 538, below.)

The primary focus of this essay, however, is not on the similarities between the views of Frege and of Russell but on their differences. It is no part of my concern to deny the similarities indicated above; they are real, and central to the thought of each of our philosophers. Nor do I mean to cast in doubt the natural pairing of Frege with Russell. On the contrary: it is because their views are in some ways so similar, and the pairing so natural, that differences between them are of great interest. Let me briefly outline my discussion of some of these differences.

I begin, in §2, with a rather well-known difference. Frege distinguishes the sense (*Sinn*) of an expression from its reference (*Bedeutung*),[4] whereas Russell denies that any such distinction is fundamental. I connect this difference with aspects of Russell's epistemology; in particular, with the fact that he takes acquaintance – a direct and unmediated relation between the mind and a known object – to be the foundation of all our knowledge. These views of Russell's pose significant difficulties. In the period before 'On denoting' he attempted one kind of resolution of these difficulties, putting forward what I shall call 'the theory of denoting concepts'. This theory accepts a distinction, for some expressions, which is in some ways akin to Frege's distinction between sense and reference; it is the subject of §3. The next section deals with the theory of descriptions, which Russell put forward in 'On denoting' and held thereafter. §5 elaborates on the way in which that theory

[3] See 'Thoughts', pp. 68–9 of the original printing; pp. 362–3 of *CP*. The expression 'third realm' is in this same passage.

[4] *Bedeutung* is sometimes translated as 'meaning', but Russell sometimes uses the word 'meaning' for something akin to Fregean *Sinn*. In this chapter, 'reference' is used throughout for *Bedeutung*.

enables Russell to avoid any analogue of the Fregean distinction. Central to Russell's answer is the idea that most apparent referring expressions are not genuine referring expressions; in particular, that there are no *complex* referring expressions. Functional expressions, such as '2 + 3' or 'the father of Alexander the Great', are, on the face of it, complex referring expressions. In accordance with what we have just said, Russell's new (post-1905) view cannot accept these expressions as primitive; they must, rather, be defined as needed. This point leads in turn to a further issue. For Frege, the function-argument method of analysis is fundamental. Since Russell does not take functions as primitive, he cannot agree with Frege on this central point. §6 concerns this difference, and the conception of the world that underlies Russell's idea of analysis. It also takes up the question of how, consistent with this conception, Russell can define functions. Finally, in §7, I discuss ways in which the metaphysical differences which have occupied us in earlier sections make a difference to the logics of Frege and of Russell. Throughout these discussions I devote more space to Russell than to Frege.

Before beginning the comparison and contrast outlined above, I shall very briefly discuss the question of the influence of Frege on Russell. Russell's work in the philosophy of mathematics does not begin until the mid-1890s; his anti-psychologism, his development of a system of logic and his logicism all post-date his rejection of Idealism in 1899. By this time most of Frege's works were already in print. (Volume II of *Grundgesetze* and the three late essays 'Thoughts', 'Negation' and 'Compound thoughts' form the main exceptions, together, of course, with those of his works which were not published at all in his lifetime.) In view of this chronology, and of the doctrinal overlap indicated in the first paragraph of this essay, one might be inclined to think that Russell learned a great deal from Frege. Further plausibility accrues to this idea from similarities in the techniques used at certain points in the attempt to reduce mathematics to logic, including the technique for the definition of number, the so-called Frege-Russell definition of number.

According to Russell, however, the main lines of his philosophical views, his logic, and his attempt to reduce mathematics to that logic, were all laid down before he studied Frege's work. He completed the main text of *The Principles of Mathematics* on the last day of December 1901. By his own account he had looked at some of Frege's work before that date, but had not studied it with the care needed to understand it. In June 1902 he wrote his famous letter to Frege, announcing the discovery of the contradiction in Frege's logic (i.e. of what is now known as 'Russell's paradox'). That letter makes it sound as if his close study of Frege's work is just beginning: 'I have known of your Basic Laws of Arithmetic for a year and a half, but only now have I been able to find the time for the thorough study I intend to devote to your writings'.[5] Similarly, in the Preface to *The Principles of Mathematics*, dated December 1902, he says: 'Professor Frege's work, which largely anticipates my own, was for the most part unknown to me when the printing of the present work began'.[6] For this reason, he says, he discusses Frege's work in detail in an appendix, written while the main body of the work was at press. Later in the Preface, he acknowledges the influence of Cantor and of Peano and says: 'If I had become acquainted sooner with the work of Professor Frege, I should have owed a great deal to him, but as it is I arrived independently at many results which he had already established' (p. xviii). In later works, looking back on this period, he tells the same story.[7]

It would be easy to be sceptical, even cynical, about Russell's account of what he learned from Frege. What evidence there is, however, seems to favour it. Without pretending to have a definitive view, I am inclined to take Russell's account at face value, and to think that the decisive influences on Russell, from his rejection of Idealism to the writing of *The Principles of Mathematics*, were G. E. Moore, in metaphysics, Peano, in logic, and Cantor and Weierstrass in mathematics. To begin with, Russell was always generous in his acknowledgements; there is no reason at all to think he would make

[5] Russell to Frege, 16 June 1902. The correspondence is published in *WB*. The passage quoted here is at p. 213. I largely follow the English translation by Hans Kaal in *PMC*: the passage quoted here is at p. 130 of that work.

[6] Bertrand Russell, *The Principles of Mathematics* (Cambridge: Cambridge University Press, 1903), p. xvi.

[7] See Bertrand Russell, 'My mental development', in Paul Arthur Schilpp (ed.), *The Philosophy of Bertrand Russell* (Evanston, Ill.: Northwestern University, 1944), especially p. 13; and *My Philosophical Development* (London: Allen and Unwin, 1959, p. 66).

an exception in this one case. More important, perhaps, the internal evidence strongly suggests that Russell first developed his logic by building on what he learned from Peano, rather than by following Frege. The logic of *The Principles of Mathematics* strikes anyone who had studied Frege with care as clumsy, or perhaps even confused. The idea that this logic was developed by beginning with Peano, by contrast, seems entirely plausible.

Taking Russell's account at face-value, however, does not mean that we should conclude that he owes nothing at all to Frege. Frege, Russell, and Peano did not live in separate intellectual worlds. There is some reason to believe that Russell may have first come across the idea for his definition of number (which is also Frege's) in a 1901 essay by Peano (who discusses the idea, but rejects it). And Peano, presumably, had read Frege's *Grundgesetze*, since he wrote a review of it in 1895.[8] Russell's logic, moreover, developed significantly after he wrote *The Principles of Mathematics*, and there is every reason to think that Frege's influence, along with the continuing influence of Peano, was important in this development. This influence is, indeed, explicitly acknowledged; in the Preface to *Principia Mathematica*, Whitehead and Russell say: 'In all questions of logical analysis, our chief debt is to Frege.'

2

Let us begin our main discussion with a disagreement between Frege and Russell that occurs in their correspondence. The issue arose from a discussion of truth. In a letter dated November 1904. Frege had said: 'Truth is not a component part of a thought, just as Mont Blanc with its snowfields is not itself a component part of the thought that Mont Blanc is more than 4,000 metres high.'[9] Russell's reply ignored the issue about truth, which was the point of Frege's remark (and with which he agreed), and seized on the incidental illustration to articulate his objections to Frege's distinction between sense and reference:

[8] See Gregory H. Moore, Editor's Introduction, *The Collected Papers of Bertrand Russell*, vol. III, especially p. xxvii.

[9] *WB*, p. 245, *PMC*, p. 163.

I believe that in spite of all its snowfields Mont Blanc itself is a component part of what is actually asserted in the proposition [*Satz*] 'Mont Blanc is more than 4,000 metres high'. We do not assert the thought, for this is a private psychological matter: we assert the object of the thought, and this is, to my mind, a certain complex (objective proposition [*objectiver Satz*], one might say) in which Mont Blanc is itself a component part. *If we do not admit this, then we get the conclusion that we know nothing at all about Mont Blanc* ... In the case of a simple proper name like 'Socrates', I cannot distinguish between sense and reference; I see only the idea, which is psychological, and the object. Or better: I do not admit the sense at all, but only the idea and the reference.[10]

This passage indicates very general differences in the underlying philosophical views of Frege and of Russell.

Consider the judgement expressed by the sentence 'Mont Blanc is over 4,000 metres high'. Each of Frege and Russell holds that in making this judgement we are somehow related to an objective non-linguistic entity – we 'grasp' it (*fassen* is Frege's word). Frege calls this entity a 'thought'. Russell speaks of such an entity as an 'objective proposition' (*objectiver Satz*, in the letter to Frege); for him a thought is 'a private psychological matter'. Thus far the differences are perhaps only terminological, but the next point is substantial. For Russell, a proposition, what we are most directly related to in making judgements, will in paradigmatic cases *contain* the entity we are talking about. It is explicit in the above passage that Mont Blanc is a constituent – a 'component part' – of the proposition expressed in the judgement. For Frege, by contrast, thoughts do not contain the entities themselves, the subjects of our judgement. The constituents of Fregean thought are the senses of expressions that refer to the entities we mean to be talking about – not those entities themselves.[11]

[10] *WB*, pp. 250–1, *PMC*, p. 169. The emphasis here is added. Russell makes a very similar point in the 1911 essay, 'Knowledge by acquaintance and knowledge by description'. He discusses 'the view that judgements are composed of something called "ideas"', and says: 'in this view ideas become a veil between us and outside things – we never really, in knowledge, attain to the things we are supposed to be knowing about, but only to the ideas of those things' (Russell, *Collected Papers*, vol. VI, p. 155).

[11] There are some reasons to be hesitant in attributing to Frege the idea that thoughts have constituents at all. The attribution is supported by some of Frege's texts, however, and certainly facilitates the comparison between Russell and Frege that is my concern here.

Russell's view can be elaborated and illustrated by briefly considering his attitude towards truth and facts. Truth, for him, is an indefinable property of propositions (as, of course, is falsehood); a fact is simply a proposition which is true. In this view, he retains something like the ordinary notion of a fact, as consisting perhaps in an object's having a certain property, or standing in certain relations to one or more other objects. These 'objective complexes', as Russell calls them, are made up of one or more objects, together with some of their properties or relations.[12] And true propositions are identified with such entities. Thus Russell says:

People imagine that if *A* exists, *A* is a fact; but really the fact is '*A*'s existence' or 'that *A* exists'. Things of this sort, *i.e.* 'that *A* exists' ... I call *propositions*, and it is things of this sort that are called *facts* when they happen to be true.[13]

Here again we see, in a slightly different context, the view that a proposition about a particular object will, paradigmatically at least, contain that object, just as one might naturally think of a fact as containing, or made up of, an object (together perhaps with a property of the object). If the proposition is true, then it simply *is* the fact; if the proposition is false, then it is, so to speak, just like a fact except that it happens not to be true. The proposition is equally real in either case.

So far we have elaborated a little on Russell's opposition to Frege about the way that names function: for Russell, the presence of a name in a sentence implies, at least in paradigmatic cases, that the sentence expresses a proposition which contains the named object. We have as yet, however, seen no reasons for this opposition. The vital clue here, I think, is given by the sentence emphasized in the passage quoted above: 'If we do not admit this, then we get the conclusion that we know nothing at all about Mont Blanc ...' The emphasis here should be on the 'about' rather than on the 'know'. The issue is not one of our having *correct* beliefs about Mont Blanc, but rather one of our having beliefs which are genuinely *about* that

[12] The phrase 'objective complex' occurs, for example, in an essay dated June 1905 called 'The nature of truth', first published in Russell's *Collected Papers*, vol. IV, pp. 492–506; see p. 495.

[13] *Ibid.*, p. 492.

mountain at all. (I shall speak of this sort of issue as epistemological, since it is not merely about how things are but also about our relation to them. This is perhaps an extension of the usual sense of the word.) Let us suppose, with Frege and Russell, that the sentence 'Mont Blanc is over 4,000 metres high' expresses an objective entity, and that we do indeed 'grasp' that entity. How does that grasping enable us to believe something about the actual snowy mountain itself? For Russell, it does so because the entity that we grasp *contains* that mountain as a constituent. Frege's view, if we express it in these alien terms, must be quite different: that what we most directly or immediately know or grasp has as a constituent (perhaps) the sense of the expression 'Mont Blanc'. But how, in virtue of grasping that entity, do we know something about the mountain, which is altogether distinct from it? From Russell's point of view this question – 'the *in-virtue-of* problem', we might call it – presents a severe difficulty; his view attempts to avoid that difficulty by insisting that, at least in paradigmatic cases, we grasp propositions which contain the very entities which they are about.

These issues must be seen in the context of epistemology. Throughout the period which is our concern, Russell takes it that knowledge is at bottom a matter of a direct and unmediated relation between the mind and the known object. (Clearly nothing of the sort holds for Frege.) Russell insists that there is such a relation, and that it plays the fundamental role in knowledge. It is only by being in direct contact with some external object that the mind able to know anything at all outside itself. 'External' here does not carry its usual spatio-temporal implications: it means only non-mental, or outside the mind. Russell has no qualms at all about assuming that we also have this kind of knowledge of purely abstract entities. On the contrary: he applies his basic picture of knowledge both to abstract objects and to concrete. That distinction, indeed, is relatively unimportant to his thought during the time with which we are concerned. For the first few years of that period he holds that all entities *subsist* or have being; some have the additional property of *existing*, i.e. (roughly, being in space and time). Our being in a direct epistemic relation to an entity does not, in this view, require that it should exist, in this sense.

Russell thus postulates a fundamental epistemic relation holding between a mind, on the one hand, and an object – existing or merely subsisting – on the other hand. After 1905 Russell calls this relation *acquaintance*, and it comes to play an increasingly explicit role in his thought. But even before 1905, from his rejection of Idealism onwards, it is an essential element in his philosophy. In the Preface to the *Principles of Mathematics*, for example, he says:

> The discussion of indefinables – which forms the chief part of philosophical logic – is the endeavour to see clearly, and to make others see clearly, the entities concerned, in order that the mind may have that kind of acquaintance with them which it has with redness or with the taste of a pineapple. (p. xv)

Russell speaks here of our knowledge of simple sensory qualities to suggest the directness and immediacy which are characteristic of his notion of acquaintance. There have, of course, been philosophers – including his Idealist opponents – who thought that not even simple sensory qualities are in fact known in the direct and immediate way that Russell wants to convey. Such qualities, however, may at least *seem* to be known in that sort of way, and this may be enough to achieve his rhetorical purposes here.

I shall speak of Russell's insistence on a direct and unmediated relation between the mind and the known object as his *direct realism*; I shall include under this head the idea that propositions paradigmatically contain the entities they are about.[14] This view, or nexus of views, must, I think be traced to Russell's rejection of Idealism. The Idealists had insisted that knowledge is mediated by a complex structure, which is also (or therefore) the structure of the world; our knowledge of this structure thus gives us knowledge of the world which is purely rational in its basis. Russell, following G. E. Moore, had cut through all such considerations by insisting, to the contrary, that the most basic sort of knowledge is direct and

[14] It might be said that the term 'direct realism' is inappropriate, because Russell comes to believe that we do *not* have direct knowledge of ordinary objects – tables and chairs and other people, and the like. By 1912, his view is that our knowledge of these things is indirect, mediated by our knowledge of sense-data and universals (which are known directly). I use the term 'direct realism' because it emphasizes the fact that his view is always that *some* entities must be known directly and immediately, even though his views about *which* entities are known directly changes over time. Still, there are no doubt uses of the term according to which Russell's view, at least in the second half of the period we are concerned with, would not count as direct realism.

unmediated. The presence of an intervening structure would, from that point of view, simply mean that our knowledge failed to attain its desired object. We would end up knowing not the object itself but rather only the intervening structure. There is, of course, much more to be said about the origin of this view of Russell's, but that would take us aside from the comparison of Frege with Russell. We shall therefore treat Russell's direct realism, in the sense indicated, as more or less an axiom of his thought.[15]

3

Russell's direct realism seems to give a clear and straightforward answer to the question how the propositions we express manage to be about the entities they are about: they are about them in virtue of containing them. Presumably our 'grasping' a proposition implies our 'grasping' its constituents; presumably it is this that allows our thought to get right through to those objects, which are the things that we mean to be talking about. This picture was, I think, his underlying instinctive view throughout the period which is our concern – the view towards which he was always attracted, and which he tended to assume. It faces, however, great difficulties. Russell attempted to resolve those difficulties in one way in the period from 1901 until June 1905, when he came across the fundamental idea of 'On denoting';[16] thereafter he resolved them in a quite different way. These two different ways of responding to difficulties in the underlying picture go along with differences in the view that Russell takes of analysis, and related matters, and are therefore of quite general significance. In this section I shall briefly discuss the first method of resolution and its concomitants; in the next section I shall turn to the second.

[15] For a much more detailed discussion of this and of related issues, see Peter Hylton, *Russell, Idealism, and the Emergence of Analytic Philosophy* (Oxford: Oxford University Press, 1990), especially ch. 4. The fact that Russell is reacting against neo-Hegelian Idealism, whereas Frege is not, is itself an important point of contrast between the two, and connected with others. I shall not, however, go into this matter further in this essay.

[16] The first statement of the new view is in a manuscript entitled 'On fundamentals', published for the first time in *Collected Papers*, vol. IV, pp. 360–413; the manuscript is dated '1905', and the words 'begun June 7' are on the first folio.

Let us begin with the difficulties facing the underlying picture. It is undeniable, one might suppose, that I understand propositions about Socrates; but it may appear as quite implausible that I stand in some direct epistemological relation to him, for he no longer exists. The case of Pegasus or the present King of France, who have never existed, may seem to be worse. So Russell must accept that I can be in direct epistemological contact with what we might call non-existent *concreta* – entities which are of the right kind to exist, but happen not to. This consequence is something that Russell was for a time willing to accept, making heavy use of the distinction, to which we have already alluded, between existence and subsistence. Pegasus, though he does not exist (roughly, is not in space and time), does, Russell thinks *subsist* (is nonetheless real). And Russell was, as we have said, willing to accept that we can stand in direct epistemological relations to non-existent *concreta* (as well as to other non-existent objects, those that we would call abstract objects). So he was, for a time, willing to accept this sort of apparently implausible consequence of his direct realism. (As we shall see, however, this is a point on which he changed his mind, even before 'On denoting'.)

There is, however, another sort of difficulty, which he never accepted. Suppose I say, for example, 'Every natural number is either odd or even.' The underlying picture of direct realism might suggest that I am expressing (and grasping) a proposition which contains all of the infinitely many the natural numbers. Russell was willing to be agnostic about whether there in fact *are* any such infinitely complex concepts. But he denied that we can grasp propositions that have this sort of infinite complexity (see *Principles of Mathematics*, §72). That we grasp infinitely complex propositions was too implausible for Russell to accept, even in the most extreme and unrestrained phase of his realism. So the issue of *generality* – how we can, for example, grasp a proposition about all the natural numbers – is one which does not fit neatly into his direct realism. It is this issue which first forces upon Russell some modification of his direct realism.[17]

[17] In the Preface to *The Principles of Mathematics*, 2nd edn (London: Allen and Unwin, 1937), he speaks of his work on the philosophy of dynamics, and says: 'I was led to a re-examination of the principles of Geometry, thence to the philosophy

An unqualified version of direct realism thus serves as a paradigm for Russell. He relies on it and presupposes it at many points, and makes statements which seem to imply this unqualified view. The passage we saw in the letter to Frege is an example. But it is always a modified or qualified version which he explicitly advocates. He takes it that the most direct way in which a proposition can be about an object is simply by containing it; but he recognizes that we must have some way of making sense of cases in which a proposition is about an entity or entities which it does *not* contain; in such cases we might speak of the proposition's being *indirectly* about the entity. (In these terms we can say that Frege's view is one in which there is only indirect aboutness: a thought is about an object in virtue of containing the relevant sense. But of course these terms of description are Russell's, and quite foreign to Frege's thought.)

From 1900 or 1901 until June 1905 the modification to the underlying picture – Russell's way of accommodating indirect aboutness – is what I shall call the theory of denoting concepts. This doctrine simply accepts that direct realism does not hold in all cases; it allows a large class of exceptions to the general rule that the entity which a proposition is about is contained in the proposition; the general rule functions as a paradigm in Russell's thought, but certain cases are allowed to violate it. For certain kinds of phrases Russell accepts a distinction in some ways analogous to Frege's distinction between sense and reference. The analogue of the sense of an expression is what he calls the *denoting concept* which it expresses, or as he later comes to say, its meaning; the analogue of the reference is the denotation of the expression, or object it denotes – if it does in fact succeed in denoting something.[18] The phrases to which Russell

of continuity and infinity, and thence, with a view to discovering the meaning of the word *any*, to Symbolic Logic.' (p. xvii). The question of 'the meaning of the word *any*' is exactly what I am calling the issue of generality.

[18] Here there is a point which, though more or less incidental to our discussion, is in other contexts quite crucial. It is not implied by Russell's other views about denoting that a denoting concept must always succeed in denoting; it is entirely consistent with his view that such a concept should not in fact denote anything. At some moments he recognizes and accepts this point quite explicitly; see, for example, *The Principles of Mathematics*, §73. (At other moments, however, he seems to imply the opposite; see §427 of the same work.) For further discussion, see Hylton, *Russell, Idealism*, especially chs. 5 and 6. The point made in passing

initially applies this distinction are descriptions, both definite descriptions such as 'the President of the USA in 2000' and indefinite descriptions, such as 'any prime number'. Where such a phrase occurs in a sentence, that sentence is taken to express a proposition which contains not the corresponding object or objects but rather a concept which *denotes* that object or those objects; the proposition contains a denoting concept but is about – indirectly about – the denoted object or objects. Here there is an in-virtue-of problem. How, in virtue of containing a denoting concept, is the proposition *about* an entity wholly distinct from it, an entity which we do not in any sense 'grasp'? To this question Russell has no answer: the relation of denoting is simply asserted to have that effect.[19]

Using this theory, Russell hopes to account for generality by (roughly speaking) treating a phrase such as 'any natural number' – or 'any object' – as representing a denoting concept. In this attempt he is unsuccessful; the theory proves unable to give a coherent account of multiple generality.[20] The theory was, however, more successful in resolving other difficulties. Russell uses it, for example, to explain how true identity statements can be informative: at least one of the expressions flanking the identity symbol must be a denoting phrase (see *Principles of Mathematics*, §64, pp. 63–4). And Russell came to see that the theory could be extended to cover proper names (ordinary proper names, as opposed to what Russell later called 'logically proper names') quite generally. This extension resolves the issue of names which appear to name concrete existing objects, but where

here undermines one still very common account of Russell's motivation for adopting the theory of descriptions.

[19] There are, however, passages in Russell's writings, not written for publication, which suggest that he was attempting to find an explanation of denoting in terms of propositional functions. There is, however, no sign that he ever found a way of doing this which satisfied him – unless, indeed, one thinks of the theory of descriptions as being such an explanation. See especially, *Collected Papers*, vol. IV, pp. 340, 342. (In this essay I have not attempted to do justice to all the intricacies of Russell's thought suggested by his unpublished work.) In this note I am indebted to correspondence with Russell Wahl.

[20] He says: 'Thus *x* is, in some sense, the object denoted by *any term*; yet this can hardly be strictly maintained, for different variables may occur in a proposition, yet the object denoted by *any term* is, one would suppose, unique' (*Principles of Mathematics*, §93, p. 94). I am here attempting to do no more than indicate the difficulties which Russell encounters.

in fact there is no such object ('Pegasus' or 'Vulcan', for example). In *The Principles of Mathematics* Russell had denied that there are any such names: names which seemed to name nothing were said to name non-existent but still subsistent (and thus real) entities. But in fact the theory of denoting concepts has the resources to avoid that conclusion; it can thus avoid non-existent *concreta*, and the idea that we can be acquainted with such things.[21]

Russell himself, in Appendix A of the *Principles of Mathematics*, says that Frege's distinction between sense and reference is 'roughly, though not exactly, equivalent' to his own distinction between a denoting concept and the denoted object (see §476, p. 502). The most obvious difference is that Frege applied the distinction very widely, whereas for Russell it was far more restricted. The clear point of similarity is that in each case we have what we might speak of as a *representational* element in the object of judgement (Frege's thought, Russell's proposition). A paradigmatic subject–predicate proposition for Russell, one *not* containing a denoting concept, does not contain something which *represents* its subject; rather the subject itself is contained in the proposition. But when we employ a description we express a proposition which contains an element which does in this sense *represent* the subject; this element is of course the denoting concept corresponding to the description, for that denoting concept is not itself the subject of the proposition, not what the proposition is about.[22] Frege's senses, if we think in such terms about them, are clearly representational in the same sort of way: a thought is not about the senses which (perhaps) make it up, but rather about the references (if any) of the expressions whose senses they are.[23]

The theory of denoting concepts strongly suggests a picture according to which the structure of a proposition is, in general,

[21] For Russell's acknowledgement of these points, see, in particular, his 'The existential import of propositions', *Collected Papers*, vol. IV, pp. 486–9.

[22] Of course there can be propositions which have denoting concepts as their subjects, but such a proposition must contain not that denoting concept which it is about, but rather some other denoting concept which denotes it.

[23] This rather cumbersome way of speaking is necessary because for Frege it is an *expression* which has a sense and (in the usual case) a reference. For Russell, by contrast, it is the denoting concept, not a linguistic item, which denotes the object.

quite closely related to the structure of a sentence which expresses it. (It may be that Russell was in part led to the theory because he already held the general picture.) The proposition expressed by the sentence 'Every natural number is either odd or even', according to the theory of denoting concepts, expresses a proposition which contains a component corresponding to the words 'every natural number'. This component is of course a denoting concept, and for further progress in analysing the sentence we need to consider that denoting concept and its function. When we analyse the sentence, to gain insight into the structure of the proposition which it expresses, we retain its grammatical structure. The point is quite general: grammatical structure is taken as a good, though not infallible, guide to the structure of the underlying proposition; each word or semantic unit is assumed to correspond to an element in the proposition. Thus Russell says:

> The study of grammar … is capable of throwing far more light on philosophical questions than is commonly supposed by philosophers. Although a grammatical difference cannot be uncritically assumed to correspond to a genuine philosophical difference, yet the one is *primâ facie* evidence of the other … Moreover it must be admitted, I think, that every word occurring in a sentence must have *some* meaning … The correctness of our philosophical analysis of a proposition may therefore be usefully checked by the exercise of assigning the meaning of each word in the sentence expressing the proposition. On the whole, grammar seems to me to bring us much nearer to a correct logic than the current opinions of philosophers. (*Principles of Mathematics*, §46, p. 42)

The picture of analysis which this suggests is one which will go word by word, or phrase by phrase, rather than sentence by sentence. For the most part it will be taken for granted that a word or phrase in a sentence corresponds to some element in the proposition expressed by the sentence; the interesting question will then be as to the nature of that element. (Is it, for example, a denoting concept, and if so of what kind?) There is here no a general contrast between grammatical structure, or surface structure, and underlying or logical structure. On the contrary: we can, for the most part, read off the underlying structure from the structure of the sentence. To put essentially the same point a different way: language is conceived as a largely transparent medium, through which propositions may be perceived without systematic distortion; the transparency of the medium makes it possible largely to ignore it.

These ideas, like Russell's reliance on the notion of acquaintance, can be put in the context of his opposition to Idealism, and especially to the monism which he attributed to F. H. Bradley. Pluralism, the existence of many distinct things which (at least sometimes) stand in relations to one another, is immediately suggested by our ordinary discourse, by the surface of our language. If the surface of language is a generally reliable guide to the underlying structure, then propositions will indeed contain a plurality of objects in relation to one another. So Russell's opposition to Idealism gives him reason to hold that there is no systematic distortion here, that the grammatical structure of a sentence is in general a good guide to the underlying structure of the proposition which it expresses. This is an idea which, as we shall see, is in very marked contrast to the view he held after June 1905.

4

Russell's famous essay 'On denoting' rejects the theory of denoting concepts, and argues for the theory of descriptions. The essay contains detailed arguments against the theory of denoting, arguments which we shall not examine here.[24] The crucial thing to note about them is that they all operate within the context of Russell's direct realism. Within that context the theory of denoting concepts is an anomaly from the outset; once Russell sees how to avoid that theory he is very ready to do so. A crucial shift from the earlier view is that now Russell takes the idea of generality – 'the variable', as he says – as primitive and unexplained. The major motive for the theory of denoting was to explain generality – roughly, by treating the

[24] The interpretation of these arguments is very controversial. For a general account, see again, ch. 6 of the work cited in note 4. For an attempt to come to terms with the text in detail see Michael Pakaluk 'The interpretation of Russell's "Gray's *Elegy*" argument', in A. Irvine and G. Wedeking (eds.), *Bertrand Russell and Analytic Philosophy* (Toronto: University of Toronto Press, 1993). See also Harold Noonan, 'The 'Gray's Elegy' argument – and others', in R. Monk and A. Palmer (eds.), *Bertrand Russell and the Origins of Analytic Philosophy* (Bristol: Thoemes, 1996), pp. 65–102; and Michael Kremer, 'The argument of "On denoting"', *Philosophical Review*, 103 (1994), pp. 249–97.

phrase 'any object' as expressing a denoting concept. But, as we have seen, the theory of denoting did not in fact succeed in this task; Russell now abandons the goal entirely, and simply takes generality for granted, as primitive and unexplained. (But see note 29, below, for a qualification to this statement.)

Presupposing generality, Russell is then able to explain indefinite descriptions in the familiar manner: 'Every prime number is odd' is explained as 'For any object *x*, if *x* is a prime number then *x* is odd', and so on. He had seen the possibility of doing this as early as 1902, but at that stage it had not influenced his philosophical views. Definite descriptions presented more of a challenge; it was Russell's seeing how to treat them in the analogous way that made it possible for him to develop the new view. The analogous treatment of 'The President of the USA in 2000 was a Democrat' is to explain it as 'There is an object *x* such that *x* served as President of the USA in 2000 and *x* was a Democrat, and for every object *y*, if *y* served as President of the USA in 2000 then *y* is identical to *x*.' More briefly and idiomatically: 'There is one and only one thing which served as President of the USA in 2000, and it was a Democrat.'

The sentence we started with above is certainly about President Clinton. As analysed, however, it expresses a proposition which does not contain that man; it is *indirectly* about him. So one might think that here too, as in the theory of denoting, there is a violation of Russell's direct realism. But in fact this is not so: here there is no in-virtue-of problem. Here the idea of indirect aboutness does not rely on a mysterious relation of denoting, introduced only for this purpose. It relies, rather, on familiar ideas. The sentence is about Clinton because it contains a predicate, 'served as President of the USA in 2000', which holds of him and of no one else. This explanation uses the idea of a predicate's holding of, or being true of, an object; this is not an idea which is mysterious or objectionable in the same way that the idea of denoting is. In particular, it is not an idea introduced ad hoc to solve – or to label – this particular problem; it is, rather, an idea which is needed for quite general purposes in almost any account of language.

So one way of putting the point of the theory of descriptions is that it is to explain in a transparent and wholly unmysterious way what the theory of denoting 'explains' in a mysterious and ad hoc fashion: how a proposition succeeds in being about entities which it does not contain. Since the entities contained in a proposition that I understand must be entities with which I am acquainted, the theory equally explains, in non-mysterious fashion, how I can understand propositions about entities with which I am not acquainted. Russell says, for example: 'All thinking has to start from acquaintance; but it succeeds in thinking *about* many things with which we have no acquaintance.'[25] This idea is not new in Russell's thought in 1905. What is new is that he now has an explanation of indirect aboutness which does not appeal to an unexplained representational element. The explanation is not question-begging or ad hoc, and does not raise an in-virtue-of problem[26].

Russell now has no hesitation in extending this analysis to many phrases which grammatically are proper names, and treating them as if they were disguised or truncated definite descriptions.[27] He is thus left with a very small category of genuine (or logically) proper names; for those names, unlike others, their occurrence in a sentence does indicate that the sentence expresses a proposition in which the corresponding object occurs. Logically proper names can only be used to name objects with which the speaker is acquainted, and from 1905 on Russell holds that each person is acquainted only with a limited range of entities. (The range gets more limited as

[25] Bertrand Russell, 'On denoting', *Mind* (1905), p. 480; *Collected Papers*, vol. IV, p. 415.

[26] It might be thought that, by doing this, 'On denoting' vindicates direct realism. Certainly this is one of the aims of that work, but we should not exaggerate the extent to which it succeeds. The crucial qualification here is one which we have already mentioned: the theory leaves Russell wholly without an explanation of generality. According to the new theory, generality is involved in almost everything we say, yet it is entirely unclear how it fits into the picture of direct realism. It is all the more important to stress this point, in view of the fact that generality was the strongest of Russell's original motives for introducing the theory of denoting. In some writings after 'On denoting', which were not intended for publication, Russell speaks of 'On denoting' not as eliminating denoting but rather as reducing it all to a single case, that of the variable. See Hylton, *Russell, Idealism*, pp. 255ff.

[27] Russell speaks of the name 'Romulus' as 'a sort of truncated description' in the sixth of his 'Lectures on the philosophy of logical atomism', *Collected Papers*, vol. VIII, p. 213. In *The Problems of Philosophy* (Oxford: Oxford University Press, 1912, reset 1946) he says: 'Common words, even proper names, are usually really descriptions. That is to say, the thought in the mind of a person using a proper name correctly can generally only be expressed explicitly if we replace the proper name by a description.' (p. 54).

time goes by; this trend started before 'On denoting', and is to some extent independent of it.) So philosophical analysis is required to show that sentences about other entities are only indirectly about them, and to work out what such sentences are directly about. Only entities with which we are acquainted can occur in propositions we can grasp. Almost all of our knowledge appears to violate this dictum, and so must be analysed to show that it does not in fact do so. Russell's position thus commits him to an extensive programme of analysis which is, in the broad sense, epistemologically driven: by the need to show how we are able to think about various entities with which we are not acquainted. It is this programme which issues in such works as *Our Knowledge of the External World*. (Such a programme of analysis, it need hardly be said, has no analogue in Frege's work; Frege's general philosophical views simply do not give rise to a need for anything of the sort.)

The theory of descriptions assumes enormous importance for Russell. He quickly comes to hold that we are acquainted with almost none of the concrete objects that we take ourselves to know about. (His reasons essentially have to do with the possibility of error and illusion. In the case of abstract objects he is more willing to accept that we are acquainted with the things we appear to know about.) So most of what we take to be our knowledge about things is descriptive knowledge, not knowledge by acquaintance. And all such knowledge, so he holds from 1905, is to be explained along the lines laid down by the theory of descriptions. The theory of descriptions is *the* method of analysis, and hence of the first importance for Russell's epistemology.

One immediate consequence of the developments we have been discussing is a complete repudiation of the idea that (surface) grammar is, in general, a good guide to the form of the underlying proposition. We briefly examined this idea, as it occurs in the *Principles of Mathematics*, and saw that it is a natural concomitant of the theory of denoting concepts. In that theory, a subject-predicate sentence, with a description (definite or indefinite) for the subject, is taken to express a proposition with subject-predicate form, with a denoting concept taking the place of the subject. In the new theory, however, from 1905 onwards, such a sentence is taken to represent a proposition with a wholly different form. (A sentence containing a definite description expresses an existentially quantified proposition.)

From 1905 on, Russell's work is full of warnings that the structure of a sentence, its surface grammar – is almost always misleading as to the form of the underlying proposition. The goal of analysis remains, as before, the production of a sentence which accurately reflects the proposition expressed by the original sentence. But now the emphasis is very much on the *form* of the proposition (and, hence, of the sentence produced by the process of analysis). This form will not in general be the same as that of the sentence being analysed; this will, indeed, hardly ever happen. Logical forms become the focus of analysis.

This change in turn has a consequence which may at first sight appear paradoxical. Precisely because it is misleading, language, which Russell had previously more or less ignored, becomes an increasing concern. When he thought of language as a more or less transparent medium, through which the proposition could be readily perceived, Russell could afford to pay it no special attention; he proceeded at once to talk of the underlying proposition, his true concern. But after 1905 he has to be self-conscious about language, if only to avoid being misled by it.[28] Before 1905 all of Russell's remarks about language (in the sense of the actual words) are casual and superficial, not in any sense part of a theory of language. After 1905 this begins to change. With the notion of an *incomplete symbol* we have, for the first time in Russell's work, a technical term which is quite explicitly and exclusively linguistic: some symbols are incomplete, but no constituents of propositions are (in anything like the same sense) incomplete. Russell is driven to pay attention to language precisely because of its misleadingness; one might say that it is here that 'Philosophy of Language', in something approximating its modern sense, comes into being.

[28] Thus: 'There is a good deal of importance to philosophy in the theory of symbolism, a good deal more than at one time I thought. I think the importance is almost entirely negative, i.e., the importance lies in the fact that unless you are fairly self-conscious about symbols, unless you are fairly aware of the relation of the symbol to what it symbolizes, you will find yourself attributing to the thing properties which only belong to the symbol' ('Lectures on the philosophy of logical atomism', p. 166). Later in the paragraph Russell says that good philosophers think about the real philosophical concerns, as opposed to symbols, for a minute every six months, whereas bad philosophers never do.

Another consequence of the new paradigm of philosophical analysis is a shift in the role played by the idea of acquaintance, or rather the reinforcement of a shift which was already underway. This is a complex and subtle matter, and concerns shifts in Russell's attitude as much as real changes in doctrine. In the *Principles of Mathematics* the notion of acquaintance had functioned, more or less, as a 'dependent variable': if Russell's philosophical analyses made it expedient for him to claim that we are acquainted with a certain entity, then he would make the claim.[29] After that work the notion of acquaintance comes increasingly to impose independent constraints upon analysis; the results of a preliminary philosophical analysis are to be checked by seeing whether we are in fact acquainted with the entities which, according to the analysis, we must be. The results of this checking were not wholly independent of the exigencies of the analysis; still Russell is increasingly restrictive in his view of what entities we are acquainted with. It is for this reason that one finds in Russell's work after the *Principles of Mathematics* appeals to 'inspection', which is meant to remind us with which entities we are actually acquainted; these facts are supposed to constrain philosophical theorizing.

This trend towards greater psychological realism about acquaintance begins before 'On denoting', but the new view greatly encourages the trend. Before 1905, analysis takes the form of a sentence for granted, and aims to clarify our understanding of the parts; the analysis is complete, presumably, when we have a clear understanding of each of the parts of the sentence. It may not always be evident how we are to know when the analysis is complete, but at least it makes sense to think that each step is bringing us closer to the fully analysed sentence, and thus to the form of the proposition itself. But after 1905 the process of analysis may at any step reveal a wholly new logical form. There is no particular reason to think that the seventeenth step in the progressive analysis of a sentence is closer to the real form of the proposition than is the thirteenth. How are we to know that we have reached the terminus of analysis, if we cannot easily think of ourselves as getting closer and closer to it?

[29] This may overstate the matter to some extent, but not by much. I owe the comparison with the idea of 'dependent variable' to Andrew Lugg.

The notion of acquaintance comes to provide an answer to this question: the terminus of analysis is reached when we have a sentence where each term refers to an object with which we are acquainted. But of course this answer presupposes that acquaintance functions, at least to some extent, as an 'independent variable', and is not simply answerable to the needs of the analysis.

Over the last few paragraphs we have been emphasizing the difference that the theory of descriptions makes to Russell's view of how analysis proceeds, and something of the wider significance of this shift. But it is also important to stress that there is an underlying continuity in Russell's conception of analysis. It is, one might say, the same sort of question he is trying to answer before 1905 and after, even though the answers he gives are not the same. The question is: what are the constituents of this proposition? Russell continues, that is to say, to conceive of a proposition as a complex entity made up of simpler entities, in something like the way a wall is made up of bricks. This general conception largely survives even Russell's adoption of the so-called 'multiple relation theory of judgement', which involves his abandoning the idea that propositions exist as objective entities independent of us (we shall discuss the new theory at more length shortly). The question of the constituents of a proposition is simply reframed, to ask about the constituents involved in a judgement; the underlying conception does not seem to change. We shall return to these matters in in §4.

5

As we have seen, the theory of descriptions was, for Russell, a way of defending his direct realism (at least if one does not focus on the issue of generality). Let us come at this from a different angle, by seeing exactly how the theory of descriptions enables Russell to avoid any version of the distinction between sense and reference for singular referring expressions. In the next section we shall draw on these ideas to articulate a further consequence of Russell's position, concerning the notion of a function; this is, again, directly relevant to the contrast between Frege and Russell.

I shall consider two kinds of reason for holding that there must be a distinction analogous to Frege's distinction of sense from reference, and argue that in each case the theory of descriptions enables Russell to avoid that reason. The first kind of reason is straightforward: there are empty names, names which name nothing, such as 'Vulcan' and 'Pegasus'. If understanding a name consists in being related to an object, then it would seem that one cannot understand an empty name. Yet we do seem to understand sentences containing such names. Frege accounted for this by saying that in such a case the name has a sense, and hence it is possible to understand it even though it lacks a reference. He takes it to be a consequence of this view that a sentence containing such a name will also have a sense, and hence be capable of significant use, but lack a reference, i.e. lack a truth-value.

Russell's approach is quite different. He claims that (apparent) names are of two wholly different kinds.[30] On the one hand there are logically proper names, which function simply as labels which the speaker affixes to objects with which he or she is acquainted. These names function in accordance with Russell's paradigm; they are, however, very rare, at least in our ordinary language. On the other hand there are all the other (apparent) names. These are not, by Russell's standards, genuine names at all. Sentences in which they appear are to be analysed in accordance with the theory of descriptions, and in the analysed sentences the apparent names do not appear. (I shall sometimes call these apparent names 'descriptive names', just to have a label for them.)

Now it is Russell's view that logically proper names *cannot* be empty: If I can use a word as such a name then I am acquainted with its bearer, and this is not possible unless there is such an entity. I thus have an epistemological guarantee that the name is not empty. Names which lack this guarantee are not logically proper names, but merely descriptive names. When a sentence containing an (apparent) name of this latter sort is analysed we obtain a sentence in which the given name does not appear at all. (Hence Russell's view that these names are not genuine names at all: they do not survive analysis.) So for Russell there is no problem of empty names. Genuine names, logically proper names, cannot be empty. Other

[30] Note that this kind of contrast can be drawn, to much the same effect, within the theory of denoting concepts. I am here not concerned to contrast that theory with the theory of descriptions, but rather to contrast the latter with Frege's view.

apparent names are not really names at all, and hence cannot be empty names. (Russell's approach also has the advantage that a sentence containing an empty descriptive name will have a truth-value; avoiding truth-valueless sentences in this way which will make for a smoother logic. Frege achieves the same end by stipulating a referent for any singular referring expression which would otherwise be empty.)

There is also a second sort of reason for introducing some version of distinction between sense and referencewhich is rather more complex, and shows up in various ways. The underlying point could be put like this: two singular referring expressions with the same referents, such as 'Socrates' and 'the teacher of Plato', may nonetheless have different semantic roles. So saying that it refers to a certain object cannot be the whole story about the semantic role of such an expression, and understanding such an expression cannot consist simply in being in some relation to its referent. Therefore the semantics of such an expression must take account not only of what Frege calls the reference of the expression but also of something else, and this something else will be at least analogous to what Frege calls the sense of the expression. Let us flesh out this argument by seeing why we cannot take a singular referring term's referring to the object that it refers to as the whole story about its semantic role. (Our doing this will also indicate what the idea of 'semantic role' comes to here, for we should not take that idea to be self-evident in this context.)

Consider a true statement of identity, such as 'Socrates is the teacher of Plato'. It is clear that someone may understand the sentence without knowing whether it is true – or while being convinced that it is false.[31] This possibility seems to be straightforward, and to arise in quite ordinary cases. But if the two expressions flanking the identity symbol have the same semantic role, then such a case would appear to be impossible, or at least quite anomalous. The sentence seems to convey information, whereas the sentence 'Socrates is Socrates' does not. If the two (apparent) singular referring expressions, 'Socrates' and 'the teacher of Plato', have the same semantic role, then it is hard to see how this can

[31] Frege begins 'On sense and reference' by talking about cases of this sort.

be so. Similarly, there is the phenomenon now known as 'referential opacity'.[32] John may believe that Socrates died from drinking hemlock, while not believing that the teacher of Plato died from drinking hemlock. So John's understanding of each expression must involve more than simply a relation to object to which it refers; the word 'Socrates' must have a different semantic role from that of the phrase 'the teacher of Plato'. Yet another way of getting at what is, I think, the same underlying issue has to do with inference. From 'All teachers are wise' and 'Plato had exactly one teacher' we may immediately infer 'The teacher of Plato is wise'. But we cannot, from the same premises, infer 'Socrates is wise'. If the two expressions had the same semantic role, however, then we should be able to do so.

These considerations may be put in more Russellian terms by speaking of objects occurring in propositions, rather than of semantic roles. It is highly implausible to think of our sentence 'Socrates is the teacher of Plato' as expressing a proposition which simply contains the same object twice over, along with the notion of identity. The two expressions function differently for us: we understand them differently, they may play different roles in our expression of belief, and in the inferences we recognize as valid. So we cannot happily think of them as indicating the presence of the same object in a proposition, unless we think of our grasp of propositions and their constituents as itself mediated. But that would undercut Russell's direct realism, the aim of which is precisely to avoid the idea that there is anything mediating between us and the objects that we hope to think about.

On a straightforward, or superficial, view of what things count as singular referring expressions, then, we cannot think of coreferential singular expressions as always playing the same semantic role. In Russellian terms, such expressions cannot be thought of as always merely indicating the presence of the corresponding

[32] The term is W. V. Quine's; see *From a Logical Point of View* (Cambridge, Mass.: Harvard University Press, 1961), p. 142. As he acknowledges, he draws on Russell's use of the term 'transparent' in Appendix C of the second edition of *Principia Mathematica*, 3 vols. (Cambridge: Cambridge University Press, 1925–7), vol. I, p. 665. The underlying point is, again, made by Frege in the first few pages of 'On sense and reference'.

object in the proposition. Frege deploys the distinction between sense and reference to resolve all of these problems. How is Russell to resolve them, without resorting to any analogous distinction? As in the case of empty names, the distinction between logically proper names and merely apparent names (descriptive names) is crucial. In a fully analysed sentence, he holds, no descriptive names occur; hence the question of the semantics of such names does not arise. And for real names, logically proper names, Russell simply denies the applicability of the pressures which might lead one to make some version of the distinction between sense and reference. If a given speaker has two logically proper names for a given object, then that speaker will be aware that the two names name the same object. Acquaintance gives us complete and unmediated knowledge: you cannot be acquainted with the same object twice over and not know it, for there are no *ways* of being acquainted with an object. Again, a logically proper name lacks any semantically significant structure, and gets its meaning, for a given speaker, simply by being a label for an object with which that speaker is acquainted. A logically proper name thus has no semantic structure which can be exploited in inference; it is in this sense a simple referring expression, not a complex referring expression.

Let us put these points another way. The considerations we examined seem to show that there must be more to the semantics of a singular referring expression than the fact that it picks out a certain object. Russell's logically proper names form an exception, but in general there is a need for an account at the level of sense as well as for an account at the level of reference.[33] One way to understand this is in terms of semantic complexity. In the case of definite descriptions this complexity is right on the surface, for they are made up of semantically significant parts. Russell assimilates descriptive names to definite descriptions, treating them as covertly complex in the same way. On his account, then, the apparent need for sense arises from the semantic complexity of most singular referring

[33] For an elaboration of this point, see Hylton, 'Functions and propositional functions in *Principia Mathematica*', in Irvine and Wedeking (eds.), *Russell and Analytic Philosophy*, pp. 342–60; reprinted as essay seven of Peter Hylton, *Propositions, Functions, and Analysis: Selected Essays on Russell's Philosophy* (Oxford: Oxford University Press, 2005), especially §2.

expressions. The semantics of a complex referring expression cannot be understood simply in terms of what it refers to; its semantic complexity must also be taken into account.

Now Russell's theory of descriptions avoids this argument by simply denying that there are any complex referring expressions. This, I think, is in part what he means by saying that descriptions have no meaning in isolation:[34] what is being denied is that such phrases are referring expressions. Definite descriptions look for all the world like complex referring expressions, but it is not hard to see how the theory of descriptions avoids treating them as such. A phrase of the form 'The *F*' is accorded a meaning only in the context of a sentence, in which we say something of the form 'The *F* is *G*'. And this sentence is analysed as having the underlying structure: there is an object which is *F*, that object is also *G*, and no other object is *F*. Here, in the analysed form, we have occurrences of the predicate '… is *F*' but not of the (apparent) complex referring expression 'the *F*'. For Russell, complex referring expressions are merely apparent, misleading superficial features of language which do not correspond to anything in the underlying structure.

In Russell's view, then, the only genuine referring expressions (for a given speaker) are those which are simple, i.e. lacking semantically significant structure, and which get their significance by referring to entities which are objects of acquaintance (for that speaker). These features ensure that for those expressions no analogue of the distinction between sense and reference is called for in the case of such expressions. Apparent referring expressions which do not meet these criteria are to be analysed away – to be shown to be merely apparent. The theory of descriptions supplies the means of analysis here.

6

To this point we have discussed Russell's direct realism, his consequent rejection of any analogue of Frege's distinction between sense

[34] In Lecture VI of the 'Lectures on the philosophy of logical atomism' Russell says that incomplete symbols, by which he means to include descriptions, 'have absolutely no meaning in isolation, but merely acquire a meaning in context' (*Collected Papers*, vol. VIII, p. 221). In 'On denoting' he says that such phrases 'never have any meaning in themselves' (*ibid*., vol. IV, p. 416).

and reference, and his use of the theory of descriptions to mitigate what would otherwise be the implausibilities of this view. This nexus of Russellian views is closely connected with sharp differences from Frege on fundamental metaphysical issues.

Let us begin with conceptions of philosophical analysis, for this is of the first importance. 'Analysis' here is no mere convenience, not a merely pragmatic point of philosophical method. The correct method of analysis is the correct way to understand the world; this corresponds to – and reveals – the fundamental nature of the world. For Frege, the method of analysis is function and argument. His notion of a function is essentially a clarified and extended version of the familiar mathematical notion, and he takes it as philosophically primitive. Concepts are treated as special cases of functions: they are those functions whose values are always truth-values. So a predicate such as '... is prime' is taken to stand for a function. Applied to some objects this function yields the truth-value *True* as its value; applied to others it yields the truth-value *False*. The idea of a function's taking one object as argument, and yielding another as value, is simply taken for granted here. There is no sensible question as to why a certain function applied to a given argument yields the value it does: that it does so is the unexplained fact in terms of which other things are to be explained.

These general Fregean views are sharply opposed to Russell's; the issues which we examined in the last section are directly relevant to this opposition. Functional expressions, if taken as primitive, give rise to complex referring expressions. The expression '2 + 3', if taken at face-value, picks out the number five, and does so in a complex way. Saying what the expression picks out is clearly very far from being a full and adequate account of its semantic function. To understand the functioning of the phrase we need a distinction between sense and reference. So if the general notion of a function is fundamental, a semantic account must deal with (something analogous to) sense as well as reference. Since Russell rejects any such idea, he also denies that functional expressions in general are primitive. Hence nothing like Frege's function-argument analysis is available to him as a fundamental way of understanding the world. (We shall see at the end of this section that Russell does take as primitive the notion of a *propositional* function, and we shall consider why the reasons he has against taking functions in general as primitive do not apply to this special case.)

The contrast that I am drawing between Russell and Frege, then, is this. Frege takes the notion of a function as primitive; his doing so commits him to a distinction between sense and reference. But Russell denies that there is such a distinction. He therefore cannot accept the general notion of a function as primitive, and cannot accept Frege's fundamental mode of analysis. This leaves us with two questions. First, what is Russell's fundamental method of analysis? The answer takes us immediately to his ontological views, especially about the nature of complexity. Second, if Russell does not take functions as primitive, how does he account for them? The two questions are connected: the first provides the constraints within which the second must be answered. Russell must have an understanding of functions which is compatible with his general view of the nature of the world. I shall consider the two questions in turn.

Russell's view of analysis is based on his atomistic conception of the world. He sees it as made up of simple objects standing in relations to one another. What appear to be complex objects are to be understood as simpler objects standing in certain relations to one another. The complex object is made up of simpler objects as a whole is made up of its parts. (This relationship sometimes seems to be understood in a very literal way: a complex object is made up of simpler objects in the way in which a wall is made up of bricks.) Propositions provide a crucial example. A proposition, for Russell, paradigmatically *contains* the objects which it is about; they are the parts, and the proposition is the whole. Propositions here, however, are more than an example. That certain objects stand in certain relations is itself a proposition. So by treating all complexity as the complexity of relations and relata Russell is implying that all complexity is propositional complexity. (We shall enter a partial qualification to this point shortly.)

The theory of descriptions, seen from this perspective, eliminates an apparent exception to the idea that all complexity is propositional complexity, namely complex denoting concepts. It is for this reason, I think that Russell throughout 'On denoting' speaks of 'denoting *complexes*'; it is the complexity, as well as the denoting, that he is concerned to eliminate. This terminology may also reflect the idea that denoting is not wholly eliminated, but rather reduced to one simple case, that of the variable. (See note 26, above.) A phrase such as 'the present King of France' is not explained by saying that it indicates the presence in the proposition of a complex object, a denoting concept. Rather, it is explained in terms of the logical form of the whole proposition: there is one and only one object such that it currently reigns over France, and that object has whatever property the sentence ascribes. The semantic complexity of a definite description is thus accounted for in terms of the complexity of the complete proposition, not in terms of the complexity of any constituent part of it.

I have been speaking here of 'propositional complexity', the kind of complexity that is characteristic of a proposition. A change in Russell's views is relevant here. Sometime between 1906 and 1909 he comes to adopt what he calls 'the multiple relation theory of judgement'.[35] According to this theory, the notion of a proposition is not fundamental; it is replaced as the fundamental metaphysical idea by the notion of a fact. Propositions are explained in terms of facts, rather than vice versa. Russell continues, however, to think of all complexity as arising from simple objects standing in relations to one another. Under both the old view and the new view, this is the kind of complexity which typifies propositions. The difference it that, according to the new view, this sort of complexity is to be understood as being, at bottom, the complexity of a fact. For our general comparison between Russell and Frege this change is, I think, of relatively little importance. It does, however, make Russell's view in two significant ways less like Frege's. First, Russell now abandons the idea that the bearers of truth and falsehood are objective and mind-independent entities. Second, Russell had earlier held, with Frege, that truth is indefinable; with the new view, he advocates a version of the correspondence theory of truth.

[35] Russell first adumbrates, but does not endorse, this view in the 1906 essay 'The nature of truth'; the introduction to the first edition of *Principia Mathematica*, however, explicitly advocates the new view. It is perhaps worth adding that this view does not seem to fit with the logic of *Principia Mathematica*, which quantifies over propositions; also that Russell never found a version of the view that satisfied him for very long.

Russell thus conceives the world as consisting of complex objects made up of simpler objects and, ultimately, presumably, of absolutely simple objects.[36] His dominant mode of analysis is, accordingly, the decomposition of a whole into its parts. Frege sees the world as divided into functions and objects. One consequence of this difference concerns the stratification of the universe into ontological categories. Functions and objects are naturally conceived of being of distinct ontological kinds, with functions themselves coming in various levels which are similarly distinct: first-level functions apply to objects, second-level functions to first-level functions, and so on. Russell's fundamental metaphysical instincts are to deny any such distinctions; a whole is not naturally thought of as being of a different ontological kind from its parts – a wall is not of an ontologically distinct category from the bricks which compose it. The dominance of the part-whole metaphor suggests that there are no fundamental ontological distinctions, that all entities are of the same general kind. In the *Principles of Mathematics*, indeed, Russell argues that no fundamental ontological distinctions are tenable: everything is, in Russellian jargon, a term, that is, very roughly, capable of being a logical subject in the simplest kind of subject-predicate proposition. To deny that something is a term is, he claims, logically self-refuting, since *a* appears as a term in the proposition expressed by '*a* is not a term'.[37] I speak here of Russell's metaphysical instincts because he is forced, by the need to escape the paradox which bears his name, to acknowledge fundamental distinctions, in the form of the theory of types. Those distinctions, however, always seem to be imposed, for the purpose of avoiding the paradox, upon a structure in which no such distinctions exist. For Frege, by contrast, the distinctions between function and object, and among functions of various levels, are built in to his thought

[36] 'Presumably' because Russell does, at least at one point, suggest that it would be possible to maintain that analysis is infinite, 'that complex things are capable of analysis *ad infinitum*', though he does not accept this view. See the discussion at the end of the second of the 'Lectures on the philosophy of logical atomism', *Collected Papers*, vol. VIII, p. 180.

[37] We might phrase this by saying that Russell takes very seriously the concept horse problem, whereas Frege wants to dismiss it as due to a mere awkwardness of language. See Appendix A of the *Principles of Mathematics*, especially §§481–3.

from the outset. (We shall return to these points in the next section, putting them in the context of the logics developed by Frege and by Russell.)

An illustration of Russell's view, and an important fact in its own right, is the difficulty that he faces in accounting for the unity of the proposition. The constituents of a proposition 'placed side by side', Russell says, 'do not reconstitute the proposition' (*Principles of Mathematics*, §54). 'A proposition is essentially a unity, and when analysis has destroyed the unity, no enumeration of constituents will restore the proposition' (*loc. cit.*). How is this unity to be understood? From within Russell's early post-Idealist metaphysics the unity of the proposition can be neither avoided nor explained.[38] Frege, by contrast, is not troubled by any analogous problem. For him there is no issue about how judgements are possible, about how concepts and objects unite. From a Russellian perspective, it might appear that he is simply ducking a problem, but in fact I think we have here an indication of how different his presuppositions are from Russell's. Let us focus on Frege's Context Principle: 'it is only in the context of a proposition that words have any meaning'.[39] This principle, as I understand it, implies that the notions of an object, and of a concept, are not to be understood independently of one other, and of the role that concept-expressions and object-expressions have in forming complete sentences.[40] On this kind of reading, Frege presupposes the notion of judgement as fundamental, and understands both concepts and objects in terms of it. For him there thus can be no question as to how these separate and independent entities can form a unity, since they are not correctly thought of as separate and independent at all.

[38] It may have been the ramifications of this issue that were responsible for the major change in Russell's metaphysics that took place when he adopted the multiple relation theory of judgement. See Thomas G. Ricketts, 'Truth and propositional unity in Early Russell', in Juliet Floyd and Sanford Shieh (eds.), *Future Pasts: The Analytic Tradition in 20th Century Philosophy* (Oxford: Oxford University Press, 2001), pp. 101–23.

[39] *Gl*, §62; cf. also p. x and sections 60, 106.

[40] For this line of interpretation see, for example, Thomas G. Ricketts, 'Objectivity and objecthood: Frege's metaphysics of judgement', in L. Haaparanta and J. Hintikka (eds.), *Frege Synthesized: Essays on the Philosophical and Foundational Work of Gottlob Frege* (Dordrecht: Reidel, 1986), pp. 65–95.

A page or two back, we saw that there is a clear ontological difference between Frege and Russell: Frege sees the world as divided into functions (of various levels) and objects; Russell, with a view dominated by the part-whole metaphor, rejects functions, and cannot easily adopt any such distinctions at a fundamental level. Our recent discussion, however, suggests that as well as this ontological difference there is also a difference in the very notion of ontology that is at issue here. Russell holds what one might call an object-based metaphysics: for him the existence of an object is a fundamental and independent fact, the idea of an object's existing or not existing makes sense by itself, in isolation from other ideas.[41] For Frege, by contrast, the fundamental ideas are those of truth and falsity, and of a judgement as that to which truth and falsity can be ascribed.[42] Here ontology is derivative: questions of existence are to be settled primarily by seeing what is required for the judgements that we make, and to account for the way those judgements behave in inferences that we make.

These metaphysical differences are connected with differences in epistemology – not just in the answers to epistemological questions, but also in the questions themselves. For Russell, as we have emphasized, the notion of acquaintance is crucial. The idea of an object's existing or not existing draws on our (supposed) capacity for acquaintance, our ability to stand in a direct cognitive relation to an object. Our knowledge and understanding must all ultimately to be explained in terms of this relation. This imperative defines a philosophical task: since most of our knowledge and understanding seems to concern things which are not objects of acquaintance, we need to show how it can be explained in terms of acquaintance.

[41] As we saw, Russell distinguishes existence from subsistence in his early post-Idealist work (even as late as the 1912 *Problems of Philosophy*; see p. 100). Here, however, I use the word 'existence' here broadly, to encompass both ways of being.

[42] Thus Frege says: 'What is distinctive about my conception of logic is that I begin by giving pride of place to the content of the word "true", and then immediately go on to introduce a thought as that to which the question "Is it true?" is in principle applicable. So I do not begin with concepts and put them together to form a thought or judgement; I come by the parts of a thought by analysing the thought.' This passage is from notes that he wrote about his thought for Ludwig Darmstaedter, and is published in *NS*, p. 273; I follow the translation in *PW*, p. 253.

A foundationalist epistemology is thus implicit in Russell's general view. He assumes that all knowledge is based on our acquaintance with certain objects, some of them abstract (he is somewhat open-minded about exactly which objects, and changes his mind about this over time). So he then needs to show how, and to what extent, the knowledge and understanding which we take ourselves to have can be explained on this basis, and thereby justified. Here there is a very sharp contrast with Frege. Frege does not seem to be at all concerned to raise questions about the basis of our knowledge, how it is acquired and what ultimately justifies it. Nor does his fundamental view seem naturally to generate such questions. (Unless, of course, such questions are inevitable and thus naturally generated by any serious thought; the point is that nothing peculiar to *Frege's* thought naturally generates such questions.) Frege seems, rather, to think of the philosophical task as primarily one of systematizing knowledge, setting out the relations of justification which hold among the various items we know. Axiomatization, of Euclidean geometry, for example, serves as a partial paradigm here, but in the ideal this model would be extended both deeper, to include the underlying logic, and wider, to include all systematic knowledge. In this way the body of our knowledge will be given greater clarity, and our understanding of exactly what it is that we know may be modified in the process. Russellian foundationalist questions, however, have no place in Frege's work; nor does scepticism play any role for in his thought.[43]

These sorts of differences are, of course, most evident in the case of our knowledge of mathematics and logic, for these subjects are at the centre of Frege's concerns. For Russell, as we saw, our knowledge of these subjects must be based on acquaintance.[44] Philosophical analysis may be required to show you *which* abstract objects play

[43] For elaboration of these ideas, see §I of Thomas J. Ricketts, 'Frege's 1906 foray into metalogic', *Philosophical Topics*, 25 (1997), on which my discussion in this paragraph draws.

[44] Russell changed his mind about this under the influence of Wittgenstein. Beginning with his lectures, 'The philosophy of logical atomism', given early in 1918, he speaks of the truths of logic as 'tautologies'; see the end of Lecture V. This tendency is more marked in the book he wrote later that year, *Introduction to Mathematical Philosophy*, where the position is somewhat elaborated. In these

a fundamental role – whether, for example, it is numbers or classes or propositional functions. But the fundamental abstract objects, whichever they turn out to be, are known by acquaintance. The objects are out there, and we are capable of standing in a direct cognitive relation to them. Russell's version of realism about abstract objects is thus backed up by his epistemology. Nothing similar can be said about Frege, and this has been taken to cast Frege's realism about abstract objects in doubt; those who take Russellian views as paradigmatic may indeed find Frege's realism less than robust. It would, however, be more accurate to say that in the context of different epistemological and metaphysical views, what realism comes to also differs.

As we have seen, the notion of a function cannot be primitive for Russell; functional expressions must be explained in other terms. It is to this explanation that I now turn. What Russell does is to define functional expressions in general in terms of expressions for what he calls 'propositional functions'. A propositional function is, very roughly, the non-linguistic correlate of an open sentence, i.e. a sentence containing one or more variables.[45] In a footnote in the Introduction to the first edition of *Principia Mathematica*, Whitehead and Russell say explicitly: 'When the word "function" is used in the sequel, "propositional function" is always meant'.[46] And *30 of that work is devoted to showing how non-propositional functions – descriptive functions, as they are there called – can be introduced on the basis of propositional functions. Roughly the idea is this: we do not begin by presupposing, say, the two-place plus function; we begin with the three-place propositional function represented by 'ADD(x, y, z)'. (Where this is read as 'The sum of x and y is z', so that 'ADD(2, 3, 5)' is a true sentence, 'ADD(5, 3, 2)' a false

works, however, the new idea sits very uneasily alongside the earlier position, so that it is hard not to think that Russell is simply using the Wittgensteinian form of words without really having thought it through, or even without really understanding it. It is the earlier position which I attribute to Russell here.

45 I thus claim that Russell uses 'propositional function' to refer to abstract objects, rather than using it to refer to linguistic objects, or in such a way that is unclear which sort of object he means to be referring to. This claim is controversial; for some defence of it, see Hylton, *Russell, Idealism*, especially pp. 217ff.

46 A. N. Whitehead and B. Russell, *Principia Mathematica*, 3 vols. (Cambridge: Cambridge University Press, 1910–13), vol. I, p. 39.

one, and so on.) The plus function, '$x + y$' is then introduced by definition:

$$x + y \text{ is defined as: the object } z \text{ such that } \textsc{add}(x, y, z)$$

This technique enables us to define an n–1 place descriptive function on the basis of any n-place propositional function which satisfies the relevant uniqueness condition: that for any given selection of n objects in places corresponding to the arguments of the descriptive function there should be exactly one object which makes the propositional function true. (Each definition of this sort, one for each non-propositional function that we want, will of course employ a definite description; this, I suspect, does something to explain the importance that Russell attributes to definite descriptions.)

The method of defining functions (descriptive functions, in Russell's sense) from propositional functions is technically quite straightforward. (No function is defined unless the propositional function satisfies the appropriate uniqueness condition, but this is the desired result.) What is problematic is to see exactly why Russell is willing to accept propositional functions as primitive, while he is not willing to accept functions in general as primitive. Clearly he is not thinking of propositional functions simply as a special case of functions, as a species of the genus *function*: but why not? How do propositional functions, in his view, differ from descriptive functions?

Recall the reason that Russell cannot accept a functional expression, such as '2 + 3' at face value, as a complex referring expression. Doing so would give rise to a need for an account of the semantic role of such phrases which requires some distinction analogous to the Fregean distinction between sense and reference. The reason for this is that the phrase has a semantic complexity which is not to be found in the object which it picks out. Thus if there were no more to the semantics of the phrase than its picking out a certain object, we would have no way of taking account of that complexity. This would make it impossible to understand the role that the phrase plays in language (in inferences, in particular). But propositional functions are in the relevant way unlike functions in general. A phrase expressing a propositional function, '$x + y = z$', for example, gives rise to sentences, '2 + 3 = 5', for example. On Russell's view, a sentence is related to – expresses, picks out – a proposition.[47] And a proposition *does* possess the requisite complexity.

Saying of the expression '2 + 3' that it refers to the number five is far from an adequate account of its semantics for there is, so to speak, no complexity in the number five which corresponds to the complexity of the functional expression. There is no way to understand the complexity in terms of relations and relata, of parts and wholes. Saying of '2 + 3 = 5' that it expresses the proposition that two plus three equals five, however, is, from a Russellian point of view, quite a different matter. For propositions are complex in just the ways that are needed. In particular, a proposition which is the value of a propositional function applied to a given object as argument will *contain* that object. And the resulting proposition has the same form as the propositional function. (Indeed we might think of a propositional function as simply being the form of a number of propositions.) Two propositions which are the values of a single propositional function have something in common in virtue of that fact. And from the proposition we can figure out of which propositional functions it is a value, for the proposition has a kind of complexity which marks its relation to the propositional functions of which it is a value. The propositional function, we might say, is *recoverable* from the proposition. None of these points applies to functions in general.[48]

The facts indicated above show why propositional functions will, while functions in general will not, fit into Russell's metaphysics. A function takes an object as argument and yields as value an object which bears no obvious systematic relationship to the argument or to the function itself; in particular, the value may be simple and unanalysable. A propositional function, by contrast, takes an object as argument and yields as value an object of a special kind – a proposition – which *does* have such systematic relationships: it contains the argument, and has the same structure as the propositional

[47] Although I speak here of propositions, and objects as constituents of propositions, what I say holds good also, *mutatis mutandis*, of the view that Russell holds after he adopts the multiple relation theory of judgement, briefly discussed above.

[48] These matters are discussed in somewhat greater detail in my 'Functions and propositional functions', in *Principia Mathematica*, reprinted as essay seven of Hylton, *Propositions, Functions, and Analysis*.

function.[49] What is unexplained in the case of a function – that *that* object taken as argument should yield *this* object as value – is transparent in the case of propositional functions.

7

The differences between Frege and Russell emphasized in the previous sections are relevant to the accounts that each gives of logic. One point is this. For Frege there is, from the outset, a fundamental difference in kind between functions and objects, with concepts defined as a special case of functions. The idea of a concept's applying, or not applying, to itself is, for him, intrinsically absurd. A consequence of this is that no analogue of Russell's paradox arises directly from his fundamental metaphysics. By the same token, however, the ontology of that metaphysics is too weak to carry out the reduction of arithmetic to logic.[50] For that purpose it is necessary to bolster the fundamental ontology with an additional assumption. It is for this reason that Frege's system of logic in the *Grundgesetze* includes the notorious Axiom V, which asserts (roughly) that for every concept there is a corresponding object. This axiom gives Frege's system of logic the power necessary to carry out the logistic reduction, but it also, notoriously, leads his system into contradiction.

For Russell's logic the situation is reversed. The power needed to carry out the reduction is intrinsic to the underlying metaphysics, and it is the paradox that has to be blocked in more or less ad hoc fashion. The part-whole metaphor supports the idea that, at the most fundamental level, there are no different kinds of entity. The idea of a propositional function's being applied to itself to yield a proposition is not one that is obviously ruled out by the basic metaphysics; Russell's paradox thus threatens that metaphysics itself. Paradox is avoided by the theory of types, which is uneasily superimposed on the underlying metaphysics. The theory of types

[49] On a Fregean account, by contrast, a sentence has both a sense (the thought it expresses) and a reference (its truth-value). Frege argues for the distinction between sense and reference for sentences in 'Function and concept'. It is striking, from our point of view, that this argument proceeds by taking for granted the notion of a function.

[50] In particular, nothing guarantees that every natural number has a successor distinct from it.

is based on the idea that a propositional function *presupposes* the propositions which are its values. Russell argues that it follows from this that a propositional function cannot be a constituent of any of its values. Since a proposition presupposes its constituents, if a propositional function were a constituent of one of its values we would have that proposition both presupposing and being presupposed by the propositional function; this he holds to be absurd. The crucial consequence of this is that we cannot apply a propositional function to itself and obtain a proposition. These ideas, however, rely upon a notion of presupposition which is unexplained and which seems, indeed, to be at odds with Russell's object-based metaphysics.[51]

Let us now turn to a rather different issue, still having to do with the logics of Frege and of Russell and with the difference between Fregean functions (and hence also concepts) and Russellian propositional functions. The latter, as we saw, are complex structured entities, whereas Fregean concepts are not. It is tempting to phrase this point about Fregean concepts by saying that concepts true of exactly the same objects are identical. This is misleading, because identity in Frege's view is a *first-level* concept: it applies only to objects, not to concepts. Frege does, however, say explicitly that co-extensiveness is the analogue for functions (including concepts) of the notion of identity.[52] Two predicates which apply to the same objects are thus, on Frege's account, like two names which pick out the same object; nothing in the logic will turn on the difference between such predicates. Frege's logic is thus, in one sense of that word, *extensional* from the outset:[53] his fundamental entities, concepts, have their identity-conditions (or rather the analogue of

[51] One might take this as a partial vindication of Frege's reaction to *Principia Mathematica*: he complains that he does not understand Russell's notation for propositional functions, and the (related) use of the word 'variable'. See his letter to Jourdain, undated draft of a letter sent on 28 January 1914, and the letter dated 28 January 1914, *PMC*, pp. 78–84.

[52] See 'Comments on sense and reference', in *NS*, p. 132, *PW*, p. 122.

[53] In the strictest, and clearest, sense, it is perhaps only *contexts*, not entities or logics, which can be said to be extensional or non-extensional: a context is extensional when replacing an expression in that context with any coreferential expression results in a whole with the same truth-value, or the same reference, as the original. The usage I follow here, however, is a common and natural way of extending the terminology.

identity-conditions) given by the objects of which they hold. His Axiom V partially undoes the concept–object distinction, by asserting that for every concept there is a corresponding object, but it does not impose extensionality, for Frege's concepts are already extensional.[54]

For Russell, however, the situation is quite different. A Russellian proposition is a complex structured entity, in some ways (though not others) more akin to the Fregean sense of a sentence than to its reference. At the most fundamental level, Russell's logic is thus not extensional. Propositional functions, moreover, have the same sort of complexity as propositions: it makes sense to say of a propositional function that it contains a given object or (crucially) that it contains a variable with a given range. This fact about propositional functions, moreover, is not adventitious. On the contrary, this is what makes propositional functions acceptable to Russell, whereas functions *simpliciter* are not. This fact is also what makes it comprehensible that Russell's theory of types is what Ramsey called a *ramified* theory: one in which two propositional functions applicable to entities of the same type may themselves be of different types. When a propositional function contains a quantifier which itself ranges over propositional functions, then on Russell's account it presupposes all those propositional functions. Hence, by the doctrine that lies at the basis of Russell's theory of types, such a propositional function cannot itself be one of those within the range of the quantifier. Hence it must be of higher type.[55]

The mathematical work of *Principia Mathematica* is of course done in extensional terms – it is done in terms of classes, which for Russell, as for everyone else, are extensional entities (in the sense in which we are using that word). Symbols for classes, however, are in that work a mere *façon de parler*, introduced by definitions which enable us to eliminate them (though at the cost of great complexity and prolixity) from any context in which they can legitimately

[54] For this reason, Frege's logic without Axiom V might be thought of as equivalent to what Ramsey described as *Simple Type Theory*; the latter, however, allows for unlimited ascent up the hierarchy of types, whereas it is by no means clear that Frege would have been willing to accept an analogous ascent up the hierarchy of objects, functions of objects, functions of functions of objects, and so on. (For this latter point I am indebted to Warren Goldfarb.)

[55] In this paragraph I am indebted to David Kaplan.

occur. The purpose of the definition is to give us the appearance of extensional entities with which to work, since the reduction to mathematics demands such entities. Russell's definition of classes should thus not be compared with Frege's Axiom V; as a mere definition, it adds no genuine power to the system. (Power is added to the system by the Axiom of Reducibility, which guarantees that for every propositional function there is a coextensive propositional function of the lowest type; this propositional function may, in effect, be treated as the class corresponding to the given propositional function.)

This extensional superstructure, however, is imposed upon a system which in its foundations is intensional through and through: Russellian propositions are not identical when they have the same truth-value, and his propositional functions are not identical when they hold of the same objects. It is this feature of the underlying logic of *Principia Mathematica* which has led some (perhaps most notably Church) to try and exploit it as a logic of such intensional notions as 'believes that'. It would be a mistake, however, to think that these intensional elements arise from any interest on Russell's part in that kind of logic. They arise, rather, from just those fundamental features of his philosophy which we have emphasized in contrasting his view with that of Frege.[56]

[56] Besides the particular debts indicated in other notes, I am indebted to Cora Diamond and Thomas Ricketts for their comments on earlier drafts.

14 Inheriting from Frege: the work of reception, as Wittgenstein did it

1 INTRODUCTION

We might compare Wittgenstein's relation to Frege with Frege's own relation to Kant. Frege's conception of arithmetic developed in great part as a critical response to Kant's, but he wanted it to be quite clear that his criticisms were not those of a petty fault-finding spirit vis-à-vis 'a genius to whom we must all look up with grateful awe' (*Gl*, §89). In Frege's criticisms of Kant one can see his sense that the pursuit of issues raised by Kant must be of the greatest value. He attempted to hold on to Kant's insights, sharpening them when he could, and removing what he took to be extraneous or in tension with Kant's most fruitful ideas. It is precisely that combination of great respect and deeply serious criticism, criticism the seriousness of which is itself expressive of respect, which we find mirrored in Wittgenstein's relation to Frege.[1]

[1] On Wittgenstein's reverence for Frege, see Erich Reck, 'Wittgenstein's "great debt" to Frege: Biographical traces and philosophical themes', in Reck (ed.), *From Frege to Wittgenstein: Perspectives on Early Analytic Philosophy* (New York: Oxford University Press, 2002); P. T. Geach, 'Saying and showing in Frege and Wittgenstein', *Acta Philosophica Fennica*, 28 (1976), pp. 54–70, and Editor's Preface, in *Wittgenstein's Lectures on Philosophical Psychology, 1946–47*, ed. Geach (Chicago: Chicago University Press, 1988), pp. xi–xv; and James Conant, 'On going the bloody *hard* way in philosophy', in J. Whittaker (ed.), *The Possibilities of Sense* (Basingstoke: Palgrave, 2002), pp. 85–129. For a different view of Wittgenstein's relation to Frege, see P. M. S. Hacker, 'Frege and the early Wittgenstein', in Hacker, *Wittgenstein: Connections and Controversies* (Oxford: Clarendon Press, 2001), pp. 191–219, and 'Frege and the later Wittgenstein', in *ibid*., pp. 219–42. On Frege in relation to Kant, see Joan Weiner, *Frege in Perspective* (Ithaca, N.Y.: Cornell University Press. 1990) on Frege's weaknesses as a reader of Kant. See Thomas Ricketts, 'Frege, the *Tractatus*, and the logocentric predicament', *Noûs* 19 (1985), pp. 3–15, at p. 15, for some comments on whether Frege is quite as loyal to Kant on geometry as he takes himself to be. See John MacFarlane, 'Frege, Kant, and the logic in logicism',

Throughout his life, Wittgenstein was enormously influenced by Frege. Frege's writings shaped, to a great extent, the problems Wittgenstein confronted in his own thought – and not just the problems, but also methods of approach, and ideas about what could count as a satisfactory solution. Frege's courage as a philosopher clearly inspired Wittgenstein; his own conception of what philosophy might demand of one reflects his view of *Frege's* response to those demands.

If we compare Wittgenstein's earlier and later writings, we can note changes in the form taken by Frege's influence. Many of the themes of the *Tractatus*, and central conceptions within it, can be seen to have their roots in Frege: an obvious and important example is Wittgenstein's use[2] of Frege's principle that one should not ask for the meaning of a word except in the context of a proposition. After stating the Context Principle, Wittgenstein applies it in his account of what an expression is; and his characterization of expressions as marks of common form and content in the propositions in which they occur is then tied directly to the structure and methodology of the *Tractatus* as a whole. The *Tractatus* treats propositions which have some element of form and content in common as values of a propositional variable, so every expression can be presented by a variable the values of which are the propositions having *that* expression in common. A generalization of that idea underlies a basic organizational principle of the book: any shared logical feature of propositions can be presented by a variable the values of which are the propositions with that feature; so the fundamental variable, which has as its values *all* propositions, will make clear what is common to *all* that can be said, and will in that way make clear the nature of thought. The idea that we can, by working out that fundamental variable, make clear the essence of the expression of thought depends on Wittgenstein's use of Frege's Context Principle.[3]

Philosophical Review, 111 (2002), pp. 25–65 for Frege's conception of logic in relation to Kant's.

[2] See Ludwig Wittgenstein, *Tractatus Logico-Philosophicus*, 3.3. Hereafter *TLP*.

[3] The Context Principle takes on a somewhat different role in Wittgenstein's later philosophy; see, e.g., Ludwig Wittgenstein, *On Certainty*, ed. G. E. M. Anscombe and G. H. von Wright, trans. Denis Paul and G. E. M. Anscombe (Oxford: Blackwell, 1968), §§347–51. For discussion of the Context Principle in Wittgenstein's later philosophy, see Erich Reck, 'Frege's influence on Wittgenstein: Reversing metaphysics via the Context Principle', in W. W. Tait (ed.), *Early Analytic Philosophy*

A less obvious but equally important Fregean element in the *Tractatus* is its idea of philosophical method as including elucidation (*Erlaüterung*); this idea is closely connected with the role in the *Tractatus* of the notion of 'showing' in contrast with 'saying'. Wittgenstein draws on Frege's appreciation of the fact that the sentences used in some kinds of philosophical explanations (for example, explanations of the distinction between concepts and objects) do not say what they seem to be intended to say, and indeed may say nothing coherent; they may nevertheless help us to reach a kind of understanding.[4]

There are also some central points in the *Tractatus* at which Wittgenstein explicitly criticizes Frege, for example, 4.063, 4.431 and 5.02 (in all three of which Wittgenstein is concerned with Frege's treatment of propositions as names), and 5.4 (criticizing Frege's and Russell's account of logical constants). In the writings leading up to the *Tractatus* and in the *Tractatus* itself, we can see Wittgenstein working his way into a deeper understanding of the tensions within Frege's approach and attempting (as Thomas Ricketts and Michael Potter make clear) to resolve those tensions; we can also see him (as Peter Geach suggests) shaping his own understanding of logic and language through a rejection of one central Fregean idea (the idea

(Chicago: Open Court, 1997), pp. 123–85. Reck emphasizes that the Context Principle does not, for Frege, replace the idea of a word as having meaning in isolation with the idea that a word has meaning in a sentence, the sentence itself being then considered in isolation. That point, that the relevant context is not a sentence taken in isolation, applies also to Wittgenstein's earlier and later uses of the principle, but in different ways. For further discussion of these issues (and further bibliographical suggestions), see Martin Gustafsson, *Entangled Sense: An Inquiry into the Philosophical Significance of Meaning and Rules* (Uppsala: Universitetstryckeriet, 2000); also James Conant, 'Wittgenstein on meaning and use', *Philosophical Investigations*, 21 (1998), pp. 222–50.

4 See Joan Weiner, 'Theory and elucidation: The end of the age of innocence', in Juliet Floyd and Sanford Shieh (eds.), *Future Pasts: The Analytic Tradition in Twentieth-Century Philosophy* (New York: Oxford University Press, 2001), pp. 43–65; and James Conant, 'The method of the *Tractatus*', in Reck, *From Frege to Wittgenstein*, pp. 374–462. Conant notes that Max Black's commentary on the *Tractatus* has contributed to the failure to see Wittgenstein's indebtedness to Frege on the issue of explanations which turn out to be nonsense but which can nevertheless be helpful. Black greatly distorts Frege's views in attempting to make out a contrast between Wittgenstein and Frege on this issue. On Wittgenstein's indebtedness to Frege concerning these points, see Geach, 'Saying and showing', and 'A philosophical autobiography', in H. A. Lewis (ed.), *Peter Geach: Philosophical Encounters* (Dordrecht: Kluwer, 1991), pp. 1–25.

of sentences as proper names), while keeping hold of Frege's other insights.[5]

Within Wittgenstein's later philosophy, the influence of Frege is perhaps most easily seen in Wittgenstein's discussions of the philosophy of mathematics. Again and again in his lectures and writings on philosophy of mathematics, he is responding to passages in Frege which he must have known virtually by heart, and which he took to express in a particularly clear way philosophical conceptions of great importance. See, for example, the many references in *Lectures on the Foundations of Mathematics* (1976) and in *Remarks on the Foundations of Mathematics* (1978) to Frege's discussions (in the Preface to *Grundgesetze*) of the objectivity of logic.[6] While Wittgenstein's explicit discussions of Frege on the objectivity of logic are in the context of lectures or writings on the foundations of mathematics, the issue is one of vital significance for all of Wittgenstein's later thinking. This is because of the importance,

[5] See Ricketts, 'Frege, the Tractatus and the logocentric predicament'; Michael Potter, *Wittgenstein's Notes on Logic* (Oxford: Oxford University Press, 2009); and Geach, 'Saying and showing'. Potter emphasizes the significance of Frege's writings and of conversation with Wittgenstein for the development of Wittgenstein's ideas from 1912 onwards. For a discussion of Geach's account of the relation between Frege's thought and the *Tractatus*, see Cora Diamond, 'Throwing away the ladder', *Philosophy*, 63 (1988), pp. 5–27, reprinted in Diamond, *The Realistic Spirit: Wittgenstein, Philosophy, and the Mind* (Cambridge, Mass.: MIT Press, 1991), pp. 179–204. Warren Goldfarb has raised questions about the claims made by Geach, Ricketts and me about Frege's influence on the *Tractatus*; see his 'Wittgenstein's understanding of Frege: the pre-Tractarian evidence', in Reck, *From Frege to Wittgenstein*, pp. 185–200. On Goldfarb's questions, see also the last paragraph of §1 of this chapter and the Appendix; also Thomas Ricketts, 'Wittgenstein against Frege and Russell', in Reck, *From Frege to Wittgenstein*, pp. 227–51; and Potter, *Wittgenstein's Notes*. For a discussion of Wittgenstein's criticisms of Frege at *TLP*, 4.442, 4.063 and 4.064, and the corresponding sections of the *Notes on Logic*, see Potter, *Wittgenstein's Notes*, and Ian Proops, 'The early Wittgenstein on logical assertion', *Philosophical Topics*, 25 (1997), pp. 121–44; I discuss some related points in Cora Diamond, 'Truth before Tarski: after Sluga, after Ricketts, after Geach, after Goldfarb, Hylton, Floyd and Van Heijenoort', in Reck, *From Frege to Wittgenstein*, pp. 252–79.

[6] Ludwig Wittgenstein, *Lectures on the Foundations of Mathematics, Cambridge, 1939*, ed. Cora Diamond (Ithaca: Cornell University Press, 1976), pp. 198, 201–2, 214, 230–1; and *Remarks on the Foundations of Mathematics*, ed. G. H. von Wright *et al.*, trans. G. E. M. Anscombe (Oxford: Blackwell, 1978), pp. 89, 95, 234, 241. Another passage to which Wittgenstein referred repeatedly is that in which Frege says (in the context of a discussion of Schubert on number) that the geometrical line connecting two points is there before we draw it. See *CP*, p. 264.

in Wittgenstein's later method, of imagining forms of activity and thought different from ours: what would go wrong if we were to have a different logic or mathematics – or were to engage in language games departing greatly from our actual ones?

In the rest of this essay I consider three elements in the complex relation between Wittgenstein's thought and Frege's. In the Appendix I discuss a further element in that complex relation and some questions, recently raised by Warren Goldfarb, about Frege's influence on Wittgenstein's earlier thought (see note 5). Some of Goldfarb's questions and criticisms concern things I said in an early version of this essay. In order that the issues here may be clear, I have made no substantial changes in those passages in this essay about which Goldfarb has questions.

2 ON DISTINGUISHING SHARPLY ENOUGH

In 'What is a function?' Frege criticized the then current understanding of the word 'function' in Analysis. What the word stands for had been explained by appeal to the idea of a variable magnitude or to the notion of supposed indefinite or variable numbers designated by variables – as if, in addition to 3 and π, which are constants, there were *variable* numbers; as if numbers could be divided into definite ones and indefinite ones.[7] We are now so accustomed to the Fregean conception of a function and to the idea that the letter 'x' which we write in the argument-place does not designate an 'indefinite number' or anything else, that we may fail to see the depth of insight in Frege's conceptual moves. There is, though, an interesting passage in one of his discussions of formalism in geometry which makes clear the nature of the insight. I shall quote it at length; it concerns the special case of expressions for concepts but comes from a general criticism of common confusions in our thought about functions and concepts.

Concept-words offer [an] occasion where it may seem that ambiguous signs are necessary. If we think that the word 'planet' designates at one time the Earth, at another Jupiter, then we should take it to be ambiguous. But … it does not stand to the Earth in the relation of sign to thing signified. Rather, it designates a concept, and the Earth falls under it. No ambiguity is to be

[7] *CP*, pp. 285–92.

found here. Let us suppose that the word 'planet' is unknown and that we wanted to designate the appropriate concept. We might then perhaps hit upon the idea of using the proper name 'Mars' for it, and might find unreasonable the demand that the word 'Mars' be given a determinate meaning: as wide a range of interpretations as possible ought to be kept open for this name. But as a concept-word, 'Mars' would have to be just as unambiguous as it would have been as a proper name. Do not say that as a concept-word it has no determinate meaning, or that it refers to an indeterminate object. Every object is determinate; 'indeterminate object' is contradictory, and wherever this expression occurs, we can be quite certain that a concept is what is really meant. We cannot say that the proposition '$x > 0$' assigns an indeterminate object, an indeterminate number, to the letter 'x' as its meaning. Rather, what is designated here is a concept: *positive number*; nor is 'x' introduced as a sign for this concept; it merely takes the place of the proper names (number-signs) of objects that may perhaps be subsumed under the concept. Thus the appearance of ambiguity arises only out of an insufficient understanding, in that proper names and concept-words are not distinguished sharply enough.[8]

Frege says: 'proper names and concept-words are not distinguished *sharply enough*'. We treat the letter indefinitely indicating an object which may fall under the concept as if it were a proper name, but of a *peculiar kind* of object. We make the difference *too slight* – and that is at the root of the philosophical confusion. Here now is Wittgenstein, in *Philosophical Investigations*:

Thinking is not an incorporeal process which lends life and sense to speaking, and which it would be possible to detach from speaking, rather as the Devil took the shadow of Schlemiehl from the ground. – But how 'not an incorporeal process'? Am I acquainted with incorporeal processes, then, only thinking is not one of them? No; I called the expression 'an incorporeal process' to my aid in my embarrassment when I was trying to explain the meaning of the word 'thinking' in a primitive way.

One might say 'Thinking is an incorporeal process', however, if one were using this to distinguish the grammar of the word 'think' from that of, say, the word 'eat'. Only that makes the difference between the meanings look *too slight*. (It is like saying: numerals are actual, and numbers non-actual, objects.) An unsuitable type of expression is a sure means of remaining in a state of confusion. It as it were bars the way out.[9]

[8] *CP*, p. 307.

[9] Ludwig Wittgenstein, *Philosophical Investigations*, ed. G. E. M. Anscombe and R. Rhees, trans. G. E. M. Anscombe (Oxford: Blackwell, 1958), §339.

Wittgenstein's diagnosis of what is wrong with the appeal to the idea of 'incorporeal processes' in our explanation of the meaning of 'thinking' runs parallel to Frege's diagnosis of the appeal to the idea of 'indefinite objects' in our explanations of the meaning of the letters we use as variables. We call to our aid, in our embarrassment (*Verlegenheit*, which could here also be translated as 'perplexity') an expression which is formed in such a way as to ensure that we remain entrapped by words: we imagine a classification *within* objects, of definite and indefinite ones, or *within* processes, of corporeal and incorporeal ones.

An irony of the passage from *Philosophical Investigations* is that it applies the Fregean kind of diagnosis to Frege's own distinction between actual (*wirklich*) and non-actual objects, and to his view of numbers as non-actual and numerals as actual objects. (For numbers as non-actual objects, see, e.g., *Gl.* §§26–7.) The application of the Fregean diagnosis to Frege's own ideas can also be seen earlier, in the *Tractatus*. Wittgenstein insists against Frege (and Russell too) that logic is not a science distinguished from other sciences by its generality; he rejects any treatment of the logical connectives which treats them as a kind of function-sign in either Frege's or Russell's sense; and he rejects the idea that sentences are proper names distinguished from other proper names by what they name. In all these cases we are unable to reach a clear understanding because we fail to make the difference between the cases deep enough. We attempt to treat a significantly different case on a familiar model, taken with modest modifications. The *Tractatus* and the stages of its development provide us with examples of Wittgenstein repeatedly seeking to break out of the hold on our thought of differences conceived not deeply enough.[10] Perhaps the most important of these cases in Wittgenstein's early thought is that of the distinction between using signs to express what we

[10] For discussion of a pre-*Tractatus* use of this sort of argument, applied to Russell, see Ricketts's account of Wittgenstein's distinction between forms and names. As Ricketts makes clear, Wittgenstein's critique of Russell is that he had failed to make the difference between objects and relations deep enough, treating relations as essentially a species of objects. Thomas Ricketts, 'Pictures, logic and the limits of sense in Wittgenstein's *Tractatus*', in Hans Sluga and David G. Stern (eds.), *The Cambridge Companion to Wittgenstein* (Cambridge: Cambridge University Press, 1996), p. 72.

wish to say is so and the expressiveness internal to signs as used to say anything whatever.[11]

The idea that philosophical confusion may stem from making a difference too slight goes with two further important ideas: first, of a contrast between kinds of differences, and second, of the need for a technique by which we will be able to reveal the character of the differences which we misconceive by taking them to be 'too slight'.

Although both Frege and Wittgenstein speak of the confusions with which they are concerned in terms of degree – the slightness of a difference, the sharpness of a distinction – their idea is not that we have, as it were, taken a difference like that between apples and potatoes and reduced it to a difference merely between two sorts of apples. Potatoes and apples are distinguished by their properties: apples have properties potatoes lack, and vice versa; but the sharp differences which Frege and Wittgenstein want us to note are not a matter of contrasting properties. It is precisely the attempt to represent the differences *as* a contrast in properties that they reject: the idea that the difference between n and π is a matter of n being an indefinite and π a definite number (*definiteness* being a property which only some numbers have), and the idea that the difference between thinking and digesting is in the property of *corporeality*, which only some processes have. Differences in logical kind are not differences in properties.[12] For both Frege and Wittgenstein the difference in logical kind *between* differences in logical kind and apple–potato differences is tied to a difference in the way in which different sorts of differences can be put before us. In this connection Frege's invention of his concept-script (and, in particular, of the quantifier-variable notation) had a twofold importance: first, it made clear the kind of difference there is between function and object and between concept and object (thus enabling us to avoid the

[11] See Michael Kremer's discussion of failure to make the difference between 'showing' and 'saying' deep enough, in his 'The purpose of Tractarian nonsense', *Noûs,* 35 (2001), pp. 39–73; see also Kremer, 'The cardinal problem of philosophy', in Alice Crary (ed.), *Wittgenstein and the Moral Life: Essays in Honor of Cora Diamond* (Cambridge, Mass.: MIT Press, 2007), pp. 143–76, and Ricketts, 'Pictures, logic', pp. 92–4.

[12] See Ludwig Wittgenstein, *The Blue and Brown Books* (Oxford: Blackwell, 1964), p. 19.

sort of confusion exemplified by the idea of variables as standing for indefinite numbers), and, second, it provided a model of *how* differences in logical kind could be made perspicuous. The influence on Wittgenstein's thought is of great importance. In the *Tractatus* he accepted the idea that differences in logical kind can be made clear in a good notation. The distinction, central in the *Tractatus*, between what can be said and what shows itself in language, builds on Frege's understanding of the difficulties in explaining the distinction between functions and objects.[13] And a central methodological issue for Wittgenstein's later thought is this: if, as he came to believe, differences in logical kind could not in general be made perspicuous by a conceptual notation, how could philosophy make them clear and thus enable us to avoid the kind of confusion arising from 'making the difference appear too slight'? Here we have a good example of how a central problem for Wittgenstein's thought can be seen to be shaped by ideas which were originally Frege's.

Again and again in his later writings, Wittgenstein turns to issues pivoting on our making differences appear too slight. We should, I am suggesting, see the impress of Frege on such very late remarks as these:

> 'Mental' for me is not a metaphysical, but a logical epithet …
>
> There are inner and outer concepts, inner and outer ways of considering human beings. Indeed there are also inner and outer facts – just as there are for example physical and mathematical facts. But they do not stand to each other like plants of different species. For what I have said sounds like someone saying: In nature there are all of these facts. Now what's wrong with that?[14]

Wittgenstein tries to show us the roots of metaphysical confusion, and of metaphysics itself, in our taking the 'logical nature' of

[13] It also builds on Frege's understanding of different sorts of variables. See Thomas Ricketts, 'Generality, meaning and sense in Frege', *Pacific Philosophical Quarterly*, 67 (1986), especially the discussion, p. 180, of concept and object not being species of any genus. That they are not species of a genus goes with there not being any variable indicating both concepts and objects; see also Geach, 'Saying and showing' and 'A philosophical autobiography'; Weiner, 'Theory and elucidation'; and Conant, 'The Method of the *Tractatus*'.

[14] Ludwig Wittgenstein, *Last Writings on the Philosophy of Psychology*, vol. II, ed. G. H. von Wright and H. Nyman, trans. C. G. Luckhardt and M. A. E. Aue (Oxford: Blackwell 1992), p. 63.

things as a matter of 'metaphysical species', thought of on the analogy with natural species. A similar argument underlies his repeated criticisms of the exploration of the mathematical realm, thought of on the analogy of exploration discovering new species of plants and animals. Although he is here very far from Frege's views about mathematics, his approach is nevertheless profoundly dependent on what he learned from Frege about the importance of not making differences too slight.[15]

3 ON OVERCOMING BY TAKING SERIOUSLY

Ernst machen mit der formalen Arithmetik, das ist sie überwinden… (*Gg*. §137)

(*To take formal arithmetic seriously is to overthrow formal arithmetic.*)

Frege's philosophical writings have some characteristic features of which we are usually aware but which we may not connect with his importance as a thinker. We can hardly help noting how frequently he responds at length and with ridicule to the confused views he wants to combat. Attention to these features of his writings, and to the ideas about philosophical confusion with which they are connected, will let us see an important kind of influence which Frege had on Wittgenstein.

I start by putting the quotation at the beginning of this section into its context. It comes from Frege's discussion, in volume II of *Grundgesetze*, of formalism in arithmetic. His *Überwindung*[16] of formal arithmetic ends this way:

[15] For Wittgenstein on our making the difference between physical and logical nature too slight (and correspondingly making the difference between physical and logical impossibility too slight) see G. E. M. Anscombe's account in 'The reality of the past', in Anscombe, *Metaphysics and the Philosophy of Mind* (Oxford: Blackwell, 1981), pp. 103–19, especially p. 114.

[16] I have used the German word because '*Überwindung*' and '*überwinden*' have a significant place in subsequent philosophical writing. I want, in particular, to make a connection with *TLP*, 6.54. The *Überwindung* of Wittgenstein's own sentences in the *Tractatus* is achieved by treating them with full seriousness; they then reveal themselves to be nonsense. I am grateful to Richard Rorty for drawing my attention to the subsequent use of '*Überwindung*' by Carnap, in connection with the overcoming of metaphysics, and by Heidegger.

Formal arithmetic can remain alive only by being untrue to itself.* Its semblance of life is facilitated by the haste with which mathematicians usually hurry over the foundations of their science (if indeed they have any concern for them), in order to reach more important matters. Many things are omitted completely, others briefly touched on, nothing performed in detail. Thus a theory may appear secure which would immediately reveal its weaknesses upon any serious attempt at consistent elaboration. This shows the road to refutation. We need only follow the lines of thought further, to see where they lead. To take formal arithmetic seriously is to overthrow formal arithmetic; and that is what we have done.

[*A fancier of paradox might say: the correct interpretation of the formal theory consists in interpreting it incorrectly.][17]

There are two extremely interesting and closely related ideas in the quoted passage: (1) formal arithmetic remains alive only by being untrue to itself (*sie sich selbst untreu wird*); (2) the way to criticize it is to take it seriously, to provide for it the fidelity to itself which it lacks. Here we see an important and original conception of philosophical criticism. Frege thinks that there is a kind of evasiveness, a failure of integrity, a failure in thinking through their own view, in the formalists. In response to this evasiveness, the critic must take what he is criticizing *with greater seriousness than it takes itself.* If you say that numbers are tangible signs, or that arithmetic is simply a game with signs which have no content beyond what is provided by the rules for manipulating the signs in the game, then, *if you are serious*, this and this and this is what you are saying.

Frege's treatments of philosophers with whom he disagrees are not always accompanied by remarks about evasiveness. But there is in his conception of what it is to think about logic and mathematics a demand that we take what we say seriously; what he (frequently) demonstrates by his criticisms is that those whom he is criticizing have not met that demand. And he does this by taking the thinkers whom he is criticizing more seriously than they take themselves. He acts, as it were, as a mirror, in which what has been hidden from a thinker by his own evasions is open to view.

[17] *Gg*, vol. II, §137. Frege has a second, much longer footnote to the passage, which I have omitted. It is concerned with the writings of H. v. Helmholtz; among its criticisms of Helmholtz is his utter failure to distinguish things which are totally different.

Consider Frege's criticism in *Grundlagen*, §27, of Schloemilch's view that number is the idea of the position of an item in a series. But if the number two were an idea, each of us would have his own: there would not be just one number two, but mine and yours and everybody else's – we should perhaps have many millions of twos. There might be conscious and unconscious twos. Worse yet, 'as new generations of children grew up, new generations of twos would continually be being born, and in the course of millenia these might evolve, for all we could tell, to such a pitch that two of them would make five'. These are the wonders to which we are led if we take seriously the suggestion that the number two is an idea.

One might read Frege in such passages merely as criticizing theories which he takes to be false by producing elaborate arguments of a reductio type. He certainly *is* doing that, but I am drawing attention to something else, which I think was important for Wittgenstein, namely the conviction, underlying Frege's method, that those who put forward these false theories are attracted to certain conceptions, but the attraction is superficial. They can be shown that they do not want what they think they want by being shown what it is. The attraction of these views for us is not separable from a failure in us to mean what we say *all the way*. So our failure in philosophical thought about logic or mathematics can be exposed by letting us see what it really would be to mean what we say.

Frege's ideas about evasiveness in thought about logic and mathematics are more explicit in *Grundgesetze* than in *Grundlagen*. The Introduction to *Grundgesetze* is quite interesting in this regard. It is usually read as exhibiting very clearly Frege's own conception of logic and of objectivity, and his rejection of psychologism. But it also presents us with a very interesting phenomenology of logical confusion. The confusion of the psychological logician is particularly evident in the 'vain struggle' in which he is engaged; he attempts somehow to keep hold of the notions characteristic of a healthy understanding and to reconcile them, or to persuade himself that he has reconciled them, with an account which is, though he does not recognize this, deeply at odds with such an understanding. The vain struggle is in evidence in the psychological logician's equivocations. He uses the word 'idea' sometimes to mean something within the mental life of individuals (this is his official view); but he shifts over to a use of 'idea' to mean something set over apart from everyone in the same way, not something psychological at all. That use of the same word for two entirely different conceptions enables the psychological logician to obscure the issue and to conceal (Frege means: conceal from himself, in the first instance) the weakness of his theory. (See *Gg*, vol. I, pp. xviii–xix; compare also *Gg*, vol. II, §137, on the equivocation of which mathematicians who put forward formalist theories are guilty, while unaware of it.) Frege draws attention to another good example of this kind of self-deception in the psychological logician's theory of perception. According to that theory, both the tower which we see through a window and the window itself are retinal images. The retinal image of the tower is of course smaller than that of the window. Our common understanding includes recognition that the tower is actually larger than the window; it follows that one must either deny the common understanding or deny that the tower and the window are retinal images. In his vain struggle to accommodate the common understanding, the psychological logician says: the retinal image of the tower, *as such*, is indeed not bigger than that of the window. Here Frege's patience gives way:

> At this I almost feel like losing my temper entirely and shouting at him: 'Well then, the retinal picture of the tower is not bigger than the retinal picture of the window at all, and if the tower were the retinal picture of the tower and the window were the retinal picture of the window, then the tower would not be bigger than the window either, and if your logic teaches you differently it is absolutely worthless!' This 'as such' is a splendid discovery for hazy writers reluctant to say either yes or no.[18]

Frege is a great logician, a great thinker; because we find so interesting and important his own substantial views, we may pass by without noticing it the role in his critical thinking of a moralized, or quasi-moralized, conception of thought about logic, and its dependence on meaning something *all the way*, without dependence on equivocation and evasion.[19]

[18] *Gg*, vol. I, p. xxiii.

[19] On the moralized character of Frege's conception of what is involved in thought about logic, see also his claim (*Gg*, vol. I, p. xxv) that what stands in the way of recognition of their confusion by psychological logicians is that they '*sich auf die psychologische Vertiefung Wunder was zu Gute thun*'. Furth's translation – 'take such fantastic pride in psychological profundity' (*Basic Laws of Arithmetic*

Wittgenstein's tone, in his later philosophy, is very different; he does not pick up Frege's impatience or his sarcasm. Philosophical confusion cannot be cured by ridicule. What he does pick up is Frege's awareness of the tension within the philosophically confused view. The attractiveness of the view depends on our not meaning it all the way. Hence the right philosophical approach to the confused view is one which will reveal it for what it is. In 1931 Wittgenstein said, 'I ought to be no more than a mirror, in which my reader can see his own thinking with all its deformities so that, helped in this way, he can put it right'; much later he said that to express a false thought boldly and clearly is already an achievement: 'It's only by thinking even more crazily than philosophers do that you can solve their problems'.[20] Wittgenstein's conception of the importance of courage in philosophical thinking is tied also to Frege's; they share the view that we can reach the truth only by having the courage not to fall back on to some comforting apparent truth lying near to hand.

Wittgenstein spoke of our having *an urge to misunderstand* the workings of our language. His philosophical style was informed by resistance to that 'urge to misunderstand'. Here there is a connection with Frege's style: for them both, resistance to the urge towards misunderstanding is not just a motive but something which gives a characteristic urgency and shape to their prose. Much of *Philosophical Investigations* is concerned with our tendency to take understanding, meaning and being guided by a rule (as in reading) to be specific mental processes. Resistance to the urge towards that group of misunderstandings of the workings of language must allow the urge its fullest expression, must respond, successively, to each of the forms which the misunderstanding takes when some earlier form has been revealed to be nothing we really could want to say. The tension of Wittgenstein's prose is the tension of giving the urge to misunderstand the full expression which it needs,

(Berkeley: University of California Press, 1964), p. 25) – makes explicit the moral character of Frege's judgement.

[20] Ludwig Wittgenstein, *Culture and Value*, ed. G. H. von Wright, trans. Peter Winch (Oxford: Blackwell, 1980), pp. 18, 75. For Wittgenstein on the philosopher as mirror, see James Conant, 'Putting two and two together: Kierkegaard, Wittgenstein and the point of view for their work as authors', in T. Tessin and M. von der Ruhr (eds.), *Philosophy and the Grammar of Religions Belief* (Basingstoke: Macmillan, 1995), pp. 248–331.

and of allowing that full expression to lead ultimately to recognition of what it was we had taken ourselves to want to say. This is what Frege had done, examining successively in *Grundlagen* the misunderstandings to which we are drawn when we ask ourselves what numbers are, and allowing those misunderstandings their full expression. If numbers are ideas, then really and truly my one-times-one might be one, while your one-times-one might be two. What was always *disguised* nonsense is made to show itself to be nonsense;[21] its attractiveness is thus made to disappear. Frege's ridicule and Wittgenstein's patient unravellings are not as far apart as they appear. Argument, for each of them, has (among its roles) the complex role of serving to get round our strategies of evasion; the prose of both writers has a complex kind of intellectual rhythm, as it allows the misunderstandings to take a multiplicity of forms, each of which then needs its own response.[22]

In §2, I argued that Frege's influence can be seen in the importance Wittgenstein attaches to not making a difference too slight; in §3, I have tried to show how the roots of Wittgenstein's conception of philosophy as responding to confusion (confusion which depends on a kind of failure to mean fully what we say) can be found in Frege. These two sorts of influence on Wittgenstein are not separable. Perhaps the very clearest example of Frege's insistence on the importance of not underestimating the sharpness of a distinction can be found in his prolonged ridicule of H. Schubert's account of numbers.[23] Schubert's espousing a view which he cannot mean, a

[21] See Wittgenstein, *Philosophical Investigations*, §464.

[22] Wittgenstein himself had noted the influence of Frege's style on his own style. In *Zettel*, ed. G. E. M. Anscombe and G. H. von Wright, trans. G. E. M. Anscombe (Oxford: Blackwell, 1967), §712, there is this remark, originally from the early 1930s: '(The style of my sentences is extraordinarily strongly influenced by Frege. And if I wanted to, I could establish this influence where at first sight no one would see it.)' I have not tried to establish the influence of Frege's style on specific passages, which would take great attention to detail. An example worth considering might be the long passage from Frege quoted in §2; the intellectual movement of the middle section might be felt as 'Wittgensteinian'. The argumentative structure pivots on a use of 'Do not say', responding to the inclinations characteristic of the confusion which Frege is concerned to combat. Compare the structure of Wittgenstein, *Philosophical Investigations*, §66. See also Reck, 'Wittgenstein's "great debt"', pp. 24–5.

[23] 'On Mr Schubert's numbers', *CP*, pp. 249–72.

view which cannot be meant, not seriously, is entirely dependent, Frege argues, on his overlooking or blurring every significant distinction, on his allegiance to the great principle of *never* distinguishing between what is different.

An important difference between Frege and Wittgenstein, and another similarity as well, come out if we consider this remark of Wittgenstein's:

> One cannot guess how a word functions. One has to *look at* its use and learn from that.
>
> But the difficulty is to remove the prejudice which stands in the way of doing this. It is not a *stupid* prejudice.[24]

Frege did, I think, believe that the conceptions which lead us to make a difference too slight were stupid prejudices (e.g., in favour of the actual and against the non-actual) and infatuation with our own theories.[25] But he did state clearly a vital point: when we are in the grip of such prejudices, we cannot see the differences to which we need to attend because we look in the wrong place. If we insist on looking only at observable objects, we *cannot* see the difference between numbers and numerals; if we insist on looking only at psychological processes and ideas, we *cannot* see the distinction between concept and object.[26] 'Looking in the wrong place', like 'making the difference too slight' is an important term of philosophical criticism which Wittgenstein got from Frege and made his own.

[24] Wittgenstein, *Philosophical Investigations*, §340.

[25] *Gg.* vol. I, p. xxv; see above, note 17. For Wittgenstein's view of the relation between stupidity and philosophy, see also Wittgenstein, *Culture and Value*, p. 39: 'Our greatest stupidities may be very wise.' On how something stupid can nevertheless, even through its stupidity, teach something which we need to grasp, see Wittgenstein on the stupidity of American films, *ibid.* p. 57.

[26] See *Gg.* vol. I, p. xxv. Frege also thought that our blurring of such differences as that between concept and object is due in part to the way in which language itself obscures them. So he believed that the concept-script which he had invented could serve as a means to intellectual liberation by laying clearly before us the differences hidden by language which is inattentive to the promptings of logic. For some further discussion of the influence of this idea on Wittgenstein, see Cora Diamond, 'What does a concept-script do?', *Philosophical Quarterly,* 34 (1984), pp. 343–68, reprinted in Diamond, *The Realistic Spirit*, pp. 115–44.

4 ON SENSE AND REFERENCE[27]

Frege's response to Kant is complex. Earlier I quoted an expression of his great admiration for Kant. He picks out, as among Kant's most important contributions, his distinction between analytic and synthetic judgements, and his characterization of the truths of geometry as synthetic a priori. But Frege's own disagreement with Kant about the nature of arithmetic, and his own new understanding of logic, led him to rethink the very distinction between analytic and synthetic which he took to be such a great contribution of Kant's. In §§4–7 I shall consider Wittgenstein's rethinking, in the *Tractatus*, of the distinction between sense and reference. That distinction is part of Wittgenstein's inheritance from Frege. But Wittgenstein transformed the distinction, just as Frege had transformed the distinction, inherited from Kant, between analytic and synthetic.[28] Here we need a warning in a Fregean style: we are not entitled to speak of 'the distinction between analytic and synthetic' or 'the distinction between sense and reference' if these phrases do not have unambiguous reference. The character of the inheritance is in both cases problematic if what is received is, in the hands of the inheritor, something quite different from what it was in the hands of its previous owner. In the case of 'the distinction between sense and reference' the fact that the phrase has no definite reference is particularly important, since Wittgenstein was not the only inheritor. Post-Fregean philosophers of language and semantic theorists also inherited 'the distinction' and they too have reshaped it in the process. So 'the distinction' as we find it in Wittgenstein and as we find it in neo-Fregean semantic theorizing are even further apart than are Frege's distinction and Wittgenstein's. While I shall in §§4–6 be concerned largely with the relation between Frege and Wittgenstein,

[27] I have generally used 'refer' and 'reference' for '*bedeuten*' and '*Bedeutung*', and have altered some quoted matter to make the translations uniform.

[28] Wittgenstein also transformed the distinction between analytic and synthetic, inherited from Kant and Frege. See Burton Dreben and Juliet Floyd, 'Tautology: How not to use a word', *Synthese,* 87 (1991), pp. 23–49. On Frege's reshaping of the analytic–synthetic distinction, see Weiner, *Frege in Perspective*, ch. 2. On Wittgenstein and Frege on the sense–reference distinction, see also Peter Hylton, 'Functions, operations, and sense in Wittgenstein's Tractatus', in Tait, *Early Analytic Philosophy*, pp. 91–105, reprinted in Hylton, *Propositions, Functions, and Analysis: Selected Essays on Russell's Philosophy* (Oxford: Clarendon Press, 2005), pp. 138–52.

I shall return in §7 to the question of the distance of the various inheritors from their legator.

It is essential to the distinction between sense and reference, as Frege explains and employs it, that the expressions which he calls names (including syntactically simple proper names, descriptive phrases, predicates, relational expressions, other sorts of functional expressions and sentences, but not variables or words like 'whoever', or the assertion sign) are capable of having (and properly should have) both sense and reference. We can thus speak of the 'level' of sense, the 'level' of reference; and it is part of Frege's distinction that an expression cannot have a reference unless it has a sense, since the sense is the way in which the reference is presented. What then is left of the distinction if that is gone – as it is in the *Tractatus*? Here are two of the crucial passages:

3.144 Situations can be described, not *named*.
(Names are like points; propositions like arrows – they have sense.)
3.3 Only propositions have sense; only in the context of a proposition does a name have reference.[29]

The quoted passages exclude the use of 'sense' for any kind of sign other than a sentence; they do not, as they stand, exclude talk of the reference of sentences; and in fact Wittgenstein does at some points (e.g., *TLP*, 5.451) use the word '*Bedeutung*' in an ordinary non-technical sense, so that any sign which has a determinate linguistic function can be said to have *Bedeutung*. He does, though, reject Frege's talk of the reference of sentences, and not just because he drops Frege's treatment of sentences as proper names.[30] (His reasons

[29] Cf. also Friedrich Waismann's bald statement in *Theses*, intended to be a summary of the main ideas of the *Tractatus*: 'A proposition has a *sense*, a word has a *reference*' (Friedrich Waismann, 'Theses' in B. F. McGuinness (ed.), *Ludwig Wittgenstein and the Vienna Circle*, trans. J. Schulte and B. F. McGuinness (Oxford: Blackwell, 1979), pp. 233–61, at p. 237.

[30] See, however, the 'Notes on logic', where Wittgenstein identifies the meaning (i.e., the *Bedeutung*) of a sentence and of its negation as the single fact which corresponds to both (Ludwig Wittgenstein, 'Notes on logic', in *Notebooks, 1914–1916*, eds. G. H. von Wright and G. E. M. Anscombe, trans. G. E. M. Anscombe (Oxford: Blackwell, 1961), pp. 93–4. This view is not present in the *Tractatus*; but something quite close to it is present. See my discussion in §6 of the reality with which a sentence and its negation are to be compared; the (least inclusive) reality to which a sentence is to be compared depends on the elementary sentences

for rejecting Frege's idea of the truth-value of a sentence as its reference would also be reasons for rejecting modifications of Frege's theory – which share his own rejection of the idea of sentences as proper names – in which sentences figure in a semantic theory through their semantic value, in something like the way in which names do.)

I turn now to a discussion of the philosophical issues involved in Wittgenstein's rethinking of the distinction between sense and reference, and to the question how it is connected with the complex range of disagreements with Frege which surface in the *Tractatus*. The importance of these interconnected issues was shown by G. E. M. Anscombe;[31] anyone discussing them now is in debt to that original account. As is clear from her discussion, a complete account of the role of Wittgenstein's critical response to Frege in the *Tractatus* would involve full and detailed exegesis of much of the *Tractatus*, which I cannot attempt here. I shall focus on *TLP* 5.02, which I therefore quote in its entirety:

The arguments of functions are readily confused with the indexes of names. For both arguments and indexes enable me to recognize the reference of the signs containing them.

For example, when Russell writes '$+_c$', the '$_c$' is an index which indicates that the sign as a whole is the addition sign for cardinal numbers. But the use of this sign is the result of arbitrary convention and it would be quite possible to choose a simple sign instead of '$+_c$'; in '~*p*', however, '*p*' is not an index but an argument: the sense of '~*p*' *cannot* be understood unless the sense of '*p*' has been understood already. (In the name Julius Caesar 'Julius' is an index. An index is always part of a description of the object to whose name we attach it: e.g. *the* Caesar of the Julian gens.)

If I am not mistaken, Frege's theory about the meaning of propositions and functions is based on the confusion between an argument and an index. Frege regarded the propositions of logic as names, and their arguments as the indexes of those names.[32]

of which it is composed. A reality is the existence and non-existence of states of affairs, i.e., something fact-like.

[31] G. E. M. Anscombe, *An Introduction to Wittgenstein's Tractatus*, 2nd edn (London: Hutchinson University Library, 1959).

[32] I have substituted 'index' for the Pears-McGuinness 'affix' as a translation of German *Index*. This use follows Wittgenstein's own practice in English. He gives an account of the index/argument contrast in *The Blue Book* (see *The Blue and Brown Books*, p. 21).

Here are three problems which the quoted passage presents.

(1) If we begin with Wittgenstein's contrast between argument and index, it seems obvious that '*p*' occurs in '~*p*', as Frege explains the latter, as an argument, not as an index. (In Frege's language, '*p*' will be the name of an argument; this difference in language does not affect the present issue. Where it is possible to do so without confusion, I shall follow Wittgenstein's practice rather than Frege's; see also note 74 below.)

Consider typical index constructions, like 'Forsythia', 'Dahlia', 'Abelia', or 'Hansen's disease', 'Parkinson's disease', 'Bright's disease'. In each case a name forms part of another name.[33] The inner name or index is, as Wittgenstein notes, part of the description of the thing named by the name of which it is part: Hansen's disease is *the* disease described by Hansen. But the '(·)'s disease' construction does not have the logical complexity of a functional expression. Even if Hansen were also known as Jensen, 'Jensen's disease' would not be another name for Hansen's disease. (Nor would the name lose its unambiguous reference if, some time after Hansen's disease had been named after him, Hansen had gone on to provide the description of some other disease.) No *independent* meaning has been fixed for '(·)'s disease'; it is rather that a certain number of names are formed on a particular pattern, and that it is therefore possible in many cases to work out the reference of the complex names if one knows the relevant facts about the thing whose name forms the index, as one might, e.g., identify 'Abertawe' as a name for Swansea from the presence of the index 'tawe' and some geographical knowledge. But, given Frege's account of negation, '~*p*' is not at all like 'Abertawe' or 'Hansen's disease'. Frege does provide a fixed independent functional meaning for '┬(·)', the expression he introduces for negation. And we *can* substitute for '*p*' any sign with the same reference to get a completed functional expression which will be a name of the same truth-value as that named by '┬*p*'. So

[33] The name used as index need not be a proper part of the name in which it occurs. Thus 'Granny Smith', the name of a particular woman, can be treated as an index in Wittgenstein's sense, occurring 'in' the name of a kind of apple; cf. also the breed of dog Akita: the role of 'Akita' as index is clearer if we write the name as 'Akita Inu': Akita-dog. Both the long and short forms allude to the fact that this is *the* dog from Akita province.

how on earth can Frege be criticized for having treated propositions as indexes in his account of negation?

(2) Max Black (and others following him) have criticized *TLP*, 5.02 in a related way. Black argued that Frege would certainly have agreed with Wittgenstein that logical articulation was essential to the sense of a proposition, and that Frege would also have agreed that the sense of 'not-*p*' is a function of the sense of '*p*'.[34] But Black's argument as it stands cannot show that there is anything the matter with Wittgenstein's criticism, as can be seen by appeal to Frege himself. Frege points out that one can intend to define a function sign and fail to do so. This happens, he says, if one attempts to define the addition sign but fixes a meaning for it only for the case of adding real integers. Frege's criticism of such an attempt at definition is that it does not give genuine independent meaning to the addition sign. Instead, all it does is assign meaning to certain complex signs as wholes (signs of the form '*a* + *b*' where '*a*' on its own and '*b*' on its own designate real integers), but these signs lack logical complexity. (See *Gg*, vol. II, §65.) So Frege's argument shows that an attempt to fix a function-meaning for a sign can misfire; the result will be that a certain number of complex signs contain indexes, in Wittgenstein's sense. It follows from Frege's argument that Black's objection to *TLP*, 5.02 cannot work *on its own*. Frege's clear understanding that '~*p*' must have genuine logical complexity, and his recognition that the sense of '~*p*' must be a function of the sense of '*p*', do not by themselves show that Wittgenstein's criticism misses its mark. Black's objection shows only that, if Wittgenstein's criticism is correct, it comes to this: Frege's definition of the negation sign does not give it the kind of meaning the definition was clearly intended to provide. That is, the Wittgensteinian criticism of Frege is of the same kind as Frege's criticism of mathematicians who give inadequate definitions of '+'. The examination of Black's criticism does, though, enable us to see what really is a problem: *if* Frege's definition fails to give to the negation sign the logical character he

[34] Max Black, *A Companion to Wittgenstein's Tractatus* (Ithaca, N.Y.: Cornell University Press, 1964), p. 239; cf. also Anthony Kenny, 'The ghost of the *Tractatus*' in Godfrey Vesey (ed.), *Understanding Wittgenstein* (London: Macmillan, 1974), pp. 1–13, at p. 3.

plainly intended to give it, how has such a failure occurred? How could Frege have missed the mark he had clearly in view?[35]

(3) Michael Dummett has argued that, *at least* in the case of definite descriptions, it seems impossible to reject Frege's distinction between sense and reference.[36] For, in such cases, a speaker must know how the reference of the whole expression 'is determined in accordance with its composition out of its component words': the speaker must be able to understand what the *route to* reference provided by the sense is (whether he can actually *take* the route or not). Wittgenstein rejects the idea of any name (or indeed anything but a sentence) having sense. Yet it seems impossible to read *TLP*, 5.02 save as accepting *in all but words* the idea that definite descriptions and expressions like 'Frege's mother', or '5 + 2', have sense as well as reference. Wittgenstein's contrast between argument and index uses the notion of two kinds of route to the reference of complex expressions, and one of the two appears to be Fregean sense. In the case of a name with genuine functional complexity, we have a route to the reference of the complex sign through our grasp of the functional complexity of that sign. Wittgenstein does not use the word 'sense', but the kind of understanding he refers to in the case of a functionally complex name is exactly what Dummett is talking about when he speaks of grasp of how the reference of a complex sign is determined in accordance with its composition out of its component parts. In the contrasting case, the case of names which, like 'Abertawe', contain an index, we may well have a route to reference

[35] We should also note a pattern here. In §2, I quoted Wittgenstein, *Philosophical Investigations*, §339, in which Wittgenstein turns Frege's 'making the difference too slight' criticism against Frege's own account of numbers. In *TLP*, 5.02, we see Wittgenstein turning against Frege the kind of objection Frege raises in *Grundgesetze* to certain practices of definition. In 'Throwing away the ladder', I argued that Wittgenstein takes Fregean points about the significance of the difference between signs for functions and proper names and uses them in criticism of Frege by making parallel points about the difference between sentences and proper names. In §§5–7, I am going to be concerned with the *Tractatus* idea that Frege fails to pay adequate attention to the distinction between what it is for a sentence to have a sense and what he calls sense in the case of proper names. This is another use by Wittgenstein of a Fregean kind of criticism in criticizing Frege. Wittgenstein is not just deeply Fregean as a critic of other philosophers; he is strikingly Fregean in his criticisms of Frege.

[36] Michael Dummett, 'Frege's distinction between sense and reference', in Dummett, *Truth and Other Enigmas* (London: Duckworth, 1978), pp. 116–44, at p. 122.

via our grasp of the way such signs are constructed, but this is not logical composition. The names 'Abertawe' and 'Swansea' may have exactly the same sense (supposing we allow that syntactically simple geographical names have sense), which may not involve any connection with the river Tawe. Our being able to figure out the reference of 'Abertawe' from its construction and our geographical knowledge is not a case of our having a route to reference *provided by sense*. My argument here is that the contrast Wittgenstein makes in *TLP*, 5.02 between two sorts of route to the reference of complex names appears to makes use of the Fregean notion of sense in the case of genuinely functional complexity. So it seems arguable that he does not really reject the Fregean idea of logically complex names as having both sense and reference. He just does not *admit* that he keeps it. So the third problem with which we are confronted by *TLP*, 5.02 is why, if Wittgenstein *does* retain the Fregean distinction between sense and reference as it applies to definite descriptions, is this acceptance of the distinction suppressed?

5 SENSE AND REFERENCE (CONTINUED): INDEX AND ARGUMENT

I begin my discussion with Max Black's argument, outlined above, that Frege cannot be criticized, as Wittgenstein criticizes him, for failing to make the sense of '~*p*' depend on the sense of '*p*', because Frege does recognize that dependence of sense. I argued that Black missed the point, which is not that Frege failed to recognize that there was such a relation between the sense of '~*p*' and that of '*p*' but that his account of negation, in fact, and counter to his intentions, makes the sense of '~*p*' *not* depend on the sense of '*p*'. How can that have happened?

The sense of '*p*' is the sense of a sentence; it is thus the sense that the truth-conditions of '*p*' are fulfilled. Here I follow Frege's statement in vol. I, §32 of *Grundgesetze*. If it has been determined under what conditions a sentence denotes the True, the sentence has a determinate sense, a thought, the thought that the truth-conditions are fulfilled. While Frege does go on in §32 to speak about asserted sentences, his characterization of the sense of sentences is not meant to apply only to asserted sentences: the sense of a sentence, the thought that the truth-conditions are fulfilled, is the same whether the sentence is asserted or not. The sense of a sentence is *that* something is the case.[37] (While I have given Frege's views as views about sentences, Frege himself refers to names of truth-values, correctly formed from the signs of *Begriffsschrift*.) I shall be discussing the tension between Frege's recognition that what is in question is expressions the sense of which is *that something is the case* and his characterization of these expressions as names of truth-values.

In Frege's account of negation, the sign for negation is introduced this way:

The value of the function

┬ξ

shall be the False for every argument for which the value of the function

—ξ

is the True; and shall be the True for all other arguments.

Any expression which we put into the argument place marked by 'ξ' must name some object. Suppose that we put into that argument-place a sentence, say '5 > 4'. On Frege's view, the sentence has associated with it conditions in which it names the True; in all other conditions it would name the False. Its role as filler-of-that-argument-place is fulfilled perfectly adequately if it has associated with it such conditions, and if it names one of the two truth-values. It is no part of its role, as presenting an argument to the ┬ function, that it should have as its sense *that* such-and-such conditions are fulfilled. A sign like 'the positive square root of 4', which does not

[37] This is clearer in the Jourdain-Stachelroth and Beaney translations of *Grundgesetze* than in Furth's translation. Furth uses a participial construction after the expression 'the truth-value of'; Jourdain-Stachelroth and Beaney put a colon after 'the truth-value of' and then use a propositional construction. The latter is much closer to Frege's German. See, e.g., the translations of *Gg*, vol. I, §5, and of the immediately preceding footnote, *PW*, p. 156; Michael Beaney (ed.), *The Frege Reader* (Oxford: Blackwell, 1997), p. 215; and *Basic Laws of Arithmetic*, p. 37. It may seem as if saying that the sense of a sentence is that its truth-conditions are fulfilled suggests that only true sentences have sense. But that is not the way to read Frege's statement. The thought expressed by the sentence '*ist der, dass diese Bedingungen erfüllt sind*'; we must be able to grasp the thought irrespective of its truth or falsity; and indeed a false thought can be part of a true thought.

have as its sense *that* anything is the case (its having sense is its having fully determinate conditions for naming this or that object) would do exactly the same logical kind of work if it were put into the argument-place marked by 'ξ' as is done by '5 > 4', when it is in that argument-place. That '5 > 4' expresses the sense *that something is so* is irrelevant to its logical role in that context. (Here and in what follows I use the example of '5 > 4', ignoring the fact that it is a mathematical sentence, and so would not be treated by Wittgenstein in the way in which he treats ordinary contingent sentences. It is simply a short stand-in for any sentence expressing the thought that something is so.)

We can make the same point if we distinguish between a sign's having sense, when this means that it expresses a thought that something is so (in which case I shall speak of it as having sentential sense), and a sign's having sense, when this means that there are determinate conditions associated with the sign through which it has this or that reference. It is clear that an expression can have sense in the latter sense without having sense in the former; Frege's views about sense imply that all proper names other than sentences have sense only in the latter sense. Part of Wittgenstein's criticism of Frege at *TLP*, 5.02 could then be put this way: the sentential sense of '~*p*' *cannot* be understood unless the sentential sense of '*p*' has been understood, but the Fregean account of negation does not give the *sentential sense* of '*p*' any essential role. In §6, I discuss the difference between sentential and non-sentential sense. In the rest of §5, I assume that there is such a difference, and I use it in explaining some of the features of Wittgenstein's criticism of Frege.

Wittgenstein uses the notion of an index to make a stronger point than that the sentential sense of '*p*' plays no role in the sense of '~*p*' as Frege explains it. He is taking for granted the principle that it is only when a sentence is actually used as a sentence, only (that is) when its sentential sense is essential to its role, that it *is* a genuine sentence with sentential sense. If a sign which can express a thought occurs in an argument-place in which there can also occur signs which do not express thoughts, and which do not have sentential sense, then, in that argument-place, the sign is not genuinely a sentence expressing a thought that such-and-such is so. (This is itself an unFregean application of a Fregean point about how the argument place in which a sign occurs makes clear the logical character of the sign in that occurrence.)[38]

As we saw earlier, it is difficult to understand Wittgenstein's 'index' criticism of Frege, since, in Frege's account of '~*p*', '*p*' occurs in the argument-place of a function which has apparently been defined perfectly well – so how can Wittgenstein say that on Frege's account it occurs merely as an index? We can, however, see what Wittgenstein meant if we introduce a way of writing the sentences which we put into the argument-place of Frege's ┬ function. Suppose we put, instead of '┬(5 > 4)', '┬(Object$_{5>4}$)'. The expression 'Object$_{5>4}$' is written in such a way as to make clear that it is simply a name of an object. (It does not have sentential sense.) The expression '5 > 4', which can be used as a sentence, is written as a subscript, in the same way '*c*' is written as a subscript in Wittgenstein's first example of an index, in *TLP*, 5.02. Wittgenstein says that an index, like an argument, can enable one to recognize the reference of the sign containing the index. Just as one can recognize that the reference of 'Hansen's disease' is leprosy if one knows that Hansen provided a scientific description of leprosy, one can recognize that the reference of the name 'Object$_{5>4}$' is the True, if one knows that 5 is greater than 4.

To follow Wittgenstein's criticism we have to bear in mind different ways of using '5 > 4': (1) as a sentence, in which case it is essential to that use that the sentence expresses the thought that something is the case, and (2) merely as an expression which has as its reference some object (and in that case '5 > 4' lacks sentential sense). This second case can itself be subdivided. Just as we can contrast names like 'Hansen's disease' with names like 'Hansen's right foot', where the former lacks functional complexity but the latter has it, so '5 > 4', if it occurs in a context in which it lacks sentential sense, may still have (non-sentential) functional complexity, or it may lack functional complexity altogether. Wittgenstein's point that '*p*' occurs only as index in '~*p*' as Frege defines the latter can then be understood most easily if we connect it with a notation which clearly marks any occurrence of a sign which can in other contexts have sentential sense, but which lacks it in its particular

[38] See, e.g., Frege's treatment of the occurrence of 'Vienna' as a predicate, in 'Concept and object', *CP*, p. 189.

context, and which also lacks non-sentential functional complexity in that context. The sentence sign, in such contexts, is written in this notation as an indexed proper name. Without such a notation, it will be much easier to miss the crucial distinction between uses of '5 > 4' as index and uses as sentence-with-sentential-sense. The situation is like the cases mentioned in note 30, in which the index forms the entirety of an indexed name ('Akita' and 'Granny Smith'). Just as we might more easily recognize the use of the name as index in those cases by writing 'Akita Dog' or 'Granny Smith's Apple', we can more easily recognize the use of '5 > 4' as index by writing it as a mere part of a proper name.[39]

We may if we choose use the word 'sentence' of both sentences with sentential sense and expressions which lack sentential sense but which do name truth-values; but it would be an objection, from a Wittgensteinian point of view, to such a use of 'sentence' that the word covered expressions with two quite different kinds of logical character. A corresponding objection can be made if we use the word 'thought' (as Frege does) to cover both what I have called sentential sense (which is of the form 'that such-and-such is the case' or 'that such-and-such conditions are fulfilled') and also the sense of expressions like 'the truth-value of: Caesar is dead', or of expressions like 'the truth-value of the thought expressed by the first sentence on

[39] To avoid possible confusion, I should note that the 'Object_p' notation is not meant to be a way of writing Frege's horizontal function. It is a way of writing a sentence when it occurs in an argument-place which will take non-sentential proper names. Thus '$\text{Object}_{5>4}$' would be a way of writing '5 > 4' when it occurs in the argument-place of '–ξ'; it is not itself a completed function expression. Wittgenstein does not deny that there could be a non-sentential completed function expression written '5 > 4', which denoted some object; but, as far as I can see, his view is that the existence of a sentence-with-sentential sense, '5 > 4', does not imply that there is any completed function expression *looking* the same, and having functionally determined reference, which might occur in argument-places open to proper names. My guess here about the ideas in the background of Wittgenstein's view that '5 > 4' occurs without any kind of functional complexity when it is in an argument-place open to names, is that, for it to be a functionally complex name in such contexts, a new definition would have to be given of the greater-than sign, in which it meant a function with objects as values, but Frege does not provide any new definition. And, if he were to provide a new definition, his account of negation would be threatened: it is essential to the account that the greater-than sign, when it occurs within '5 > 4' in the argument-place of the negation-sign, should not have a different definition from that of the greater-than sign in '5 > 4' on its own.

p. 234 of Frege's *Collected Papers*'; these latter kinds of phrase do not express thoughts that anything is so. If an expression is in the argument-place of a function which takes objects as arguments, it does not have sentential sense, and we cannot somehow make it express the sense that something is the case. '5 > 4', when it expresses sentential sense, and '5 > 4', when it occurs in the argument-place of Frege's ┬ξ, may *look* the same, but Wittgenstein's view was that the difference in use goes also with a difference in the character of the sign: only in use as a sentence, with sentential sense, does '5 > 4' have the kind of logical articulation essential to sentences.

In the 'Notes on logic', Wittgenstein said that in not-p, p is exactly the same as if stands alone; he added that this point is absolutely fundamental.[40] The point is connected with his criticisms of Frege at *TLP*, 5.02: what Frege's definition of negation does is make '*p*' in '*~p*' *not* be the same as it is on its own. That is what is brought out by rewriting '┬*p*' as '┬Object$_p$'. When '*p*' stands on its own, its sense is that something is so, and its truth or falsity is agreement or disagreement with the way things are. If we do use the words 'truth-value' in connection with '*p*' on its own, what is meant is tied to the asymmetry of sentential sense: sentential sense has a kind of asymmetry in contrast with the sense of complex proper names.[41] We need to consider this contrast further, since it is important to our understanding of Wittgenstein's criticisms of Frege, and his rejection of Frege's way of distinguishing between sense and reference.

6 SENSE AND REFERENCE (CONTINUED): SENTENTIAL SENSE

Consider such names as 'Frege's birthplace', which we may view as formed by completing functional expressions. There will be,

[40] Wittgenstein, 'Notes on logic', p. 97.

[41] In the *Tractatus* Wittgenstein uses the expression 'truth-value' only in speaking about Frege's views; he takes the expression to be tied to Frege's understanding of sentences as names of objects. It would have been possible to introduce a different use of the term, consistent with the views of the *Tractatus*, but Wittgenstein perhaps believed that it would have been confusing to use a term so closely tied to what he took to be misunderstandings about the kind of meaning sentences have, and the kind of signs they therefore had to be. See also P. T. Geach, 'Truth and God', *Proceedings of the Aristotelian Society, Supplementary Volume*, 56 (1982), pp. 83–97, at p. 88.

associated with any such name, its sense; its having sense is its being determinate what the conditions are in which it would name any object in the converse domain of the function.[42] Thus, the Wismar-conditions of 'Frege's birthplace' are that Frege was born in Wismar, its Jena-conditions are that Frege was born in Jena, its Frankfurt-conditions that Frege was born in Frankfurt, and so on. The name names Wismar because the Wismar-conditions are fulfilled, but it does not have as its sense *that* the Wismar-conditions are fulfilled, or *that* any conditions are fulfilled. Even if a function has only two objects as possible values, the situation is exactly the same. Each completed function expression will have a pair of jointly exhaustive conditions attached to it. Suppose we have a function defined this way:

> The value of the function $f(x)$ shall be Wismar if x was born in Wismar; the value of the function $f(x)$ shall be Frankfurt if it is not the case that x was born in Wismar.

Then '*f*(Frege)' has associated with it the Wismar-condition that Frege was born in Wismar, and the Frankfurt-condition that it is not the case that Frege was born in Wismar. '*f*(Frege)' does not express the fulfilment of either its Wismar or its Frankfurt conditions. While we can investigate the state of things to find out which condition is fulfilled, this will not be a *comparison* of '*f*(Frege)' with reality, and neither the name nor its sense could be said to agree or to disagree with the way things are.

Wittgenstein rejects Frege's view that a sentence is a completed function-expression naming a truth-value. This is because any such expression would have exactly the kind of symmetrical sense – symmetrical with respect to two values – which 'f(Frege)' has. Its 'having a truth-value' would be something entirely different from the being true or false of a sentence. The root of the problem is that the Fregean view makes us unable to understand the kind of sense which Frege himself recognized that sentences have.[43] In the section

[42] This is on the model of truth-conditions. Any completion of a concept-expression or relational expression has associated with it the conditions in which it names the True; it names the False in all other conditions. Similarly one can speak of any completion of any function-expression as having associated with it conditions in which it names each of the objects which the function can have as its value.

[43] Peter Sullivan has argued a closely related point about how Frege's conception of sentences as complex names of truth-values makes it impossible for his account

of *Grundgesetze* quoted earlier, Frege says that well-constructed 'names of truth-values' have as their sense that their truth conditions are fulfilled; but Wittgenstein's view is that expressions which have *that such-and-such is so* as their sense cannot be *names of truth-values*. If we take any expression for a function which always has one or other of two objects as its value, and complete it, the result is a name the sense of which is of the same general type as other completed function-names naming objects, with which there is associated a set of conditions determining when the name names any of the objects which the function can have as value. Its sense is not that the conditions are fulfilled in which it names some one of these objects. Such a name may have 'truth-conditions' if this means: conditions in which it names a particular object; but it does not have truth-conditions in the sense in which a sentence does. A sentence has sense in that it is determinate how it is to be compared with reality, i.e., what counts as agreement or disagreement; the truth-conditions are the conditions of agreement; the thought it expresses is that those conditions are fulfilled. Sentential sense is thus expressible in a 'that'-clause. The determinacy of the sense of a functionally complex name, in contrast, is independent of there being anything that counts as 'agreement' or 'disagreement' with reality; its sense is not expressible in a 'that'-clause.[44] (The expression

to do justice to his own understanding of truth ('The sense of "a name of a truth-value"', *Philosophical Quarterly* 44 (1994), pp. 476–81). Frege's mature account in *Grundgesetze* 'recognizes no forms of complexity, characteristic of thoughts and their expression, which carry with them [the] intrinsic involvement with truth' which is essential to Frege's own conception of truth. The intrinsic involvement with truth, central in Sullivan's discussion of the issues, is inseparable from the kind of opposition of sense that there is between a sentence and its negation; see on this also Ricketts, 'Wittgenstein against Frege and Russell', pp. 243–4.

[44] This way of putting the contrast between the sense of complex names and that of sentences does not amount to a 'correspondence theory of truth' for sentences. I am suggesting that both Frege and Wittgenstein engage in what Thomas Ricketts speaks of as redescription of assertion and judgement, where the redescription is intended to 'highlight distinctive features' of our linguistic practices; see Thomas Ricketts, 'Objectivity and objecthood: Frege's metaphysics of judgment', in L. Haaparanta and J. Hintikka (eds.), *Frege Synthesized: Essays on the Philosophical and Foundational Work of Gottlob Frege* (Dordrecht: Reidel, 1986), p. 72; see also §7 of this essay. My way of putting the contrast between sentential sense and the kind of meaning that complex names have is meant to be part of such a redescription, based on Wittgenstein's criticism in the *Tractatus* of the Fregean redescription as misleading.

'truth-conditions' is thus capable of ambiguity. To avoid the ambiguity we might speak of sentences as having agreement-conditions; a proper name which has associated name-of-object conditions may have (as one case among others) naming-the-true conditions.) There are readers of Frege, like Tyler Burge, who have held that Frege was able to make a sharp distinction, using his horizontal function, between genuine sentences and mere completed function expressions not constructed as completions of the horizontal function sign, but it is not (on the account I am giving) possible to take the horizontal function as the basis of a distinction between genuine sentences and mere functionally complex names. An expression formed from completing the horizontal function sign with a complex name of a truth-value does not differ from the argument-sign in sense or reference; they have the same name-of-object conditions, and neither has genuine agreement-conditions.[45]

We have arrived at a version of the issue which Anscombe takes as central in her explanation of Wittgenstein's criticisms of Frege. What is required, she asks, for the 'T's and 'F's which we write in a truth-table in the columns under the propositions of a truth-functional proposition to have the same significance as the 'T's and 'F's giving the truth or falsity of the truth-functionally composite proposition itself?[46]

Anscombe goes into details of this issue; here I shall focus on one central element in Wittgenstein's move away from Frege on sense and reference. Wittgenstein, I have argued, was concerned with the need for the inner sentence in '~*p*' to occur with *sentential* sense. If, instead, what occurs in that position is a name, it will have associated with it, not a sentential sense, but (at most) conditions in which the name is a name of this or that object. The inner sentence must supply, as argument, a 'discrimination' of facts:[47] what counts as things being as the sentence represents them, and what counts as things not

[45] For an attempt to treat the horizontal as the basis of a logical distinction between sentences and complete names, see, e.g., Tyler Burge, *Truth, Thought, Reason: Essays on Frege* (Oxford: Clarendon Press, 2005), pp. 21–2; for the way in which such an attempt founders, see William W. Taschek, 'Truth, assertion, and the horizontal: Frege on "the essence of logic"', *Mind*, 117 (2008), pp. 375–401, part 4.

[46] See especially Anscombe, *Introduction*, pp. 51–3.

[47] See Wittgenstein, 'Notes on logic', p. 99 for this use of 'discrimination'.

being so. The functional character of '~*p*' is its reversing, not name-of-an-object conditions, but agreement/disagreement conditions; and so conditions of the latter sort must be supplied by the inner sentence. In the case of '5 > 4' and '~5 > 4', the truth or falsity of the sentences is determined by the same reality, namely the relative magnitude of 5 and 4.[48] Because '5 > 4' has a determinate comparison with *that* reality, agreeing or disagreeing with it, the reversal of that comparison (comparison with the same reality, but with agreement and disagreement reversed) is also a fully determinate comparison with a particular reality, a function of the comparison provided by '5 > 4'. This is how the sense of '~*p*' is a function of the sense of '*p*'.

It may seem as if Frege's account meets the requirement that the truth-conditions of the negation of '*p*' are the opposite of those of '*p*'; for '5 > 4' and '┬5 > 4', on his account, are true or false according as 5 > 4 is the true/is not the true, each sentence being made true by what makes the other not true. And indeed otherwise '┬' would not even appear to be the sign for negation. But Wittgenstein's view is that Frege's account does not accomplish what he intended. For it to succeed, three things must be the case: (1) '5 > 4' must have the same sense as '5 > 4 is the true' (which is writable as 'Object$_{5>4}$ is the true'); (2) '┬5 > 4' (which is writable as 'Object$_{5>4}$ is not the true') must genuinely be the negation of '5 > 4 is the true'; and (3) '┬5 > 4' must have the same sense as '5 is not greater than 4'.[49] 'Object$_{5>4}$ is not the true' is genuinely the negation of 'Object$_{5>4}$ is the true', but the sense of 'Object$_{5>4}$ is the true' is not the same as that of '5 > 4' (they

[48] I here follow Anscombe's suggestion (G. E. M. Anscombe, 'Grammar, structure, essence', *Arete* 12.2 (2000), pp. 113–20, at pp. 116–17) that Wittgenstein's references to *die Wirklichkeit* be translated, not as 'reality', but as 'the reality'. The point (in the sections with which she is concerned) is not that a picture depicts reality (reality in general), but that a picture depicts a particular reality. Different pictures may depict the same reality, but represent it differently: 'p · q' depicts the same reality as 'p v q'. The situation is, however, complicated: at some points in the *Tractatus*, '*die Wirklichkeit*' has to be taken to mean *die gesamte Wirklichkeit* (the sum-total of reality). Any picture which depicts some reality is thereby a picture depicting any reality which includes the former reality; hence every picture depicts *die gesamte Wirklichkeit*. Cf. also Wittgenstein's comments (Ludwig Wittgenstein, *Letters to C. K. Ogden with Comments on the English Translation of the Tractatus Logico-Philosophicus*, ed. G. H. von Wright (Oxford: Blackwell, and London: Routledge and Kegan Paul, 1973), p. 27) about the translation of *TLP*, 4.023.

[49] For the last step, see *Gg*, vol. I, §6, also *CP*, p. 150.

have entirely different agreement-conditions); and 'Object$_{5>4}$ is not the true' does not have the same sense as '5 is not greater than 4'. No determination of the sense of the latter is given by determining the sense of '┬5 > 4'; see on this last point *TLP*, 4.431. If we use a notation which does not distinguish between 'Object$_{5>4}$' and '5 > 4'-the-sentence-with-sentential-sense, i.e., if we write them both as '5 > 4', it is very easy to slide between them; we then may treat them as if they had truth-conditions in the same sense, and identical truth-conditions at that. That is what can make it appear that Frege's account of negation meets the requirement that the agreement-conditions of '*p*' and '~*p*' are genuinely opposite: the same properties or relations of things must be relevant to the determination of the truth/falsity of both, but in opposite ways. No treatment of negation as a Fregean concept can meet that requirement.

If the negation sign is merely a reverser of comparisons, then (as Wittgenstein notes) '*p*' could be used to express what '~*p*' is now used to express, and '~*p*' what '*p*' is used to express, without any change in the meaning of any sign. In this new use, '*p*' and '~*p*' are to be compared with the same reality as before; what is changed is merely what, for each of the two sentence-signs, counts as agreement and what counts as disagreement with that reality. The negation sign will be used in exactly the same way in both uses of the pair of sentence-signs. Its fixed use is to reverse a comparison, to form the contradictory of a sentence;[50] it does not have a sense or reference of its own. (It can be said to have *Bedeutung* if this means no more than that it has a determinate linguistic role.) The other signs in the two sentences also do not change in meaning. Their meaning determines (as in the previous use) the reality with which the sentences are to be compared; and that has not changed.

The possibility of the 'reversed' use of the sentence and its negation indicates how far we have moved from Frege's understanding of the sense–reference distinction. On Wittgenstein's view, the reality to which both '*p*' and '~*p*' are related is a function of the fixed reference of the (simple) names in the two sentences.[51] But the reference

[50] See Geach, 'Truth and God', p. 89; note also Geach's correction of Max Black's misunderstandings of the *Tractatus* on this matter. See also Diamond, 'Truth before Tarski', on reversibility of sense.

[51] The reality would depend on the arrangement of the names as well as on their reference. The reference of the simple names is not independently available, but is rather a feature internal to the use of sentences which have determinate

of the parts does not determine the truth or falsity of either sentence, because *that* depends on the 'direction' of the comparison with the reality. A sentence and its negation, which are opposite in sense (in agreement/disagreement conditions), and opposite in truth or falsity,[52] are not distinguished by the reference of any of their parts. Dummett describes as an essential part of Frege's notion of reference that the truth-value of a proposition is fully determined by the reference of its parts.[53] That feature is then *not* present in Wittgenstein's reworking of the distinction.[54]

There is one further important difference between Frege and Wittgenstein on sense and reference. The sense of a truth-function of '*p*' is held by both Frege and Wittgenstein to be a function of the sense of '*p*'. But sticking to that functional relation, on Wittgenstein's view, requires giving up the Fregean view that *in general* the sense of a sentence is a function of the sense of its parts. And giving up that latter Fregean view goes with Wittgenstein's limitation of the term 'sense' to what I have called sentential sense. There appears to be room on his account for some non-sentential expressions to have sense (and I have argued that this seems to be implicit in *TLP*, 5.02), but what it is for them to have a sense (if we do speak that way) is entirely different from what it is for a sentence to have sense: the sense of a complex name lacks 'directionality'; and the sense of a sentence (the agreement-conditions) is in any case not completely

agreement-conditions and which stand in determinate logical relations to each other. What the simple names 'stand for' is whatever it is sentences so used show they are about.

52 We can say 'opposite in truth-value', if we recall that truth and falsity, on Wittgenstein's view, are not values of a function.

53 Michael Dummett, *Interpretation of Frege's Philosophy* (London: Duckworth, 1981), p. 150; see also Burge, 'Truth, thought, reason', p. 22.

54 This difference between Frege and Wittgenstein is connected with another important difference between them, concerning the logical character of the signs which result, on the one hand, from removing one or more names from a complex name of an object, and, on the other hand, from removing one or more names from a sentence. Wittgenstein follows Russell in distinguishing sharply between propositional functionality and non-propositional functionality. This difference has ramified consequences within the *Tractatus* which I cannot discuss here. On the determination of the truth-value of a proposition by the reference of its parts, see below, §7. Goldfarb (in 'Wittgenstein's understanding of Frege') takes the contrast between Wittgenstein's and Frege's understanding of functionality to be central in his account of Wittgenstein's criticism of Frege at *TLP*, 5.02; on Goldfarb's account see Appendix.

functionally fixed by any feature of the meaning of its parts. It is dependent on a feature of the sentence's use, its direction of comparison with the reality to which the sentence is related, and *that* can be changed without any change in meaning of any part of the sentence.[55]

7 SENSE AND REFERENCE (CONCLUDED)

In Part VII, I am concerned with the philosophical issues involved in Wittgenstein's rethinking of the distinction between sense and reference. It will be useful first to consider the aim of Frege's and Wittgenstein's accounts of negation, and the background of shared views.

Part of Frege's inheritance from Kant, passed on to Wittgenstein, is the idea of the primacy of judgement; tied to this is the view that what a sentence expresses is something that *can* be the content of a judgement. I take Wittgenstein also to share with Frege an understanding of what talk of sense and reference is responsible *to*. But what such talk is responsible to depends on the sort of philosophical activity to which that talk belongs and on its aim. Here I turn to a remark of Peter Geach's, that there is, in all informative discourse, an inchoate understanding of the logical interconnections between truth on the one hand and the 'reversibility' of the sense of our sentences. About this understanding, he says that it 'can be clarified or sharpened by logical and philosophical training, but there can be no question of analysis or explicit definition'.[56] Geach himself engages in the kind of philosophical activity, of articulating what is involved in informative discourse, that is suggested by his

[55] We should also note that the implied allowance of sense, of a sort distinct from sentential sense, to some expressions other than sentences does not extend to the simple names of the *Tractatus*. (There is no 'route' to the reference of simple names.) The complex expressions which might be said to have non-sentential sense can, unlike simple names, fail to have reference because of what is or is not the case. The possibility of their lacking reference does not, within the *Tractatus*, lead to there being sentences which lack truth-value, since Wittgenstein accepts Russell's theory of descriptions. This implied allowance of sense belongs to what is essentially an abbreviated way of speaking about whole sentences; see the Appendix and note 78.

[56] Geach, 'Truth and God', p. 94.

remark. There is a similar conception of a philosophical activity of articulation of informative discourse in Thomas Ricketts's reading of Frege. Ricketts has described the overall aim of Frege's discussions of thought and language as *redescription* of certain features of our linguistic practices, redescription intended to bring out sharply various aspects of Frege's conception of objectivity.[57] In the practices of assertion and inference, and especially in the recognition, within these practices, of logical constraints, there is embodied an understanding of what is objective and independent of psychology, in contrast with what belongs merely to the psychological life of individuals (our ideas, our emotional responses). As Ricketts reads Frege, Frege's talk of sense and reference is not meant to go behind practices of assertion and inference, but is answerable to those practices, so far as they embody recognition of the distinction between objective and subjective.[58]

Wittgenstein, as I am reading him, is engaged in a similar kind of redescription of linguistic practice, bringing to attention some of its salient features. We can see Wittgenstein's treatment of sense and reference and Frege's as alternative articulations; and this is to see them as intellectual alternatives, but intellectual alternatives of a different kind from alternative theories. Wittgenstein's talk of sense and reference is meant to be answerable to what is implicit in ordinary informative discourse; and so far as what Wittgenstein says is meant as a criticism of Frege, it takes Frege's talk of sense and reference to be answerable to the same practices. The possibility of such criticism depends on both of them recognizing as salient some central logical features of those practices; it is the possibility of Wittgenstein's bringing out how Frege's own account fails to do justice to those central logical features of our practices that the account was meant to articulate.[59]

Wittgenstein takes his own way of thinking about the notion of sentence-sense to be Fregean, as we can see at *TLP*, 4.431: what a sentence expresses is its truth-conditions.[60] His distinction

[57] Ricketts, 'Objectivity and objecthood', p. 72.

[58] See Ricketts, 'Generality'; cf. also Michael Kremer, chapter 7, this volume, §8.

[59] In the last two paragraphs, I have drawn on my discussion of philosophical method in Diamond, 'Truth before Tarski', pp. 256–7.

[60] See Appendix for some further discussion of this point.

betweeen index-complexity and function-argument-complexity reflects his agreement with Frege about what it is for an expression in a sentence to have reference: its having reference in a particular sentence is inseparable from the role of the whole sentence in inferences. (This is how we distinguish the occurrence of 'Hansen' in 'Smith has Hansen's hat' from its occurrence in 'Smith has Hansen's disease'.) If we start from these Fregean views, there is indeed a prima facie case for taking sentences which occur truth-functionally in other sentences to have reference in something like the way a proper name does. But there is a question how far the analogy between sentences and names goes, how far there is an analogy between sentential and non-sentential functionality. Frege himself lets the analogy shape his way of developing the distinction between sense and reference, while Wittgenstein appeals to Fregean ideas in disentangling the distinction from the analogy. What I am concerned with here is that we can see the value of some such distinction in a description of sentence-functionality (and its relation to inference), and can see how Wittgenstein's criticisms of Frege's way of making the distinction do not depend on importing some new aim into the discussion, but on the argument that Frege was mistaken in taking his own aims to be compatible with the understanding of sentences as proper names. Truth-functional inference depends on sentences, when they occur truth-functionally in other sentences, *not* occurring there with the kind of meaning characteristic of complex proper names. That at any rate was Wittgenstein's view.[61]

Of logical constraints on judgement-making, the most important concerns contradiction. Here there is a more specific agreement between Frege and Wittgenstein: our understanding of what it is for a sentence '~*p*' to express the thought contradicting that expressed by another, '*p*', depends on a functional relation between the two. The story we tell about the functional relation must enable us to see how '~*p*', by containing the sentence '*p*', expressing the thought

[61] I am not concerned here with the further ramifications within the *Tractatus* of Wittgenstein's view of truth-functional inference. What is important is that the idea that truth-functional occurrences of sentences are occurrences of sentences with sentential sense (not of expressions with the logical role of proper names) makes possible the *Tractatus* account of logical inference as not in need of laws of inference.

it does, itself expresses the opposite thought. Call this the common complex aim of their discussions of negation.

Wittgenstein's criticisms of Frege take for granted not just the common complex aim, but also the idea that talk of sense and reference subserves that aim: no account of sense and reference which results in an incoherent account of the functional relation between '*p*' and '~*p*' does what talk of sense and reference overall is meant to do.

The common complex aim requires that '*p*' in '~*p*' should be the same as it is on its own: the story told about negation will make clear what that 'being the same' is.[62] Frege's own account was intended to meet that demand: it was supposed to be met by the sense and reference of '*p*' on its own being the same as the sense and reference of '*p*' in '~*p*'. And what it is for '*p*' to be true is supposed to be the same, whether it occurs in '~*p*' or on its own. Wittgenstein's objections to Frege are not just that he gives an inadequate account of negation but that the account does not do what Frege himself intended it to do; hence the importance in his criticism of the point that Frege does not make the sense of '~*p*' a function of the sense which '*p*' has on its own. Wittgenstein's criticism takes for granted the common complex aim; that '*p*' in '~*p*' must have a fully sentential sense is essential to that aim. An account of negation will fail if, according to it, '*p*' in '~*p*' does not express *that* something is the case but is merely an expression naming an object. Despite Frege's intention that the sense of '~*p*' be functionally dependent on that of '*p*', his treating of sentences as occurring in argument-places open to names deprives them of sentential sense, and so frustrates his intentions. Frege's explanation of his negation sign, reflecting as it does his failure to distinguish sharply enough between sentential sense and the kind of meaning complex names have, is thus taken by Wittgenstein to be a serious flaw in Frege's construction of a conceptual notation. Wittgenstein's reshaping of the distinction between sense and reference is thus part of a larger argument which we can see in the

[62] Another essential part of the story concerns assertion, and whether Frege's view implies that unasserted sentences lack genuine verbs. '*p*' on its own will not be the same as '*p*' in '~*p*' on any account in which the former has and the latter lacks a genuine verb. For discussion of these issues see Proops, 'The early Wittgenstein', and Diamond, 'Truth before Tarski'.

Tractatus: the argument that Frege's own aims are undermined by his analogy between sentences and complex proper names.[63]

The problems to which Wittgenstein draws attention cannot be solved by minor alterations in Frege's account of negation, which result in sentences being distinguished from proper names, but not distinguished sharply enough. Thus, for example, if we were to deny that sentences *name* truth-values, and if we were to allow no non-sentences into the argument-place of the negation sign, but if we otherwise kept Frege's account intact, Wittgenstein's criticisms would still apply. The fundamental criticism was that '~*p*', as Frege explained it, was not dependent on the *sentential* sense of '*p*', on its having truth-conditions which are genuinely agreement-conditions, not just conditions in which one rather than the other of two values is associated with '*p*'. Whether it 'names' the value is not important; what it is for it to *have* the value is. If '*p*' in '~*p*' is not being used as a sentence with sentential sense (if its sentential sense is irrelevant to its role as argument), then we ought to avoid writing it in such a way as to make it look as if it really is the sentence '*p*', because that creates the illusion that there is genuine functionality. If '*p*' in '~*p*' merely supplies a semantic value, then it should be written like this: 'SV_p', to make clear that '*p*', the sentence with sentential sense, is not what is in the argument-place.

Michael Dummett disputes the point for which I have been arguing. He explains two distinct notions of truth-value: one notion is required for the understanding of what is involved in asserting a sentence which stands on its own and the other is, he says, required in explaining the role of sentences as constituent parts of truth-functionally compound sentences. He says then that there is no a priori reason why the two notions of truth-value should coincide.[64] He does not comment on that claim; it seems to be meant to be an obvious or unquestionable point. But in fact it defines a point of view; he is giving up the demand that '*p*' in '~*p*' should be the same as it is on its own. (Here I follow Wittgenstein in taking it that, if what it is for a particular asserted sentence to have a truth-value is not the same as what it is for the identical-looking unasserted sentence, forming

[63] See Geach, 'Saying and showing' and 'Truth and God'.

[64] Michael Dummett, *Frege: Philosophy of Language* (London: Duckworth, 1973), p. 417.

part of another sentence, to have a truth-value, then the resemblance between the asserted sentence and the sentence which is part of another is not a mark of a logical relation between the asserted sentence and the sentence which contains the sentence looking like the asserted sentence.) An account of sense and reference which follows Frege in recognizing the demand that '*p*' should be the same on its own and in '~*p*', and which treats sentential sense as primary, is actually closer to Frege's in some important respects (despite departing from it in the obvious ways in which Wittgenstein's does) than are neo-Fregean accounts like Dummett's, which abandon either the demand itself or the primacy of sentential sense, or both. How close a treatment of sense and reference is to Frege's is not obvious from the surface of things but depends on the philosophical place from which one is judging (which determines how one sees the connections between Frege on sense and reference and his other views).

What one takes to be essential to the conceptions of sense and reference itself reflects one's understanding of logic. To make the contrast between Dummett's approach and Wittgenstein's clear, we need to bear in mind that, for Wittgenstein, a sentence could be used to express the sense which we now express by its negation, and that that difference in the sentence's sense would involve no change in the meaning of any sign. The difference between a sentence and its negation is a difference in the 'direction' of comparison with reality; and though we have a sign, the negation sign, that reverses sense, we have no sign that indicates what the direction of comparison is; that belongs to the use of the sentence (which could be the opposite of what it is). So it is essential to what a sentence is, on Wittgenstein's view, that the truth or falsity of the sentence is not functionally determined by the meaning of the parts. Dummett describes the determination of the truth-value of a sentence by the reference of its parts as essential to Frege's understanding of reference; it is also essential to Dummett's own understanding of the role of reference in semantic theory. On Wittgenstein's view, the idea that the truth-value of a sentence is determined by the reference of its parts, far from being essential to the notion of reference, reflects failure to distinguish deeply enough between the functionality of sentences and that of complex names. That failure then makes it impossible to see clearly the character of the kind of disagreement between two sentences, in which exactly what it is for one sentence to be true is what it is for the other to be false, and vice versa. Wittgenstein's insight here is that the *heart* of what is meant by sentence-sense (the expression by a sentence of truth-conditions) is tied to the possibility of such opposition, and therefore to the possibility of *reversal* of sense, the possibility of reversing the truth-conditions of the opposed sentences (reversing the direction of comparison). Here truth-conditions are not conditions for a sentence's having a certain semantic value. The 'semantic value' account makes irrelevant to the occurrence of '*p*' in '~*p*' what Wittgenstein means by the directionality of sense; hence the 5.02 objection, that '*p*' is thereby turned into a mere index, applies. And this blocks understanding of the character of truth-functional inference (and hence blocks understanding of the character of all inference): the internality of inference to what our sentences themselves are depends on the truth-functional occurrence of a sentence being its occurrence with sentential sense.

Frege made clear, as no one had before him, how very easy it is to equivocate if one does not have a notation in which expressions with different logical roles are written differently. So (as he noted) even someone as acute as Hilbert was able to slide unwittingly between the use of 'point' for a first-level concept and its use for a second-level concept.[65] Wittgenstein's argument against Frege resembles such arguments of Frege's own: Frege is unwittingly equivocating, and can do so because his notation fails to mark clearly the logical features which belong to the use of sentences. The expression by a pair of written or spoken sentences of contradictory thoughts depends upon what the *Tractatus* speaks of as the projective use of sentence-signs. In such use, the sentence-signs express thoughts; and sentence-signs, even as parts of other sentence-signs, are in such a projective relation to the world. Only in such use do the signs have sense. The difference between sentences and names is lost to view if we fail to consider the projective use of sentence-signs. The trouble with Frege's notation is that it flattens out what belongs to the projective use of signs; we cannot see in the notation the difference between '*p*' used as a sentence, and '*p*' used merely as an index, hence the ease with which it is possible to slide unwittingly between the two uses. The notation encourages what one might think of as a kind of blindness to use, blindness to logical differences dependent on use.

[65] *PMC*, pp. 93–4.

8 CONCLUDING COMMENTS: ON INHERITING FROM FREGE

What I have said in §§4–7 puts before us a particular picture of negation, roughly this: if one once gets clear about '~*p*' as reversing the sense of '*p*', one will have got the essence of negation. At the same time, one will have made clear such further logical relations as that '~~*p*' has the same truth-conditions, the same sense, as '*p*', since its sense is nothing but the reversal of the reversal of the sense of '*p*'. In his later philosophy, Wittgenstein treats that picture as completely misleading, as itself a reflection of a kind of blindness to use.[66] Is it possible for him, within the context of his later philosophy, to continue to take seriously Frege's insistence on the separation of the logical from the psychological, the objective from the subjective? I cannot here answer that question; the point here is rather that he is guided in his later philosophy by the desire to take those distinctions as seriously as Frege did, but to avoid at the same time a mythology of what it is to take them seriously. The argument of *Philosophical Investigations*, at two extremely significant junctures, reminds us how important the conception of logic was which Wittgenstein inherited from Frege: at §108, where the question is what remains of the rigour of logic if logical concepts are not pure essences but families of more or less related structures, and at §§240–2, where Wittgenstein faces the question whether he is not giving up the Fregean distinction between what human beings agree about and what is true. An aim of Wittgenstein's in *Philosophical Investigations* is to show that he is *not* giving up that distinction, not bargaining away the rigour of logic. Just as, in the *Tractatus*, a central question is what it will be to inherit the distinction between sense and reference, a main question for Wittgenstein later is what it will be to inherit from Frege respect for the rigour of logic, and for the distinction between psychological and logical, between subjective and objective.

I have not here argued directly against Michael Dummett's view that the greatness of Frege as a philosopher lies primarily in his recognition of the foundational importance, within philosophy, of the theory of meaning, of the attempt to arrive at a general theoretical

[66] See especially the discussion, beginning at Wittgenstein, *Philosophical Investigations*, §89, of subliming the logic of our language.

understanding of what the meaning of the expressions of our language consists in. I have worked with an account of Frege as engaged, not in a theoretical undertaking of the sort sketched by Dummett, but rather in redescribing certain features of our practices of assertion and inference.[67] My aim has been to present a different *line of inheritance* from Frege. What we – now – can inherit from Frege is itself a philosophical question; part of Wittgenstein's greatness as a philosopher, one of the things he passes down to us, is how he took the question of inheritance from Frege.

APPENDIX: QUESTIONS ABOUT INFLUENCE

Warren Goldfarb begins his stimulating essay 'Wittgenstein's understanding of Frege: the pre-Tractarian evidence' (2002) by quoting the Preface to the *Tractatus*, where Wittgenstein acknowledges his debt for much of the stimulation of his thoughts to 'Frege's great works and the writings of my friend Mr. Bertrand Russell'. Goldfarb comments: 'What is less clear is the relative weights of those influences'. He argues that the documentary evidence makes it clear how Wittgenstein, starting from within a Russellian conception of judgement and truth, works his way to his own views, but that there is no such evidence that Wittgenstein worked with a comparably deep-going understanding of Frege's thought. That Wittgenstein reached views which are in some respects strikingly like some of Frege's may show, not that he was influenced by Frege, but that a Russellian approach, 'when its implications are followed out strictly', coincides closely with conclusions reached by following out similarly strictly the implications of a Fregean judgement-based approach to analysis. In the course of his discussion Goldfarb criticizes the argument of §§4–7 of this present essay and some of the claims of §1. In this Appendix I cannot deal with all the questions

[67] See Ricketts, 'Frege, the *Tractatus*, and the logocentric predicament', 'Generality, meaning and sense', 'Objectivity and objecthood', and 'Logic and truth in Frege', *Proceedings of the Aristotelian Society, Supplementary Volume*, 70, (1996), pp. 121–40; cf. also Joan Weiner, 'Has Frege a philosophy of language?', in Tait, *Early Analytic Philosophy*, pp. 249–72. If the reading of Frege which I take for granted helps us to see the relation between Frege's thought and that of Wittgenstein, that would be some evidence that the reading was similar to Wittgenstein's.

he raises, but I want first to comment on the general issue, and then to turn to two points from Goldfarb's discussion of §§4–7.

How did Frege's 'great works' stimulate Wittgenstein as he worked on the *Tractatus*? Goldfarb is right that the question cannot be settled by considering merely the similarities between Frege's views and those of Wittgenstein. To bring out the difficulty of the question it is worth looking at one particular similarity on which Wittgenstein himself commented much later. In 1939, Wittgenstein said that the truth-table schematism was not his invention but Frege's; what *was* his invention was the use of the table as a symbol for the proposition rather than as an explanation for it.[68] Frege didn't actually use a tabular schematism in explaining the content of truth-functional propositions. So we have to read Wittgenstein as having meant that Frege, in his setting out of the explanation of his symbols, had got at something significant, something which Wittgenstein thought was reflected in his own use of truth-tables in the *Tractatus*, and which made it appropriate to take the 'real' inventor of the truth-table to be Frege. What then exactly had Frege done? In *Begriffsschrift*, he had explained the content of propositions constructed, using his symbolism, from one or more propositions; in the case of constructions from two propositions, he set out the explanation by specifying which of the four combinations (of affirmation and denial) of the argument-propositions are 'allowed' by the constructed proposition. Call the point that this itself specifies a content the first truth-table point. To see how Frege's procedure shows us what is (and what isn't) included in 'conceptual content' (content relevant to inference) and to see the similarity to truth-tables as understood in the *Tractatus*, we should note that, in the course of Frege's explanations (in §7), he treats the case in which content B is affirmed and the negation of content A denied as being the case in which B is affirmed and A is affirmed. The explanation in terms of the four possible combinations of affirmed and denied contents is treated implicitly as explanation in terms of the four rows of a truth-table; Frege makes without comment a move in which a second negation of a content (in the structure of the truth-table row) simply cancels the first negation. He does not treat the case as involving an inference (on which he would have had to

[68] Wittgenstein, *Lectures on the Foundations of Mathematics*, p. 177.

comment). Here we have gone further than the first truth-table point, and arrived at what I think is the heart of the similarity that led Wittgenstein to say that it was Frege who invented truth-tables. Why? Well, suppose we have a complicated truth-functional construction from two propositions. *All* that matters, on Wittgenstein's early view, in the way the truth-table rows taken together show us the content of the constructed proposition, is, in each row, the T or F for each of the two argument-propositions and the T or F in the column for the constructed proposition. Wittgenstein had also developed another notation, the a-b notation described in the 'Notes on logic', which had the same feature: only the connection between the innermost a-b poles and the outermost signifies; so, for example, a doubly negated proposition is simply the same symbol as the original proposition. What Wittgenstein took to be visible in truth-tables and in his own a-b notation, he took (I think) to be also visible in Frege's presentation of his symbolism in §§5–7 of *Begriffsschrift*. It belongs to the character of composite propositions that their content is determined merely by the correlation of each possible combination of truth or falsity of argument-propositions with truth or falsity for the composite proposition. Call that the second truth-table point; to invent truth-tables is then to invent a notation which makes perspicuous such determination of propositional content. Frege's two-dimensional symbols can be read as constructing correlations between combinations of truth and falsity for the argument-propositions and the truth or falsity of the constructed proposition, where the truth-correlations themselves are what fix the content of the constructed proposition as a function of that of the argument-propositions. How the correlation works *between* the truth-value combinations for the arguments and the truth-value for the whole proposition does not bear on content; it is not part of what symbolizes, and parts of it can simply cancel out. (If I am right that the two truth-value points are included in what Wittgenstein meant in 1939 by 'inventing the truth-table', then 'inventing the truth-table' is distinct from what it is usually taken to be, namely the invention of a method for evaluating formulae or arguments.) When Wittgenstein writes about truth-functionality (in the 'Notes on logic') he doesn't say that Frege's (implicitly truth-tabular) explanations of his symbolism have made plain a central feature of the nature of propositions. *Was* he, though, stimulated by Frege's explanation of truth-functional propositions? Reflection on what can be read as internal to Frege's method and his explanations could lead directly to such central ideas of Wittgenstein's as the distinction between logical connectives and function-expressions. The truth-table points are themselves among Wittgenstein's framework ideas; the first is stated at *TLP*, 4.431, the second is formulated in the 'Notes on logic' ('All that is essential about molecular functions is their T-F schema'), where there is an apparent connection with *Begriffsschrift*, §7.[69] The *Begriffsschrift* method of explaining truth-functional propositions reflects Frege's conception of what is logically significant in propositional content, namely, what matters for inference. The truth-table points can be taken to be tied to Frege's notion of judgeable content, and to his understanding of the primacy of judgement. If Wittgenstein, in his recognition of the truth-table points and his understanding of their connection with the nature of the logical connectives, was stimulated by Frege's explanation of truth-functional propositions, that would be a case in which he had followed out the implications of something deep-going in Frege's approach. But his central ideas about the nature of the logical connectives could also be reached in a different way by following through implications of Russell's very different views.[70] The force which those central ideas had for Wittgenstein may have come in part from the existence of different routes apparently leading to the same place. When Wittgenstein raises problems, then, about Frege's views, from how deep an understanding of Frege's thought do the problems come? *He* is going to see what is deep and central in Frege's thought in terms of his own understanding of where and how

[69] Wittgenstein, 'Notes on logic', p. 100. Wittgenstein says there that it follows from Frege's explanations of 'not-p' and 'if p then q' that 'not-not-p' designates the same as 'p'. Why does he drag in 'if p then q'? That would be totally mysterious unless he is referring to Frege's account of the combination of 'if p then q' and 'not-p' in a single proposition, most likely the account in *Bs*, §7. In *Gg*, vol. I, §12 there is also a cancellation of a double negative in Frege's explanation, but Wittgenstein's remark about what follows from Frege's explanations seems to apply rather to the *Begriffsschrift* account. See Ricketts, 'Wittgenstein against Frege and Russell', for a discussion of the importance for Wittgenstein of the contrast between Frege's 1879 explanations of his symbolism and his 1892 explanations. Ricketts also quotes a remark from 'Notes on logic' in which Wittgenstein links his fundamental idea about the logical connectives with what I call the second truth-table point.

[70] See Ricketts, 'Wittgenstein against Frege and Russell'.

Frege leads us to the heart of what is involved in a proposition's having sense. So, although there may be, within Frege's scheme of ideas, the possibility of responding to Wittgenstein's criticisms, Wittgenstein would not himself have taken such possibilities seriously if the responses which we can envisage involve going against what comes out in (let us say, possibly) the *Begriffsschrift* quasi-truth-tabular explanations. One might say that there, in those explanations, there is (from Wittgenstein's point of view) something deeper than Frege's 'scheme of ideas' and that the 'scheme of ideas' is properly answerable to it. Goldfarb is right to emphasize that Wittgenstein's criticisms of Russell don't stand in the same relation to Russell's views as the relation of his criticisms to Frege's; but the different relation nevertheless leaves room for Wittgenstein to have been genuinely following through implications of what is present in Frege's thought. My comment here depends on a disagreement with Goldfarb's claim that in Wittgenstein's pre-*Tractatus* writings there is no hint of influence from Frege's judgement-based approach to analysis. Wittgenstein's account of composite propositions precludes any view of such propositions as built up from all their parts including the logical connectives. This leaves unanswered the question how exactly Wittgenstein's emerging ideas about logical composition exerted pressure on his whole account of the nature of propositions. (Ricketts has emphasized that Frege did, in unpublished writings that Wittgenstein would not have known, acknowledge sense 'in something like the way Wittgenstein understands this notion', i.e., that he had a conception close to Wittgenstein's of the opposition between a sentence and its negation. In *Begriffschrift*, Ricketts says, Frege took the iterability of the logical connectives to be intrinsic to them; but his development of an account of that iterability in terms of the notion of a mathematical function made it impossible for his developed theory to respect his own insights about the opposition of sense between a sentence and its negation.[71] My argument here has focused on what I have called the truth-table points, but it is closely related to that of Ricketts; my account provides another formulation of the same issues. The truth-table points are themselves closely tied to the issue raised in §§4–7: how Frege's recognition that even unasserted sentences have as their sense that

[71] *Ibid.*; see especially pp. 228–9, 244–6 and 249 note 49.

something is so (that their truth-conditions are fulfilled) can be squared with his treatment of those sentences when they occur within another sentence in a truth-functional context. My general comment here leaves intact Goldfarb's argument that the strong claims that Geach, Ricketts and I have made about Frege's influence need to be re-examined.

I cannot here consider all the questions raised by Goldfarb in his discussion of *TLP*, 5.02 and his criticism of my reading of it. I shall touch on two issues: how the passage bears on Frege's overall account of functions, and whether there is available a relatively simple reading of 5.02, alternative to mine. Goldfarb reads the passage as a dismissal of Frege's general understanding of functions; and he connects that reading with a much simpler account than mine of Wittgenstein's criticism of Frege. One disagreement which partly shapes our different treatments of *TLP*, 5.02 should be mentioned here. Goldfarb says that Wittgenstein accepts Russell's view that all complexity is propositional complexity and that all functions are propositional functions.[72] I read *TLP*, 5.02 as making an allowance for some notion of functional complexity applicable to some complex names, and I see the whole passage, then, as concerned to bring out the incoherence of any attempt to treat truth-functional complexity as the complexity of a name, taking for granted that there is some sort of possibility of functional complexity for names.

Goldfarb's argument starts from the fact that it is a consequence of Frege's understanding of functions that there is no trace in the *Bedeutung* of a completed function expression (a sentence or other proper name) of the *Bedeutungen* of the parts of the expression. This, Goldfarb argues, means that 'the occurrence of the particular sub-sentential parts in the sentence is not essential to the identity of the value of the sentence; those parts thus occur as indexes'.[73] Goldfarb notes that a possible defence of Frege would appeal to his notion of sense in order to distinguish between sub-sentential parts that are arguments of functions and those that are indexes, but Wittgenstein will not allow such an appeal to two levels of meaning.[74] So, on that

72 Goldfarb, 'Wittgenstein's understanding of Frege', p. 195.

73 *Ibid.*, p. 196.

74 Here, as elsewhere in discussion of *TLP*, 5.02, it is important to note that the argument–index distinction for Wittgenstein is applied to parts of complex signs, including signs that for Frege count as completed function signs, i.e., signs that

reading of Wittgenstein, it would follow quite simply that Frege, not being allowed by Wittgenstein to appeal to a two-level story about meaning, would be unable to account for the logical character of composite propositions, including in particular logical propositions. Goldfarb's simpler reading implies that, as Wittgenstein reads Frege, Frege cannot distinguish between the role of 'Parkinson' as index in a sentence containing 'Parkinson's disease' and its role as argument in a sentence containing 'Parkinson's birthplace', since, if Frege is not allowed the appeal to a two-level account of meaning, *no* sub-sentential part could be recognized as an argument rather than as an index. That's crucial, but I think it is questionable. Although Wittgenstein doesn't accept Frege's account of functions, I think that, for the purposes of 5.02, he is not concerned with its application to anything but those sub-sentential parts of sentences that are themselves sentences. He says, in 5.02, that both arguments and indexes enable one to recognize the *Bedeutung* of the sign of which they are part, but his idea is that they don't do this in the same way. In the case of function-argument complexity, the occurrence of the argument in the sign is not arbitrary; it is (this is the implication) essential to the way we are able to recognize what the *Bedeutung* is of the whole sign. It isn't obvious that it would follow that the argument must be essential to the identity of the value of the function; what is needed is (as I read 5.02) only that the argument is essential to recognizing what the *Bedeutung* is. So, on this reading of 5.02, it wouldn't be obvious that a Fregean understanding of functions makes arguments into indexes, even though it is true that it makes the *Bedeutung* of 'Parkinson' inessential to the identity of Edinburgh (supposing Edinburgh to be the *Bedeutung* of 'Parkinson's birthplace'). The question is whether a Fregean understanding of functions allows any account of the essential role of the argument in the recognition of what the *Bedeutung* of the whole sign is. On a Fregean account of functions, we are able to recognize some place, say Edinburgh, as *Bedeutung* of 'Parkinson's birthplace' through the occurrence of 'Parkinson' within the phrase, because a phrase

contain one or more argument-signs. What for Frege would count as an argument-*sign*, and not as an index, if the sign does indeed have genuine functional complexity, would count, if we were able to apply Wittgenstein's distinction, as an *argument*. This difference in language does not affect the issue, and so far as possible I use Wittgenstein's language.

formed on the model of '*x*'s birthplace' *bedeutet*, in accordance with a general rule, the place where the person named by the word we put for '*x*' was born. (There is a partial parallel with Wittgenstein's explanation in *The Blue Book* of how arguments differ from indexes, where he uses the example of two ways of understanding 'Bright's disease': so that it denotes a particular illness, 'Bright' then being mere index, and so that it means the disease Bright has, 'Bright' then being argument of the function '*x*'s disease'.)[75] The Fregean approach I have described is not so much an appeal to a Fregean two-level account of meaning as it is a use of the kind of thinking that might make a Fregean account of sense appear attractive, by providing an example of what could be meant by talk of how a *Bedeutung* is presented. The logical differences between index and argument would show up in a Fregean account in the effects of substitutions, which show the location of argument-places. So, for example, if one is speaking about James Parkinson, one can substitute 'James' for 'Parkinson' in 'Parkinson's birthplace' and one will be speaking of the same place; but 'James's disease' is not another name for Parkinson's disease. That way of making the distinction between argument and index is not ruled out by the absence of a trace of the *Bedeutung* of 'Parkinson' in Edinburgh.[76] So my claim is that Wittgenstein wouldn't deny that Frege can explain how the occurrence of a non-sentential argument plays an essential role in our recognition of what the *Bedeutung* is of the whole sign of which it is part. In the background of his discussion there is the possibility of interpreting in a Russellian way functions of the sort Frege took to be paradigmatic.[77] Wittgenstein's language in 5.02 abstracts, indeed, from the complexity of his own view about the contrast between argument and index, when he speaks quite simply of our recognizing the *Bedeutung* of the signs containing an index or an argument.

[75] See Wittgenstein, *The Blue and Brown Books*, p. 21.

[76] The substitution test provides a necessary but not sufficient condition for a sign to be an argument. See Frege's use of what is in effect the argument–index distinction in *Gg*, vol. II, §65, where he argues that a failure to fix the reference for every possible complex sign of the form '$a + b$' has the result that particular signs of that form are not genuinely functionally complex. Their arguments are, in Wittgenstein's language, mere indexes. That they are indexes is not revealed by the effects of substitutions, as in the 'Parkinson's disease' sort of case, but can be shown only through a complex argument.

[77] See also Hylton, 'Functions, operations and sense', pp. 95–6.

That way of talking apparently allows for logical complexity within such phrases as 'Parkinson's birthplace'; what Wittgenstein means would need spelling out in terms of a Russellian reading of the sentences containing such phrases.[78] If a sentence contains 'Parkinson's birthplace' and identifies it or ascribes a property to it, a Russellian restatement of the sentence is possible, in which the word 'Parkinson' occurs as naming something to which there is related one and only one place-of-being-born. (On my reading, then, Wittgenstein would not accept that a satisfactory 'pure' Fregean account could be given of what is involved in the occurrence of phrases like 'Parkinson's birthplace' in propositions, but he is not concerned in 5.02 with the fact that the distinction between the two phrases, so far as it is available to Frege, would, if thought through, require a different account from Frege's of what it is for such signs to occur in propositions.) As I read 5.02, then, Wittgenstein is arguing that the distinction between argument and index, so far as it may be treated, for the purposes of his argument, as available to Frege, doesn't enable Frege to deal with the case in which the argument is supposed to be an unasserted sentence; the trouble with Frege's account of functions – the trouble in view in 5.02 – is that the account allows functions in his sense to have sentences as arguments. So I read the strong claim made by Wittgenstein, that the index–argument confusion lies at the root of Frege's account of the meaning of propositions and functions, to be an abbreviated way of saying that Frege's conception of propositions as complete signs and his conception of functions on the model of arithmetical functions, taken together, give him

[78] Compare Ricketts, 'Wittgenstein against Frege and Russell'. Ricketts notes that, on the *Tractatus* view, the completion of a Fregean function-sign by a sign for its argument is a complex name, the structure of which has no representational relevance. See also Goldfarb, 'Wittgenstein's understanding of Frege', p. 195. Wittgenstein, Goldfarb says, takes the Russellian view that all complexity is propositional complexity. I take Wittgenstein to be willing to speak in an abbreviated way of phrases like 'Parkinson's birthplace' as having logical structure. They can be spoken of, in this abbreviated way, as having logical structure, so far as we consider them with their occurrence in sentences in which the parts of the phrase are recognizably expressions in the *Tractatus* sense; speaking of logical structure within the phrase is a short way of speaking of the logical structure of sentences containing the phrase, and the phrase would not be spoken of, in a fuller statement, as itself having logical structure. Thus this abbreviated way of speaking has in the background an unFregean contrast between sentences and descriptive phrases.

no options. The two together reduce propositions occurring within other propositions to mere indexes; this is a fundamental flaw in Frege's whole conception of functions and propositions. (An advantage of this reading is that it helps us to see why Wittgenstein put the argument-index criticism at 5.02. Wittgenstein's idea is that, if the distinctive character of truth-functional construction is not seen (not seen to belong to what a proposition itself is, as specified in *TLP*, 5), if it is not separated sharply from functionality as thought of by Frege, we lose our understanding of how the sense of 'not-*p*' depends on that of '*p*' (and in general of how the sense of composite propositions depends on that of their arguments) and therefore also our understanding of what is shown by the propositions of logic. These issues are connected also with the use Wittgenstein makes of Frege's context principle, which undergoes a shift in significance as he links it with his own fundamental ideas, but I cannot go further into these matters here.)

Goldfarb's account of *TLP*, 5.02 is indeed much simpler than mine, but there is a further point that should be made about the cost of that simplicity. Goldfarb's simpler account fits with his claim that Wittgenstein shares Russell's view that all complexity is propositional complexity and that all functions are propositional functions.[79] The difficulty with that reading is Wittgenstein's apparent allowance of function-argument complexity to some complex names, at *TLP*, 5.02, which appears to make use of the same argument–index distinction explained at greater length in *The Blue Book*.[80]

[79] Goldfarb, 'Wittgenstein's understanding of Frege', p. 95.

[80] I am very grateful to Warren Goldfarb for his analysis of the issues, to James Conant, Hans Sluga, Susan Haack, Thomas Ricketts, Michael Potter and Juliet Floyd for comments and discussion, and to Benjamin Bennett for help with questions about translation.

BIBLIOGRAPHY

WORKS BY FREGE

The following abbreviations are used throughout:

Bs	*Begriffsschrift*
Gl	*Die Grundlagen der Arithmetik*
Gg	*Grundgesetze der Arithmetik*

The standard English translation of *Grundlagen* is:
Foundations of Arithmetic, trans. J. L. Austin, 2nd edn (Oxford: Blackwell, 1953)
English translations of Frege's shorter works are to be found in the following four collections:

CP	*Collected Papers on Mathemtaics, Logic and Philosophy*, ed. Brian McGuinness (Oxford: Blackwell, 1984);
PMC/WB	*Philosophical and Mathematical Correspondence*, ed Gottfried Gabriel *et al.* (Oxford: Blackwell, 1980); English translation of his *Wissenschaftlicher Briefwechsel*, ed. Hans Hermes *et al.* (Hamburg: Meiner, 1976);
PW/NS	*Posthumous Works*, ed. Hans Hermes *et al.* (Oxford: Blackwell, 1979); English translation of his *Nachgelassene Schriften*, ed. Hans Hermes *et al.* (Hamburg: Meiner, 1969);
CN	*Conceptual Notation and Other Works*, trans. and ed. Terrell Ward Bynum (Oxford: Oxford University Press, 1972).

The following chronological list (which excludes his correspondence) makes no attempt to list all the English translations of Frege that have been published or all the places in which they have appeared.

'Über eine geometrische Darstellung der imaginären Gebilde in der Ebene. Inaugural-Dissertation der philosophischen Facultät zu Göttingen zur Erlangung der Doctorwürde' (Jena: Neuenhahn, 1873).

(Trans.: Kaal) 'On a geometrical representation of imaginary forms in the plane (Doctoral dissertation in the Philosophical Faculty of Göttingen)', *CP*, pp. 1–55.

'Rechnungsmethoden, die sich auf eine Erweiterung des Größenbegriffes gründen. Dissertation zur Erlangung der Venia Docendi bei der Philosophischen Fakultät in Jena' (Jena: Rommann, 1874).

(Trans.: Kaal) 'Methods of calculation based on an extension of the concept of quantity (Dissertation for the Venia docendi in the Philosophical Faculty of Jena)', *CP*, pp. 56–92.

'Rezension von H. Seeger, *Die Elemente der Arithmetik, für den Schulunterricht bearbeitet*', *Jenaer Literaturzeitung*, 1 (1874), p. 722.

(Trans.: Kaal) 'Review of H. Seeger, *The Elements of Arithmetic*', *CP*, v93–4.

'Rezension von A. v. Gall und Ed. Winter, *Die analytische Geometrie des Punktes und der Geraden und ihre Anwendung auf Aufgaben*', *Jenaer Literaturzeitung*, 4 (1877), pp. 133–4.

(Trans.: Kaal) 'Review of A. v. Gall and E. Winter, *The Analytical Geometry of Point and Line and its Application to Problems*', *CP*, p. 98.

'Rezension von J. Thomae, *Sammlung von Formeln welche bei Anwendung der elliptischen und Rosenhain'schen Funktionen gebraucht werden*', *Jenaer Literaturzeitung*, 4 (1877), p. 472.

(Trans.: Kaal) 'Review of J. Thomae, *Collection of Formulae used in the Application of Elliptical and Rosenhain Functions*', *CP*, p. 98.

'Über eine Weise, die Gestalt eines Dreiecks als complexe Grösse aufzufassen', *Sitzungberichte der Jenaischen Gesellschaft für Medizin und Naturwissenschaft*, 12 (1878), Supplement, p. xviii.

(Trans.: Kaal) 'Lecture on a way of conceiving the shape of a triangle as a complex quantity', *CP*, pp. 99–100.

Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens (Halle: Nebert, 1879).

(Trans. 1: Bauer-Mengelberg) in J. van Heijenoort (ed.), *From Frege to Gödel*, pp. 1–82.

(Trans. 2: Bynum) 'Conceptual notation, a formula language of pure thought modelled upon the formula language of arithmetic', *CN*, pp. 101–203.

'Anwendungen der Begriffsschrift', *Jenaische Zeitschrift für Naturwissenschaft*, 13 (1879), Supplement II, pp. 29–33.

(Trans.: Bynum) 'Applications of the conceptual notation', *CN*, pp. 204–8.

Logik (unpublished, 1879–91).

(Trans.: Long and White) 'Logic', *PW*, pp. 1–8.

Rezension von Hoppe, *Lehrbuch der analytischen Geometrie, I Teil*', *Deutsche Literaturzeitung*, 1 (1880), columns 210–11.

(Trans.: Kaal) 'Review of Hoppe, *Textbook of Analytic Geometry I*', *CP*, pp. 101–2.

'Booles rechnende Logik und die Begriffsschrift' (unpublished, 1880/1).
(Trans.: Long and White) 'Boole's logical calculus and the concept-script', *PW*, pp. 9–46.
[Unpublished, 1882?].
(Trans.: Long and White) '[17 key sentences on logic]', *PW*, pp. 174–5.
'Über den Briefwechsel Leibnizens und Huygens mit Papin', *Jenaische Zeitschrift für Naturwissenschaft*, 15 (1882), Supplement, pp. 29–32.
'Booles logische Formelsprache und meine Begriffsschrift' (unpublished, 1882).
(Trans.: Long and White) 'Boole's logical formula-language and my concept-script', *PW*, pp. 47–52.
'Über die wissenschaftliche Berechtigung einer Begriffsschrift', *Zeitschrift für Philosophie und philosophische Kritik*, 81 (1882), 48–56.
(Trans. 1: Bartlett) 'On the scientific justification of a concept-script', *Mind*, 73 (1964), pp. 155–60.
(Trans. 2: Bynum) 'On the scientific justification of a conceptual notation', *CN*, pp. 83–9.
'Über den Zweck der Begriffsschrift', *Jenaische Zeitschrift für Naturwissenschaft*, 16 (1883), Supplement, pp. 1–10.
(Trans. 1: Dudman) 'On the purpose of the *Begriffsschrift*', *Australasian Journal of Philosophy*, 46 (1968), pp. 89–97.
(Trans. 2: Bynum) 'On the aim of the conceptual notation', *CN*, pp. 90–100.
[Unpublished, before 1884].
(Trans.: Long and White) '[Dialogue with Punjer on existence]', *PW*, pp. 53–67.
Die Grundlagen der Arithmetik (Breslau: Koebner, 1884).
(Trans.: Austin) *The Foundations of Arithmetic: A Logico-mathematical Enquiry into the Concept of Number* (Oxford: Blackwell, 1950).
'Geometrie der Punktpaare in der Ebene', *Sitzungberichte der Jenaischen Gesellschaft für Medizin und Naturwissenschaft*, 17 (1884), Supplement, pp. 98–102.
(Trans.: Kaal) 'Lecture on the geometry of pairs of points in the plane', *CP*, pp. 103–7.
'Rezension von H. Cohen: Das Princip der Infinitesimal-Methode und seine Geschichte', *Zeitschrift für Philosophie und philosophische Kritik*, 87 (1885), pp. 324–9.
(Trans.: Kaal) 'Review of H. Cohen, *The Principle of the Method of Infinitesimals and its History*', *CP*, pp. 108–11.
'Erwiderung', *Deutsche Literaturzeitung*, 6 (1885), column 1030.
(Trans.: Kaal) 'Reply to Cantor's review of *Foundations of Arithmetic*', *CP*, pp. 122.

'Über formale Theorien der Arithmetik', *Sitzungberichte der Jenaischen Gesellschaft für Medizin und Naturwissenschaft*, 19 (1886), Supplement 2, pp. 94–104.
(Trans.: Kluge) 'On formal theories of arithmetic', *CP*, pp. 112–21.
[Unpublished, 1890–2].
(Trans.: Long and White) '[Draft towards a review of Cantor's *Gesammelte Abhandlungen zur Lehre vom Transfiniten*]', *PW*, pp. 68–71.
'Über das Trägheitsgeset', *Zeitschrift für Philosophie und philosophische Kritik*, 98 (1891), pp. 145–161.
(Trans. 1: Rand) 'About the law of inertia', *Synthese*, 13 (1961), pp. 350–63.
(Trans. 2: Kaal) 'On the law of inertia', *CP*, pp. 123–36.
Funktion und Begriff (Jena: Pohle, 1891).
(Trans.: Geach) 'Function and concept', *CP*, pp. 137–56.
'Über den Begriff der Zahl' (Unpublished, 1891/92).
(Trans.: Long and White) 'On the concept of number', *PW*, 72–86.
'Über Sinn und Bedeutung', *Zeitschrift für Philosophie und philosophische Kritik*, 100 (1892), pp. 25–50.
(Trans.: Black) 'On sense and meaning [i.e. reference]', *CP*, pp. 157–77.
[Unpublished, 1892–5].
(Trans.: Long and White) '[Comments on sense and meaning [i.e. reference]]', *PW*, pp. 118–25.
'Rezension von Georg Cantor: Zur Lehre vom Transfiniten', *Zeitschrift für Philosophie und philosophische Kritik*, 100 (1892), pp. 269–272.
(Trans.: Kaal) 'Review of Georg Cantor, *Contributions to the Theory of the Transfinite*', *CP*, pp. 178–81.
[Unpublished, 1892].
(Trans.: Long and White) 'Draft of "On concept and object"', *PW*, pp. 87–117.
'Über Begriff und Gegenstand', *Vierteljahrsschrift für wissenschaftliche Philosophie*, 16 (1892), pp. 192–205.
(Trans.: Geach) 'On concept and object', *CP*, pp. 182–94.
Grundgesetze der Arithmetik, vol. I (Jena: H. Pohle, 1893).
(Trans.: Furth) of Preface, Introduction, §§1–52, *The Basic Laws of Arithmetic: Exposition of the System*, vol. I (Berkeley: University of California Press, 1964).
'Rezension von Dr. E. G. Husserl: Philosophie der Arithmetik I', *Zeitschrift für Philosophie und philosophische Kritik*, 103 (1894), pp. 313–32.
(Trans.: Kaal) 'Review of E. G. Husserl, *Philosophy of Arithmetic I*', *CP*, pp. 195–209.
'Kritische Beleuchtung einiger Punkte in E. Schröders Vorlesungen über die Algebra der Logik', *Archiv für systematische Philosophie*, 1 (1895), pp. 433–56.

(Trans.: Geach) 'A critical elucidation of some points in E. Schröder', *Lectures on the Algebra of Logic'*, *CP*, 210–28.
'Le nombre entier', *Revue de Métaphysique et de Morale*, 3 (1895) pp. 73–8.
(Trans.: Dudman) 'Whole numbers', *CP*, pp. 229–33.
'Logik' (unpublished, 1897).
(Trans.: Long and White) 'Logic', *PW*, pp. 126–51.
'Lettera del sig. G. Frege all'Editore', *Revue de Mathématiques (Revista di Matematica)*, 6 (1896–9), pp. 53–9.
(Trans.: Kaal) 'Letter to the editor', *PMC*, pp. 112–18.
'Über die *Begriffsschrift* des Herrn Peano und meine eigene', *Berichte über die Verhandlungen der Königlich Sächsischen Gesellschaft der Wissenschaften zu Leipzig. Mathematisch-Physische Klasse*, 48 (1897), pp. 361–78.
(Trans.: Dudman) 'On Mr Peano's conceptual notation and my own', *CP*, pp. 234–48.
[Unpublished, 1897/8 or shortly afterwards].
(Trans.: Long and White) 'The argument for my stricter canons of definition', *PW*, pp. 152–6.
Über die Zahlen des Herrn H. Schubert (Jena: Pohle, 1899).
(Trans.: Kaal) 'On Mr H. Schubert's numbers', *CP*, pp. 249–72.
Grundgesetze der Arithmetik, vol. II (Jena: Pohle, 1903).
(Trans.: Geach and Beaney) of §§55–67, §§138–47 and Appendix, M. Beaney (ed.) *The Frege Reader* (Oxford: Blackwell, 1997), pp. 258–89.
(Trans.: Black) of §§86–137 in P. T. Geach and M. Black (eds.), *Translations from the Philosophical Writings of Gottlob Frege*, 3rd edn (Oxford: Blackwell, 1980), pp. 162–213.
'Logische Mängel in der Mathematik' (unpublished, 1898/9 or later, probably not after 1903).
(Trans.: Long and White) 'Logical defects in mathematics', *PW*, pp. 157–66.
'Über Euklidische Geometrie' (unpublished, 1899–1906?).
(Trans.: Long and White) 'On Euclidean geometry', *PW*, pp. 167–9.
'Über die Grundlagen der Geometrie', *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 12 (1903), pp. 319–24 and 368–75.
(Trans. 1: Szabo) 'The foundations of geometry', *Philosophical Review*, 69 (1960), pp. 3–17.
(Trans. 2: Kluge) 'On the foundations of geometry: first series', *CP*, pp. 273–84.
[Unpublished, after 1903].
(Trans.: Long and White) '[Frege's Notes on Hilbert's "Grundlagen der Geometrie"]', *PW*, pp. 170–3.
'Was ist eine Funktion?', in *Festschrift Ludwig Boltzmann gewidmet zum sechzigsten Geburtstage, 20. Februar 1904* (Leipzig: Barth, 1904), pp. 656–66.

(Trans.: Geach) 'What is a function?', *CP*, pp. 285–92.
'Über die Grundlagen der Geometrie', *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 15 (1906), pp. 293–309, 377–403 and 423–30.
(Trans.: Kluge) 'On the foundations of geometry: second series', *CP*, pp. 285–92.
'Über Schoenflies: Die logischen Paradoxien der Mengenlehre' (unpublished, 1906).
(Trans.: Long and White) 'On Schoenflies: Die logischen Paradoxien der Mengenlehre', *PW*, pp. 176–83.
'Antwort auf die Ferienplauderei des Herrn Thomae', *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 15 (1906), pp. 586–90.
(Trans.: Kluge) 'Reply to Mr Thomae's holiday *causerie*', *CP*, pp. 341–5.
'Was kann ich als Ergebnis meiner Arbeit ansehen?' (unpublished, August 1906).
(Trans.: Long and White) 'What may I regard as the result of my work?', *PW*, pp. 184.
[Unpublished, August 1906].
(Trans.: Long and White) 'Introduction to logic', *PW*, pp. 185–96.
'Kurze Übersicht meiner logischen Lehren' (unpublished, 1906).
(Trans.: Long and White) 'A brief survey of my logical doctrines', *PW*, pp. 197–202.
'Die Unmöglichkeit der Thomaeschen formalen Arithmetik aufs Neue nachgewiesen', *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 17 (1908), pp. 52–6.
(Trans.: Kluge) 'Renewed proof of the impossibility of Mr Thomae's formal arithmetic', *CP*, pp. 346–50.
'Footnotes to Philip E. B. Jourdain: The development of the theories of mathematical logic and the principles of mathematics: Gottlob Frege' (*The Quarterly Journal of Pure and Applied Mathematics*, 43 (1912), 237–269), *PMC*, 179–206.
'Logik in der Mathematik' (unpublished lecture notes, spring 1914).
(Trans.: Long and White) 'Logic in mathematics', *PW*, pp. 203–50.
[Unpublished, 1915?].
(Trans.: Long and White) 'My basic logical insights', *PW*, pp. 251–2.
'Vorschläge für ein Wahlgesetz' (unpublished, 1918). In Gottfried Gabriel and Uwe Dathe (eds.), *Gottlob Frege: Werk und Wirkung* (Paderborn: Mentis, 2000), pp. 297–313.
'Der Gedanke', *Beiträge zur Philosophie des deutschen Idealismus*, 1 (1918), pp. 58–77.
(Trans. 1: Quinton and Quinton) 'The thought: A logical enquiry', *Mind*, 65 (1956), pp. 289–311.
(Trans. 2: Geach and Stoothoff) 'Logical investigations, I: Thoughts', *CP*, pp. 351–72.

'Die Verneinung', *Beiträge zur Philosophie des deutschen Idealismus*, 1 (1918), pp. 143–57.
(Trans.: Geach) 'Logical investigations, II: Negation', *CP*, pp. 373–89.
[Unpublished, July 1919].
(Trans.: Long and White) '[Notes for Ludwig Darmstaedter]', *PW*, pp. 253–7.
'Gedankengefüge', *Beiträge zur Philosophie des deutschen Idealismu*s, 3 (1923), pp. 36–51.
(Trans.: Stoothoff) 'Logical investigations, III: Compound thoughts', *CP*, pp. 390–406.
'Logische Allgemeinheit' (unpublished, not before 1923).
(Trans.: Long and White) 'Logical investigations, IV: Logical generality', *PW*, pp. 258–62.
Tagebuch (unpublished, 23 March 1924–25 March 1924).
(Partial trans. 1: Long and White) '[Diary entries on the concept of numbers]', *PW*, pp. 263–4.
(Partial trans. 2: Mendelsohn) 'Diary: Written by Professor Dr Gottlob Frege in the time from 10 March to 9 April 1924', *Inquiry*, 39 (1996), pp. 303–42.
'Zahl' (unpublished, September 1924).
[Trans.] 'Number', *PW*, pp. 265–6.
'Erkenntnisquellen der Mathematik und der mathematischen Naturwissenschaften' (unpublished, 1924/5).
(Trans.: Long and White) 'Sources of knowledge of mathematics and the mathematical natural sciences', *PW*, pp. 267–74.
'Zahlen und Arithmetik' (unpublished, 1924/5).
(Trans.: Long and White) 'Numbers and arithmetic', *PW*, pp. 275–7.
'Neuer Versuch der Grundlegung der Arithmetik' (unpublished, 1924/5).
(Trans.: Long and White) 'A new attempt at a foundation for arithmetic', *PW*, pp. 278–81.

OTHER WORKS

Allison, H. E., *Kant's Transcendental Idealism* (New Haven: Yale University Press, 1983).
Anscombe, G. E. M., 'Grammar, structure, essence', *Arete*, 12 (2000), pp. 113–20.
An Introduction to Wittgenstein's Tractatus (London: Hutchinson University Library, 1959).
'The reality of the past', in her *Metaphysics and the Philosophy of Mind* (Oxford: Blackwell, 1981), pp. 103–19.
Antonelli, A. and May, R., 'Frege's new science', *Notre Dame Journal of Formal Logic*, 41 (2000), pp. 242–70.

Aristotle, *Politics*, ed. and trans. C. D. C. Reeve (Indianapolis: Hackett, 1998).

Arnauld, A., and Nicole, P., *Logic or the Art of Thinking*, trans. J. Vance Buroker (Cambridge: Cambridge University Press, 1996).

Avigad, Jeremy, 'Methodology and metaphysics in the development of Dedekind's theory of ideals', in J. Ferreirós and J. Gray (eds.), *The Architecture of Modern Mathematics* (Oxford: Oxford University Press, 2006), pp. 159–86.

Beaney, M., *Frege: Making Sense* (London: Duckworth, 1996).

(ed.), *The Frege Reader* (Oxford: Blackwell, 1997).

Beaney, M., and Reck, E. (eds.), *Gottlob Frege: Critical Assessments of Leading Philosphers, vol. III: Frege's Philosophy of Mathematics* (Routledge, 2005).

Bell, D., *Frege's Theory of Judgment* (Oxford: Clarendon Press, 1979).

'How "Russellian" was Frege?', *Mind*, 99 (1990), pp. 267–77.

'Thoughts', *Notre Dame Journal of Formal Logic*, 28 (1987), pp. 36–50.

Bell, E. T., *The Development of Mathematics* (New York: McGraw-Hill, 1940).

Bernacerraf, P., 'What the numbers could not be', in P. Benacerraf and H. Putnam (eds.), *Philosophy of Mathematics: Selected Readings*, 2nd edn (Cambridge: Cambridge University Press, 1983), pp. 272–94.

Bernays, P., 'Die Bedeutung Hilberts für die Philosophie der Mathematik', *Die Naturwissenschaften*, 10 (1922), pp. 93–9; trans. in P. Mancosu (ed.), *From Brouwer to Hilbert: The Debate in the Foundations of Mathematics in the 1920s* (Oxford: Oxford University Press, 1998), pp. 189–97.

Black, M., *A Companion to Wittgenstein's Tractatus* (Ithaca, N.Y.: Cornell University Press, 1964).

'Frege on functions', reprinted in E. D. Klemke (ed.), *Essays on Frege* (Urbana: University of Illinois Press, 1968), pp. 223–48.

Blackburn, S., and Simmons, K. (eds.), *Truth* (Oxford: Oxford University Press, 1999).

Blanchette, P., 'Frege's reduction', *History and Philosophy of Logic*, 15 (1994), pp. 85–103.

Boolos, G., 'Frege's theorem and the Peano postulates', *Bulletin of Symbolic Logic*, 1 (1995), pp. 317–26.

'Is Hume's principle analytic?', in G. Boolos, *Logic, Logic and Logic*, pp. 301–14; reprinted in R. Heck (ed.), *Language, Thought and Logic: Essays in Honour of Michael Dummett* (New York and Oxford: Oxford University Press, 1997), pp. 245–61.

'On second-order logic', *Journal of Philosophy*, 72 (1975), pp. 509–27.

Logic, Logic and Logic, ed. Richard Jeffrey (Cambridge, Mass.: Harvard University Press, 1998).
Boolos, G., and Heck, R. G., '*Die Grundlagen der Arithmetik* §§82–83', in G. Boolos, *Logic, Logic and Logic*, pp. 315–38.
Brandom, R. B., *Making It Explicit* (Cambridge, Mass.: Harvard University Press, 1994).
Articulating Reasons: An Introduction to Inferentialism (Cambridge, Mass.: Harvard University Press, 2000).
Burge, T., 'Frege on extensions of concepts from 1884 to 1903', *Philosophical Review*, 93 (1984), pp. 3–34; reprinted in T. Burge, *Truth, Thought, Reason: Essays on Frege*, pp. 273–98.
'Frege on knowing the foundation', *Mind*, 107 (1998), pp. 305–47.
'Frege on truth', in L. Haaparanta and J. Hintikka (eds.), *Frege Synthesized*, pp. 97–154.
'Truth and singular terms', *Noûs*, 8 (1974), pp. 167–81; reprinted in M. Platts (ed.), *Reference, Truth and Reality* (London: Routledge and Kegan Paul, 1980), pp. 309–25.
Truth, Thought, Reason: Essays on Frege (Oxford: Clarendon Press, 2005).
Burgess, J. P., *Fixing Frege* (Princeton: Princeton University Press, 2005).
'Frege and arbitrary functions', in W. Demopoulos (ed.), *Frege's Philosophy of Mathematics*, pp. 89–107.
Cantor, G., 'Über eine elementare Frage der Mannigfaltigkeitslehre', *Jahresbericht der deutschen Mathematiker-Vereinigung*, 1 (1891), pp. 75–8.
Caplan, B., and Thau, M., 'What's puzzling Gottlob Frege?', *Canadian Journal of Philosophy*, 31 (2001), pp. 159–200.
Carl, W., *Frege's Theory of Sense and Reference: Its Origins and Scope* (Cambridge: Cambridge University Press, 1994).
Carnap, R., 'Intellectual biography', in P. A. Schilpp (ed.), *The Philosophy of Rudolf Carnap*, pp. 3–84.
Clark, P., and Read, S. (eds.), *Reading Putnam* (Oxford: Blackwell, 1994).
Coffa, J. A., *The Semantic Tradition from Kant to Carnap: To the Vienna Station*, ed. L. Wessels (Cambridge: Cambridge University Press, 1991).
Conant, J., 'Elucidation and nonsense in Frege and early Wittgenstein', in A. Crary and R. Read (eds.), *The New Wittgenstein* (London: Routledge, 2000), pp. 174–217.
'The method of the *Tractatus*', in E. Reck (ed.), *From Frege to Wittgenstein*, pp. 374–462.
'On going the bloody *hard* way in philosophy', in J. Whittaker (ed.), *The Possibilities of Sense* (Basingstoke: Palgrave, 2002), pp. 85–129.
'Putting two and two together: Kierkegaard, Wittgenstein and the point of view for their work as authors', in T. Tessin and M. von

der Ruhr (eds.), *Philosophy and the Grammar of Religious Belief* (Basingstoke: Macmillan, 1995), pp. 248–331.

'The search for logically alien thought: Descartes, Kant, Frege, and the Tractatus', *Philosophical Topics*, 20 (1991), pp. 115–80.

'Wittgenstein on meaning and use', *Philosophical Investigations*, 21 (1998), pp. 222–50.

Coolidge, J. L., *The Geometry of the Complex Domain* (Oxford: Clarendon Press, 1924).

Crossley, J. N., 'A note on Cantor's theorem and Russell's paradox', *Australasian Journal of Philosophy*, 51 (1973), pp. 70–1.

Davidson, D., *Inquiries into Truth and Interpretation* (Oxford: Clarendon Press, 1984).

'On the very idea of a conceptual scheme', in D. Davidson, *Inquiries into Truth and Interpretation*, pp. 183–99.

'Radical interpretation', in D. Davidson, *Inquiries into Truth and Interpretation*, pp. 125–39.

Dedekind, R., ' Brief an Keferstein', in M.-A. Sinaceur, 'L'Infini et les nombres', *Revue d'Histoire des Sciences*, 27 (1974), pp. 251–78; trans. in J. van Heijenoort (ed.), *From Frege to Gödel*, pp. 99–103.

Essays on the Theory of Numbers, trans. W. W. Beman (New York: Dover, 1963).

Gesammelte mathematische Werke, vol. III, ed. Robert Fricke *et al.* (Braunschweig: Vieweg, 1932).

'The nature and meaning of numbers', in *Essays on the Theory of Numbers,* trans. W. W. Beman (New York: Dover, 1963).

Stetigkeit und irrationale Zahlen (Braunschweig: Vieweg, 1872); trans. in W. Ewald (ed.), *From Kant to Hilbert*, pp. 765–79.

Theory of Algebraic Numbers, trans. J. Stillwell (Cambridge: Cambridge University Press, 1996).

Was sind und was sollen die Zahlen? (Braunschweig: Vieweg, 1888); trans. in W. Ewald (ed.), *From Kant to Hilbert*, pp. 787–833.

Dehn, M., 'Die Legendre'sche Sätze über die Winkelsumme im Dreieck', *Mathematische. Annalen*, 53 (1900), pp. 404–39.

Demopoulos, W., 'Frege, Hilbert and the conceptual structure of model theory', *History and Philosophy of Logic*, 15 (1994), pp. 211–25.

(ed.), *Frege's Philosophy of Mathematics* (Cambridge, Mass.: Harvard University Press, 1995).

Demopoulos, W., and Clark, P., 'The logicism of Frege, Dedekind, and Russell', in S. Shapiro (ed.), *Oxford Handbook of Philosophy of Mathematics and Logic* (New York: Oxford University Press, 2005), pp. 129–65.

Detlefsen, M., 'Fregean hierarchies and mathematical explanation', *International Studies in the Philosophy of Science*, 3 (1988), pp. 97–116.

Devitt, M., *Designation* (New York: Columbia University Press, 1981).

Diamond, C., *The Realistic Spirit: Wittgenstein, Philosophy, and the Mind* (Cambridge, Mass.: MIT Press, 1991).

'Throwing away the ladder', in C. Diamond, *The Realistic Spirit*, pp. 179–204.

'Truth before Tarski: After Sluga, after Ricketts, after Geach, after Goldfarb, Hylton, Floyd and van Heijenoort', in E. Reck (ed.), *From Frege to Wittgenstein*, pp. 252–79.

'What does a concept-script do?', in C. Diamond, *The Realistic Spirit*, pp. 115–44.

Dodd, J., and Hornsby, J., 'The identity theory of truth: Reply to Baldwin', *Mind*, 101 (1992), pp. 319–22.

Dreben, B., and Floyd, J., 'Tautology: How not to use a word', *Synthese*, 87 (1991), pp. 23–49.

Dreben, B., and van Heijenoort, J., 'Introductory note to 1929, 1930, and 1930a', in K. Gödel, *Collected Works*, vol. I, ed. S. Feferman *et al.* (New York: Oxford University Press, 1986), pp. 44–59.

Dummett, M., 'The Context Principle: Centre of Frege's philosophy', in Max Ingolf and Werner Stelzner (eds.), *Logik und Mathematik: Frege Kolloquium, Jena 1993* (Berlin: De Gruyter, 1995).

Frege and Other Philosophers (Oxford: Clarendon Press, 1991).

'Frege and the consistency of mathematical theories', in M. Dummett, *Frege and Other Philosophers*, pp. 1–16.

'Frege and the paradox of analysis', in M. Dummett, *Frege and Other Philosophers*, pp. 17–52.

'Frege on functions', in M. Dummett, *Truth and Other Enigmas*, pp. 74–86.

Frege: Philosophy of Language (London: Duckworth, 1973; 2nd edn 1981).

Frege: Philosophy of Language (New York: Harper and Row, 1973).

Frege: Philosophy of Mathematics (London: Duckworth, 1991).

'Frege's distinction between sense and reference', in M. Dummett, *Truth and Other Enigmas*, pp. 116–44.

'Frege's *Kernsätze zur Logik*', in M. Dummett, *Frege and Other Philosophers*, pp. 65–79.

'Intellectual autobiography', in R. E. Auxier and L. E. Hahn (eds.), *The Philosophy of Michael Dummett* (Chicago: Open Court, 2007).

The Interpretation of Frege's Philosophy (London: Duckworth, 1981).

The Logical Basis of Metaphysics (Cambridge, Mass.: Harvard University Press, 1991).

'Of what kind of a thing is truth a property?', in S. Blackburn and K. Simmons (eds.), *Truth* (Oxford: Oxford University Press, 1999), pp. 264–81.

The Seas of Language (Oxford: Oxford University Press, 1993).
Truth and Other Enigmas (London: Duckworth, 1978).
'Truth', in M. Dummett, *Truth and Other Enigmas*, pp. 1–24.
'What is a theory of meaning?' parts I and II, in M. Dummett, *The Seas of Language*, pp. 1–93.
Edwards, H., 'The genesis of ideal theory', *Archive for the History of the Exact Sciences*, 23 (1980), pp. 321–78.
Enriques, F., *The Historical Development of Logic*, trans. Jerome Rosenthal (New York: Holt, Rinehart and Winston, 1929).
Etchemendy, J., *The Concept of Logical Consequence* (Cambridge, Mass.: Harvard University Press, 1990).
Evans, G., *Collected Papers* (Oxford: Oxford University Press, 1985).
'Understanding demonstratives', in G. Evans, *Collected Papers*, pp. 291–321.
The Varieties of Reference (Oxford: Oxford University Press, 1982).
Ewald, W. (ed.), *From Kant to Hilbert*, 2 vols. (Oxford: Oxford University Press, 1996).
Ewald, W., and Sieg, W. (eds.), *David Hilbert's Lectures on the Foundations of Logic and Arithmetic, 1917–1933, Hilbert's Lectures on the Foundations of Mathematics and Physics*, vol. III (Heidelberg: Springer, 2008).
Fernández Moreno, L., 'Die Undefinierbarkeit der Wahrheit bei Frege', *Dialectica*, 50 (1996), pp. 25–35.
Fine, K., 'Vagueness, truth and logic', *Synthese*, 30 (1975), pp. 265–300.
Freudenthal, H., 'The impact of von Staudt's foundations of geometry', in P. Plaumann and K. Strambach (eds.), *Geometry – von Staudt's Point of View* (Dordrecht: Reidel, 1981).
Friedman, M., *Kant and the Exact Sciences* (Cambridge, Mass.: Harvard University Press, 1992).
Furth, M., 'Editor's Introduction', in Frege, *The Basic Laws of Arithmetic: Exposition of the System*, ed. and trans. M. Furth (Berkeley and Los Angeles: University of California Press, 1964).
Gabriel, G., 'Einige Einseitigkeiten der Fregeschen Logikbegriffs', in M. Schirn (ed.), *Studien zu Frege II*.
'Frege als Neukantianer', *Kant-Studien*, 77 (1986), pp. 84–101.
'Frege, Lotze, and the continental roots of early analytic philosophy', in E. Reck (ed.), *From Frege to Wittgenstein*, pp. 39–52.
'Frege's "Epistemology in disguise"', in M. Schirn (ed.), *Frege: Importance and Legacy* (Berlin: de Gruyter, 1996).
'Logik und Metaphysik in Freges Philosophie der Mathematik', in U. Dathe (ed.), *Gottlob Frege: Werk und Wirkung* (Paderborn, 2000), pp. 25–38.

'Objektivität: Logik und Erkenntnistheorie bei Lotze und Frege', Editor's Introduction to H. Lotze, *Logik: Drittes Buch. Vom Erkennen (Methodologie)* (Hamburg: Meiner, 1989).
Gahringer, R., 'Intensional conjunction', *Mind*, 79 (1970), pp. 259–60.
Gauss, K. F., *Disquisitiones Arithmeticae*, trans. A. A. Clark (New Haven, Conn.: Yale University Press, 1965).
Geach, P. T., 'Critical notice of Michael Dummett, *Frege: Philosophy of Language*', *Mind*, 85 (1976), pp. 436–49.
'Editor's Preface', in P. T. Geach (ed.), *Wittgenstein's Lectures on Philosophical Psychology, 1946–47* (Chicago: University of Chicago Press, 1988), pp. xi–xv.
'Frege', in Elizabeth Anscombe and Peter Geach, *Three Philosophers* (Oxford: Blackwell, 1961), pp. 127–62.
'Names and identity', in S. Guttenplan (ed.), *Mind and Language* (Oxford: Clarendon Press, 1975), pp. 139–58.
'A philosophical autobiography', in H. A. Lewis (ed.), *Peter Geach: Philosophical Encounters* (Dordrecht: Kluwer, 1991), pp. 1–25.
'Preface', Gottlob Frege, *Logical Investigations*, ed. P. T. Geach (Oxford: Blackwell, 1977).
'Quine on classes and properties', in P. T. Geach, *Logic Matters* (Oxford: Blackwell, 1972), pp. 222–6.
'Saying and showing in Frege and Wittgenstein', *Acta Philosophica Fennica*, 28 (1976), pp. 54–70.
'Subject and predicate', *Mind*, 59 (1950), pp. 461–82.
'Truth and God', *Aristotelian Society Supplementary Volume*, 56 (1982), pp. 83–97.
Gentzen, G., *The Collected Papers of Gerhard Gentzen*, ed. M. E. Szabo (Amsterdam: North-Holland, 1969).
Goldfarb, W., 'Frege's conception of logic', in J. Floyd and S. Shieh (eds.), *Future Pasts: The Analytic Tradition in Twentieth-Century Philosophy* (New York: Oxford University Press, 2001), pp. 25–41.
'Logic in the twenties', *Journal of Symbolic Logic*, 44 (1979), pp. 351–68.
'Metaphysics and nonsense: On Cora Diamond's The Realistic Spirit', *Journal of Philosophical Research*, 22 (1997), pp. 57–73.
'Wittgenstein's understanding of Frege: The pre-Tractarian evidence', in E. Reck (ed.), *From Frege to Wittgenstein*, pp. 185–200.
Gray, J., *Worlds out of Nothing* (London: Springer, 2007).
Gustafsson, M., *Entangled Sense: An Inquiry into the Philosophical Significance of Meaning and Rules* (Uppsala: Universitetstryckeriet, 2000).
Haack, S., *Philosophy of Logics* (Cambridge: Cambridge University Press, 1978).

Haaparanta, L. and Hintikka, J. (eds.), *Frege Synthesized* (Dordrecht: Reidel, 1986).

Hacker, P. M. S., 'Frege and the early Wittgenstein', in P. M. S. Hacker, *Wittgenstein: Connections and Controversies* (Oxford: Clarendon Press, 2001), pp. 191–219.

'Frege and the later Wittgenstein', in P. M. S. Hacker, *Wittgenstein: Connections and Controversies* (Oxford: Clarendon Press, 2001), pp. 219–42.

Hájek, P., 'On arithmetic in the Cantor–Lukasiewicz fuzzy set theory', *Archive for Mathematical Logic*, 44 (2005), pp. 763–82.

Hallett, M., *Cantorian Set Theory and Limitation of Size* (Oxford: Oxford University Press, 1984).

'The "purity of method" in Hilbert's *Grundlagen der Geometrie*', in P. Mancosu (ed.), *The Philosophy of Mathematical Practice* (Oxford: Clarendon Press, 2008), pp. 198–255.

'Putnam and the Skolem paradox', in P. Clark and S. Read (eds.), *Reading Putnam*, pp. 66–97.

Heath, T. L., *The Thirteen Books of Euclid's Elements*, 3 vols, 2nd edn (Cambridge: Cambridge University Press, 1925).

Heck, R., 'Definition by induction in Frege's *Grundgesetze der Arithmetik*', in W. Demopoulos (ed.), *Frege's Philosophy of Mathematics*, pp. 295–333.

'The finite and the infinite in Frege's *Grundgesetze der Arithmetik*', in M. Schirn (ed.), *Philosophy of Mathematics Today* (Oxford: Oxford University Press, 1998), pp. 429–66.

'Grundgesetze der Arithmetik I §10', *Philosophia Mathematica*, 7 (1999), pp. 258–92.

'Grundgesetze der Arithmetik I §§29–32', *Notre Dame Journal of Formal Logic,* 38 (1998), pp. 437–74.

'The Julius Caesar objection', in R. Heck (ed.), *Language, Thought, and Logic: Essays in Honour of Michael Dummett* (Oxford: Oxford University Press, 1997), pp. 273–308.

'The sense of communication', *Mind*, 104 (1995), pp. 79–106.

Heidelberger, H., 'The indispensability of truth', *American Philosophical Quarterly*, 5 (1968), pp. 212–17.

Hilbert, D., *David Hilbert's Lectures on the Foundations of Geometry, 1891–1902,* ed. M. Hallett and U. Majer (Heidelberg: Springer, 2004).

Elemente der Euklidischen Geometrie: Ausarbeitung by Hans von Schaper of the lecture notes *Grundlagen der Euklidischen Geometrie*, 1899, Niedersächsische Staats- und Universitätsbibliothek, Göttingen, and the Mathematisches Institut of the Georg-August Universität, Göttingen; first published in M. Hallett and U. Majer (eds.), *Lectures*, pp. 302–406.

The Foundations of Geometry, trans. E. J. Townsend (La Salle, Ill.: Open Court, 1902).

Grundlagen der Euklidischen Geometrie, lecture notes for a course held in the Wintersemester of 1898/9 at the Georg-August Universität, Göttingen. Niedersächsische Staats- und Universitätsbibliothek, Göttingen; first published in M. Hallett and U. Majer (eds.), *Lectures*, pp. 221–301.

Grundlagen der Geometrie, lecture notes for a course to have been held in the Wintersemester of 1893/4 at the University of Königsberg. Niedersächsische Staats- und Universitätsbibliothek, Göttingen; first published in M. Hallett and U. Majer (eds.), *Lectures*, pp. 72–144.

'Grundlagen der Geometrie', in *Festschrift zur Feier der Enthüllung des Gauss-Weber-Denkmals in Göttingen* (Leipzig: B. G. Teubner, 1899), republished as ch. 5 in M. Hallett and U. Majer (eds.), *Lectures*.

Grundlagen der Geometrie, *Ausarbeitung* by August Adler for lectures in the Sommersemester of 1902 at the Georg-August Universität, Göttingen, library of the Mathematisches Institut; first published as ch. 6 in M. Hallett and U. Majer (eds.), *Lectures*.

Grundlagen der Geometrie, revised 2 edn (Leipzig and Berlin: B. G. Teubner, 1903).

Grundlagen der Mathematik, lecture notes for a course held in the Wintersemester of 1921/2 at the Georg-August Universität, Göttingen, library of the Mathematisches Institut; to be published in W. Ewald and W. Sieg (eds.), *David Hilbert's Lectures*.

'Mathematische probleme', *Nachrichten von der königlichen Gesellschaft der Wissenschaften zu Göttingen, mathematisch-physikalische Klasse* (1900), pp. 253–96.

'Les Principes fondamentaux de la géométrie', *Annales scientifiques de l'École Normale Supérieure*, 3(17) (1900), pp. 103–209.

'Über den Satz von der Gleichheit der Basiswinkel im gleichschenkli-chen Dreieck', *Proceedings of the London Mathematical Society*, 35 (1902/3), pp. 50–67.

'Über den Zahlbegriff', *Jahresbericht der deutschen Mathematiker-Vereinigung*, 8 (1900), pp. 180–5.

Hilbert, D., and Bernays, P., *Grundlagen der Mathematik* (Berlin: Springer, 1935–9).

Hintikka, J., 'On the development of the model-theoretic viewpoint in logical theory', *Synthese*, 77 (1988), pp. 1–36.

Hodes, H., 'The composition of Fregean thoughts', *Philosophical Studies*, 41 (1982), pp. 161–78.

Holton, R., 'Minimalism and truth-value gaps', *Philosophical Studies*, 97 (2000), pp. 137–68.

Hovens, F., 'Lotze and Frege: The dating of the "Kernsätze"', *History and Philosophy of Logic*, 18 (1997), pp. 17–31.

Hugly, P., 'Ineffability in Frege's logic', *Philosophical Studies*, 24 (1973), pp. 227–44.

Hylton, P., 'Functions, operations, and sense in Wittgenstein's *Tractatus*', in W. W. Tait (ed.), *Early Analytic Philosophy*, pp. 91–105; reprinted in P. Hylton, *Propositions, Functions, and Analysis: Selected Essays on Russell's Philosophy* (Oxford: Clarendon Press, 2005), pp. 138–52.

'Functions and propositional functions in *Principia Mathematica*', in A. Irvine and G. Wedeking (eds.), *Russell and Analytic Philosophy* (Toronto: University of Toronto Press, 1994), pp. 342–60; reprinted in Hylton, *Propositions, Functions, and Analysis: Selected Essays on Russell's Philosophy* (Oxford: Oxford University Press, 2005).

Russell, Idealism, and the Emergence of Analytic Philosophy (Oxford: Oxford University Press, 1990).

'The Vicious-Circle Principle: Comments on Philippe de Rouilhan', *Philosophical Studies*, 65 (1992), pp. 183–91; reprinted in Hylton, *Propositions, Functions, and Analysis* (Oxford: Oxford University Press, 2005).

Jeshion, R., 'Frege's notion of self-evidence', *Mind*, 110 (2001), pp. 937–76.

Jourdain, P. E. B., 'Gottlob Frege', a chapter from P. E. B. Jourdain, 'The development of the theories of mathematical logic and the principles of mathematics', reprinted as the Appendix to Frege, *PMC*, pp. 179–206.

Kant, I., *Critique of Pure Reason*, trans. P. Guyer and A. Wood (Cambridge: Cambridge University Press, 1998).

Logik, trans. as *Lectures on Logic*, trans. J. M. Young (Cambridge: Cambridge University Press, 1992).

Kaplan, D., 'Demonstratives', in J. Almog, J. Perry and H, Wettstein (eds.), *Themes from Kaplan* (New York: Oxford University Press, 1989), pp. 481–614.

Kemp, G., 'Frege's sharpness requirement', *Philosophical Quarterly*, 46 (1996), pp. 168–84.

Kenny, A., *Frege* (London: Penguin, 1995).

'The Ghost of the *Tractatus*', in G. Vesey (ed.), *Understanding Wittgenstein* (London: Macmillan, 1974), pp. 1–13.

'Oratio Obliqua', *Aristotelian Society Supplementary Volume,* 37 (1963), pp. 127–46.

Kirwan, C., *Logic and Argument* (London: Duckworth, 1978).

Kitcher, P., 'Frege's epistemology', *Philosophical Review*, 88 (1979), pp. 235–62.

Klein, F., *Elementary Mathematics from an Advanced Standpoint: Geometry* (New York: Dover, 1941).

Klemke, E. D. (ed.), *Essays on Frege* (Urbana, Ill.: University of Illinois Press, 1968).

Kluge, E.-H. W. (ed.), *On the Foundations of Geometry and Formal Theories of Arithmetic* (New Haven and London: Yale University Press, 1971).

Kneale, W., and Kneale, M., *The Development of Logic* (Oxford: Clarendon Press, 1962).

Kreiser, L., *Frege: Leben, Werk, Zeit* (Hamburg: Meiner, 2001).

Kremer, M., 'The argument of "On denoting"', *Philosophical Review*, 103 (1994), pp. 249–97.

'The cardinal problem of philosophy', in A. Crary (ed.), *Wittgenstein and the Moral Life: Essays in Honor of Cora Diamond* (Cambridge, Mass.: MIT Press, 2007), pp. 143–76.

'Judgment and truth in Frege', *Journal of the History of Philosophy*, 38 (2000), pp. 549–81.

'The purpose of Tractarian nonsense', *Noûs,* 35 (2001), pp. 39–73.

Kripke, S., *Naming and Necessity* (Cambridge, Mass.: Harvard University Press, 1980).

'A puzzle about belief', in A. Margalit (ed.), *Meaning and Use* (Dordrecht: Reidel, 1979).

Kuczmarski, R. J., and Flegal, K. M., 'Criteria for definition of overweight in transition: background and recommendations for the United States', *American Journal of Clinical Nutrition*, 72 (2000), pp. 1074–81.

Landini, G., 'The ins and outs of Frege's way out', *Philosophia Mathematica*, 14 (2006), pp. 1–25.

Lehmann, S., 'Strict Fregean free logic', *Journal of Philosophical Logic*, 23 (1994), pp. 307–36.

Levine, J., 'Analysis and decomposition in Frege and Russell', *Philosophical Quarterly,* 52 (2002), pp. 195–216.

Linnebo, Ø., 'Frege's proof of referentiality', *Notre Dame Journal of Formal Logic,* 45 (2004), pp. 73–98.

Linsky, L., 'Frege and Russell on vacuous singular terms', in M. Schirn (ed.), *Studien zu Frege/Studies on Frege*, vol. III (Stuttgart: Frommann, 1976), pp. 97–115.

Names and Descriptions (Chicago: University of Chicago Press, 1977).

Long, A. A., and Sedley, D. N. (eds.), *The Hellenistic Philosophers* (Cambridge: Cambridge University Press, 1987).

Macbeth, D., 'Frege and early Wittgenstein on logic and language', in E. Reck (ed.), *From Frege to Wittgenstein*, pp. 201–27.

MacFarlane, J., 'Frege, Kant, and the logic in logicism', *Philosophical Review*, 111 (2002), pp. 25–65.

Makin, G., *The Metaphysicians of Meaning: Russell and Frege on Sense and Denotation* (London: Routledge, 2000).

Mancosu, P. (ed.), *From Brouwer to Hilbert: The Debate in the Foundations of Mathematics in the 1920s* (Oxford: Oxford University Press, 1998).

Manders, K., 'The Euclidean diagram', in P. Mancosu (ed.), *The Philosophy of Mathematical Practice* (Oxford: Oxford University Press, 2008), pp. 112–83.

Massey, G. J., *Understanding Symbolic Logic* (New York: Harper and Row, 1970).

Mates, B., *Elementary Logic*, 2nd edn (Oxford: Oxford University Press, 1972).

May, R., 'Frege on identity statements', in C. Cecchetto *et al*. (eds.), *Semantic Interfaces: Reference, Anaphora, and Aspect* (Stanford: CSLI, 2001).

McDowell, J., *Mind and World* (Cambridge, Mass.: Harvard University Press, 1996).

'On the sense and reference of proper names', *Mind*, 86 (1977), pp. 159–85.

'Truth-value gaps', in J. McDowell, *Meaning, Knowledge, and Reality* (Cambridge, Mass.: Harvard University Press, 1998), pp. 199–213.

Mendelsohn, R., 'Frege's *Begriffsschrift* theory of identity', *Journal of the History of Philosophy*, 20 (1982), pp. 279–99.

Milne, P., 'Existence, freedom, identity, and the logic of abstractionist realism', *Mind*, 116 (2007), pp. 23–53.

Monna, A. F., *Dirichlet's Principle* (Utrecht: Oostheok, Scheltema and Holkema, 1975).

Nagel, E., *Teleology Revisited* (New York: Columbia University Press, 1982).

Noonan, H., *Frege* (Oxford: Polity Press, 2001).

'The "Gray's Elegy" argument – and others', in R. Monk and A. Palmer (eds.), *Bertrand Russell and the Origins of Analytic Philosophy* (Bristol: Thoemmes, 1996), pp. 65–102.

Pakaluk, M., 'The interpretation of Russell's "Gray's *Elegy*" Argument', in A. D. Irvine and G. A. Wedeking (eds.), *Russell and Analytic Philosophy* (Toronto: University of Toronto Press, 1993), pp. 37–65.

Parsons, C., *Mathematics in Philosophy* (Ithaca, N.Y.: Cornell University Press, 1983).

'Objects and logic', *Monist*, 65 (1982), pp. 491–516.

'Review article: Gottlob Frege, *Wissenschaftlicher Briefswechsel*', *Synthese*, 52 (1982), pp. 325–43.

Parsons, T., 'On the consistency of the first-order portion of Frege's logical system', *Notre Dame Journal of Formal Logic*, 28 (1987), pp. 161–8.

'What do quotation marks name? Frege's theories of quotations and that-clauses', *Philosophical Studies*, 42 (1982), pp. 315–28.

'Why Frege should not have said "The concept horse is not a concept"', *History of Phiosophy Quarterly*, 3 (1986), pp. 449–65.

Pasch, M., *Vorlesungen über neuere Geometrie* (Leipzig: Teubner, 1882).

Peirce, C. S., 'Upon logical comprehension and extension', in E. C. Moore *et al.* (eds.), *Writings of Charles S. Peirce: A Chronological Edition*, vol. II, *1862–1871* (Bloomington: Indiana University Press, 1984).

Perry, J., 'Frege on demonstratives', *Philosophical Review*, 86 (1977), pp. 474–97.

'Frege on identity, cognitive value, and subject-matter', in A. Newen *et al.* (eds.), *Building on Frege: New Essays on Sense, Content and Concept* (Stanford: CSLI, 2001).

Picardi, E., 'Frege on definition and logical proof', in *Atti del Congresso Temi e Prospettive della logica e della filosofia della scienza contemporanee. Cesean 7–10 gennaio 1987*, vol. I (Bologna: CLUEB, 1988), pp. 227–30.

'Kerry und Frege über Begriff und Gegenstand', *History and Philosophy of Logic*, 15 (1994), pp. 9–32.

Poincaré, H., *The Foundations of Science* (New York: The Science Press, 1913).

'Les géométries non euclidiennes', *Revue général des sciences pures et appliqués*, 2 (1891), pp. 769–74.

'La logique de l'infini', *Revue de Métaphysique et de Morale*, 17 (1909), pp. 461–82.

Science and Hypothesis (Walter Scott, 1905; reprinted New York: Dover Publications, 1952).

La science et l'hypothèse (Paris: Ernst Flammarion, Paris, 1902).

Science and Method, trans. F. Maitland (New York: Dover Publications, 1952).

Science et méthode (Paris: Ernst Flammarion, 1908).

La valeur de la science (Paris: Ernst Flammarion, 1905).

Potter, M., *Reason's Nearest Kin: Philosophies of Arithmetic from Kant to Carnap* (Oxford: Oxford University Press, 2000).

Wittgenstein's Notes on Logic (Oxford: Oxford University Press, 2009).

Potts, T. C., *Structures and Categories for the Representation of Meaning* (Cambridge: Cambridge University Press, 1994).

Prior, A. N., 'Is the concept of referential opacity really necessary?', *Acta Philosophica Fennica*, 16 (1963), pp. 189–99.

'Oratio Obliqua', *Aristotelian Society Supplementary Volume*, 37 (1963), pp. 115–26.

Proops, I., 'The early Wittgenstein on logical assertion', *Philosophical Topics*, 25 (1997), pp. 121–44.

Quine, W. V., *Elementary Logic* (Boston: Ginn, 1941).

From a Logical Point of View (Cambridge, Mass.: Harvard University Press, 1953).

Methods of Logic (New York: Holt, 1950).

Ontological Relativity and Other Essays (New York: Columbia University Press, 1966).

The Philosophy of Logic, 2nd edn (Cambridge, Mass.: Harvard University Press, 1986).

Pursuit of Truth, rev. edn (Cambridge, Mass.: Harvard University Press, 1992).

Word and Object (Cambridge, Mass.: MIT Press, 1960).

Ramsey, F. P., 'Universals', in F. P. Ramsey, *The Foundations of Mathematics and Other Logical Essays*, ed. R. B. Braithwaite (London: Kegan Paul, 1931).

Reck, E., 'Frege's influence on Wittgenstein: Reversing metaphysics via the context principle', in W. W. Tait (ed.), *Early Analytic Philosophy*, pp. 123–85.

'From Frege and Russell to Carnap: logic and logicism in the 1920s', in S. Awodey and C. Klein (eds.), *Carnap Brought Home: The View from Jena* (Chicago: Open Court, 2009).

(ed.), *From Frege to Wittgenstein: Perspectives on Early Analytic Philosophy* (Oxford: Oxford University Press, 2002).

'Wittgenstein's "great debt" to Frege: Biographical traces and philosophical themes', in E. Reck (ed.), *From Frege to Wittgenstein*, pp. 3–38.

Reck, E., and Awodey, S. (eds.), *Frege's Lectures on Logic: Carnap's Student Notes, 1910–14* (Chicago: Open Court: 2004).

Reid, L. W., *The Elements of the Theory of Algebraic Numbers* (New York: MacMillan, 1910).

Resnik, M., *Frege and the Philosophy of Mathematics* (Ithaca, N.Y.: Cornell University Press, 1980).

'Frege's theory of incomplete entities', *Philosophy of Scence,* 32 (1965), pp. 329–41.

Ricketts, T., 'Frege's 1906 foray into metalogic', *Philosophical Topics,* 25 (1997), pp. 169–188.

'Frege, the *Tractatus*, and the logocentric predicament', *Noûs,* 19 (1985), pp. 3–15.

'Generality, sense, and meaning in Frege', *Pacific Philosophical Quarterly*, 67 (1986), pp. 172–95.

'Logic and truth in Frege', *Aristotelian Society Supplementary Volume,* 70 (1996), pp. 121–40.

'Objectivity and objecthood: Frege's metaphysics of judgement', in L. Haaparanta and J. Hintikka (eds.), *Frege Synthesized,* pp. 65–95.

'Pictures, logic, and the limits of sense in Wittgenstein's *Tractatus*', in H. Sluga and D. G. Stern (eds.), *The Cambridge Companion to Wittgenstein* (Cambridge: Cambridge University Press, 1996), pp. 59–99.

'Quantification, sentences, and truth-values', *Manuscrito*, 26 (2003), pp. 389–424.

'Truth and propositional unity in early Russell', in Juliet Floyd and Sanford Shieh (eds.), *Future Pasts: The Analytic Tradition in 20th Century Philosophy* (Oxford: Oxford University Press, 2001), pp. 101–23.

'Truth-values and courses-of-values in Frege's *Grundgesetze*', in W. W. Tait (ed.), *Early Analytic Philosophy: Frege, Russell, Wittgenstein* (Chicago: Open Court, 1997), pp. 187–211.

'Wittgenstein against Frege and Russell', in Reck (ed.), *From Frege to Wittgenstein*, pp. 227–51.

Rumfitt, I., 'Frege's theory of predication: An elaboration and defense, with some new applications', *Philosophical Review*, 103 (1994), pp. 599–637.

Russell, B., *An Essay on the Foundations of Geometry* (New York: Dover, 1956).

The Collected Papers of Bertrand Russell, ed. K. Blackwell *et al.* (London: Routledge, 1983–).

Introduction to Mathematical Philosophy (London: Allen and Unwin, 1919).

'Knowledge by acquaintance and knowledge by description', in B. Russell, *Mysticism and Logic*, pp. 152–67.

Logic and Knowledge: Essays 1901–1950, ed. R. C. Marsh (London: Allen and Unwin, 1956).

'Mathematical logic as based on the theory of types', *American Journal of Mathematics*, 30 (1908), pp. 222–62.

Mysticism and Logic (London: Allen and Unwin, 1917).

My Philosophical Development (London: Allen and Unwin, 1959).

'On denoting', in B. Russell, *Logic and Knowledge*, pp. 41–56.

'On the notion of order', in B. Russell, *Collected Papers,* vol. III, ed. G. H. Moore (London: Routledge, 1993).

The Principles of Mathematics (Cambridge: Cambridge University Press, 1903; 2nd edn London: Allen and Unwin, 1937).

The Problems of Philosophy (Oxford: Oxford University Press, 1959).

Theory of Knowledge, vol. VIII of his *Collected Papers*, ed. E. H. Eames (London: Routledge, 1983).

Salmon, N., *Frege's Puzzle* (Cambridge, Mass.: MIT Press, 1986).

Reference and Essence (Princeton: Princeton University Press, 1981).

'Reference and information content: Names and descriptions', in D. Gabbay and R. Guenthner (eds.), *Handbook of Philosophical Logic* (Dordrecht: Reidel, 1990), pp. 409–61.

Schilpp, P. A. (ed.), *The Philosophy of Bertrand Russell* (Evanston, Ill.: Northwestern University, 1944).

The Philosophy of Rudolf Carnap (Chicago, Ill.: Open Court, 1963).

Schlotter, S., 'Frege's anonymous opponent in Die Verneinung', *History and Philosophy of Logic*, 27 (2006), pp. 43–58.

Scholem, G., *Walter Benjamin: The Story of a Friendship* (London: Faber, 1982).

Schubert, H., *Mathematical Essays and Recreations* (Chicago, Ill.: Open Court, 1910).

Searle, J., 'Proper names', *Mind*, 67 (1958), pp. 166–73.

Sellars, W., 'Abstract entities', in *Philosophical Perspectives* (Springfield: Charles Thomas, 1967).

Shapiro, S., and Weir, A., '"Neo-Logicist" logic is not innocent', *Philosophia Mathematica*, 8 (2000), pp. 160–89.

Sher, G., *The Bounds of Logic: A Generalized Viewpoint* (Cambridge, Mass.: MIT Press, 1991).

Shieh, S., 'On Interpreting Frege on truth and logic', in E. Reck (ed.), *From Frege to Wittgenstein*, pp. 96–124.

Sigwart, C., *Logic*, vol. I, trans. H. Dendy (London: Swan Sonnenschein, 1895).

Simons, P., 'Frege's theory of real numbers', *History and Philosophy of Logic*, 8 (1987), pp. 25–44.

'The next best thing to sense in *Begriffsschrift*', in J. Biro and P. Kotatko (eds.), *Frege: Sense and Reference One Hundred Years Later* (Amsterdam: Kluwer, 1995), pp. 129–40.

Sluga, H., 'Frege on meaning,' in H. J. Glock (ed.), *The Rise of Analytic Philosophy* (Oxford: Blackwell, 1997).

'Frege on the indefinability of truth', in E. Reck (ed.), *From Frege to Wittgenstein*, pp. 75–96.

Gottlob Frege (London: Routledge and Kegan Paul, 1980).

(ed.), *The Philosophy of Frege*, 4 vols. (New York: Garland, 1993).

Smiley, T., 'Sense without denotation', *Analysis*, 20 (1960), pp. 125–35.

Smith, D. E., *A Source Book in Mathematics* (New York: Dover, 1985).

Smith, H. J. S., 'On some of the methods at present in use in pure geometry', in H. J. S. Smith, *Collected Papers*, vol. I (New York: Chelsea, 1965).

Soames, S., *Beyond Rigidity: The Unfinished Semantic Agenda of Naming and Necessity* (New York: Oxford University Press, 2002).

'Truth, meaning, and understanding', *Philosophical Studies*, 65 (1992), pp. 17–35.

Stanley, J., 'Truth and metatheory in Frege', *Pacific Philosophical Quarterly*, 77 (1996), pp. 45–70.

Staudt, K. von, *Beiträge zur Philosophie der Lage* (Nuremberg: Bauer and Raspe, 1856).

Geometrie der Lage (Nuremberg: Bauer and Raspe, 1847).

Stein, H., 'Logos, logic and logistiké', in W. Aspray and P. Kitcher (eds.), *History and Philosophy of Modern Mathematics* (Minneapolis: University of Minnesota Press, 1988).

Steiner, M., *Mathematical Knowledge* (Ithaca, N.Y.: Cornell University Press, 1975).

Stenius, E., 'The sentence as a function of its constituents in Frege and in the Tractatus', *Acta Philosophica Fennica*, 28 (1976), pp. 71–84.

Stevenson, L., 'Frege's two definitions of quantification', *Philosophical Quarterly*, 23 (1973), pp. 207–223.

Stolz, B.,'Die geometrische Bedeutung der complexen Elemente in der analytischen Geometrie', *Mathematische Annalen*, 4 (1871), pp. 416–41.

Sullivan, D., 'Frege on the statement of number', *Philosophy and Phenomenological Research*, 50 (1990), pp. 595–603.

Sullivan, P., 'Frege's logic', in D. M. Gabbay and J. Woods (eds.), *Handbook of the History of Logic*, vol. III (Amsterdam: North-Holland, 2004), pp. 659–750.

'The functional model of sentential complexity', *Journal of Philosophical Logic*, 21 (1992), pp. 91–108.

'The sense of a name of a truth-value', *Philosophical Quarterly*, 44 (1994), pp. 476–81.

Tait, W. W. (ed.), *Early Analytic Philosophy: Frege, Russell, Wittgenstein* (Chicago, Ill.: Open Court, 1997).

'Frege versus Cantor and Dedekind: On the concept of number', in W. W. Tait (ed.), *Early Analytic Philosophy*, pp. 213–48.

Tappenden, J., 'Extending knowledge and "fruitful concepts"', *Noûs*, 29 (1995), pp. 427–67.

'Frege on axioms, indirect proof, and independence arguments in geometry: Did Frege reject independence arguments?', *Notre Dame Journal of Formal Logic*, 41 (2000), pp. 271–315.

'Geometry and generality in Frege', *Synthese*, 102 (1995), pp. 319–61.

'Metatheory and mathematical practice in Frege', *Philosophical Topics*, 25 (1997), pp. 213–64.

Tappolet, C., 'Truth pluralism and many-valued logics: A reply to Beall', *Philosophical Quarterly*, 50 (2000), pp. 382–5.

Tarski, A., 'The concept of truth in formalized langauges', in J. Corcoran (ed.), *Logic, Semantics, and Metamathematics* (Indianapolis: Hackett, 1958), pp. 152–278.

'On the concept of logical consequence', in A. Tarski, *Logic, Semantics, Metamathematics* (Oxford: Oxford University Press, 1956).
Taschek, W. W., 'Frege's puzzle, sense, and information content', *Mind*, 101 (1992), pp. 767–91.
'On ascribing beliefs: Content and context', *Journal of Philosophy*, 95 (1998), pp. 323–53.
'Truth, assertion, and the horizontal: Frege on "the essence of logic"', *Mind*, 117 (2008), pp. 375–401.
Tennant, N., 'A general theory of abstraction operators', *Philosophical Quarterly*, 54 (2004), pp. 105–33.
'Frege's content-principle and relevant deducibility', *Journal of Philosophical Logic*, 32 (2003), pp. 245–58.
Natural Logic (Edinburgh: Edinburgh University Press, 1978).
Thompson, B., 'Why is conjunctive simplification invalid?', *Notre Dame Journal of Formal Logic*, 32 (1991), pp. 248–54.
van Fraassen, B., 'Singular terms, truth-value gaps, and free logic', *Journal of Philosophy*, 63 (1966), pp. 481–95.
van Heijenoort, J., *Frege and Gödel: Two Fundamental Texts in Mathematical Logic* (Cambridge, Mass.: Harvard University Press, 1970).
'Frege on sense identity', *Journal of Philosophical Logic*, 6 (1977), pp. 103–8.
(ed.), *From Frege to Gödel: A Source Book in Mathematical Logic* (Cambridge, Mass.: Harvard University Press, 1967).
'Logic as calculus and logic as language', *Synthese*, 17 (1967), pp. 324–30.
Veblen, O. and Young., J. W., *Projective Geometry* (Boston: Ginn, 1910).
Veraart, A., 'Geschichte des wissenschaftlichen Nachlasses Gottlob Freges und seiner Edition. Mit einem Katalog des ursprünglichen Bestands der nachgelassenen Schriften Freges', in M. Schirn (ed.), *Studien zu Frege I: Logik und Philosophie der Mathematik* (Stuttgart: Frommann, 1976).
Waismann, F., 'Theses', in B. F. McGuinness (ed.), *Ludwig Wittgenstein and the Vienna Circle*, trans. J. Schulte and B. F. McGuinness (Oxford: Blackwell, 1979), pp. 233–61.
Wehmeier, K. and Schmidt am Busch, H.-C., 'The quest for Frege's Nachlass', in M. Beaney and E. H. Reck (eds.), *Gottlob Frege: Critical Assessments* (Routledge, 2004), pp. 55–68.
Weiner, J., 'Frege and the linguistic turn', *Philosophical Topics*, 25 (1997), pp. 265–88.
Frege Explained (Chicago: Open Court, 2004).
Frege in Perspective (Ithaca, N.Y.: Cornell University Press, 1990).
'Has Frege a philosophy of language?', in W. W. Tait (ed.), *Early Analytic Philosophy*, pp. 249–72.

'Section 31 revisited: Frege's elucidations', in E. Reck (ed.), *From Frege to Wittgenstein*, pp. 149–82.

'Semantic descent', *Mind*, 114 (2005), pp. 321–54.

'Theory and elucidation: The end of the age of innocence', in J. Floyd and S. Shieh (eds.), *Future Pasts: The Analytic Tradition in Twentieth-Century Philosophy* (New York and Oxford: Oxford University Press, 2001), pp. 43–65.

'What's in a numeral? Frege's answer', *Mind*, 116 (2007), pp. 677–716.

Weitzman, L., 'Frege on the individuation of thoughts', *Dialogue*, 26 (1997), pp. 563–74.

Whitehead, A. N., and Russell, B., *Principia Mathematica*, 3 vols. (Cambridge: Cambridge University Press, 1910–13).

Wiggins, D., 'Meaning, truth-conditions, proposition: Frege's doctrine of sense retrieved, resumed, and redeployed in the light of certain recent criticisms', *Dialectica*, 46 (1992), pp. 61–90.

Wilson, M., 'Frege: The Royal Road from geometry', in W. Demopoulos (ed.), *Frege's Philosophy of Mathematics*, pp. 108–49.

Wittgenstein, L., *The Blue and Brown Books* (Oxford: Blackwell, 1964).

Culture and Value, ed. G. H. von Wright, trans. P. Winch (Oxford: Blackwell, 1980).

Last Writings on the Philosophy of Psychology, vol II, ed. G. H. von Wright and H. Nyman, trans. C. G. Luckhardt and M. A. E. Aue (Oxford: Blackwell, 1992).

Lectures on the Foundations of Mathematics, Cambridge, 1939, ed. Cora Diamond (Ithaca, N.Y.: Cornell University Press).

Letters to C. K. Ogden with Comments on the English Translation of the Tractatus Logico-Philosophicus, ed. G. H. von Wright (Oxford: Blackwell, and London: Routledge and Kegan Paul, 1973).

Notebooks, 1914–1916, ed. G. H. von Wright and G. E. M. Anscombe, trans. G. E. M. Anscombe (Oxford: Blackwell, 1961; 2nd edn 1979).

'Notes on Logic', Costelloe version, in L. Wittgenstein, *Notebooks, 1914–1916*, pp. 91–106.

On Certainty, ed. G. E. M. Anscombe and G. H. von Wright, trans. D. Paul and G. E. M. Anscombe (Oxford: Blackwell, 1969).

Philosophical Investigations, ed. G. E. M. Anscombe and R. Rhees, trans. G. E. M. Anscombe (Oxford: Blackwell, 1958).

Prototractatus, ed. B. F. McGuinness *et al.*, trans. D. Pears and B. F. McGuinness (Routledge and Kegan Paul, 1971).

Remarks on the Foundations of Mathematics, ed. G. H. von Wright *et al.*, trans. G. E. M. Anscombe (Oxford: Blackwell, 1978).

Tractatus Logico-Philosophicus, (London: Routledge and Kegan Paul, 1922).

Tractatus Logico-Philosophicus, trans. D. F. Pears and B. F. McGuinness (London: Routledge and Kegan Paul, 1963).

Zettel, ed. G. E. M. Anscombe and G. H. von Wright, trans. G. E. M. Anscombe (Oxford: Blackwell, 1967).

Wright, C., 'Truth: A traditional debate reviewed', in S. Blackburn and K. Simmons (eds.), *Truth*, pp. 203–38.

Wright, C., and Hale, B., *The Reason's Proper Study* (Oxford: Oxford University Press, 2004).

Wundt, W., *Logik* (Stuttgart: Enke, 1880–3).

Zermelo, E., 'Untersuchungen über die Grundlagen der Mengenlehre I', *Mathematische Annalen*, 65 (1908), pp. 261–81; trans. as 'Investigations in the foundations of set theory I', in J. van Heijenoort (ed.), *From Frege to Gödel*, pp. 199–215.

INDEX

OTHER VOLUMES IN THE SERIES OF CAMBRIDGE COMPANIONS

GREEK AND ROMAN PHILOSOPHY *Edited by* DAVID SEDLEY
HABERMAS *Edited by* STEPHEN K. WHITE
HAYEK *Edited by* EDWARD FESER
HEGEL *Edited by* FREDERICK C. BEISER
HEGEL AND NINETEENTH-CENTURY PHILOSOPHY *Edited by* FREDERICK C. BEISER
HEIDEGGER 2ND EDN *Edited by* CHARLES GUIGNON
HOBBES *Edited by* TOM SORELL
HOBBES'S *LEVIATHAN* *Edited by* PATRICIA SPRINGBORG
HUME 2ND EDN *Edited by* DAVID FATE NORTON and JACQUELINE TAYLOR
HUSSERL *Edited by* BARRY SMITH and DAVID WOODRUFF SMITH
WILLIAM JAMES *Edited by* RUTH ANNA PUTNAM
KANT *Edited by* PAUL GUYER
KANT AND MODERN PHILOSOPHY *Edited by* PAUL GUYER
KEYNES *Edited by* ROGER E. BACKHOUSE and BRADLEY W. BATEMAN
KIERKEGAARD *Edited by* ALASTAIR HANNAY and GORDON DANIEL MARINO
LEIBNIZ *Edited by* NICHOLAS JOLLEY
LEVINAS *Edited by* SIMON CRITCHLEY and ROBERT BERNASCONI
LOCKE *Edited by* VERE CHAPPELL
LOCKE'S *ESSAY CONCERNING HUMAN UNDERSTANDING* *Edited by* LEX NEWMAN
LOGICAL EMPIRICISM *Edited by* ALAN RICHARDSON and THOMAS UEBEL
MAIMONIDES *Edited by* KENNETH SEESKIN
MALEBRANCHE *Edited by* STEVEN NADLER
MARX *Edited by* TERRELL CARVER
MEDIEVAL JEWISH PHILOSOPHY *Edited by* DANIEL H. FRANK and OLIVER LEAMAN
MEDIEVAL PHILOSOPHY *Edited by* A. S. MCGRADE
MERLEAU-PONTY *Edited by* TAYLOR CARMAN and MARK B. N. HANSEN

MILL *Edited by* JOHN SKORUPSKI
MONTAIGNE *Edited by* ULLRICH LANGER
NEWTON *Edited by* I. BERNARD COHEN and GEORGE E. SMITH
NIETZSCHE *Edited by* BERND MAGNUS and KATHLEEN HIGGINS
OCKHAM *Edited by* PAUL VINCENT SPADE
PASCAL *Edited by* NICHOLAS HAMMOND
PEIRCE *Edited by* CHERYL MISAK
THE PHILOSOPHY OF BIOLOGY *Edited by* DAVID L. HULL and MICHAEL RUSE
PLATO *Edited by* RICHARD KRAUT
PLATO'S *REPUBLIC* *Edited by* G. R. F. FERRARI
PLOTINUS *Edited by* LLOYD P. GERSON
QUINE *Edited by* ROGER F. GIBSON JR
RAWLS *Edited by* SAMUEL FREEMAN
RENAISSANCE PHILOSOPHY *Edited by* JAMES HANKINS
THOMAS REID *Edited by* TERENCE CUNEO and RENÉ VAN WOUDENBERG
ROUSSEAU *Edited by* PATRICK RILEY
BERTRAND RUSSELL *Edited by* NICHOLAS GRIFFIN
SARTRE *Edited by* CHRISTINA HOWELLS
SCHOPENHAUER *Edited by* CHRISTOPHER JANAWAY
THE SCOTTISH ENLIGHTENMENT *Edited by* ALEXANDER BROADIE
ADAM SMITH *Edited by* KNUD HAAKONSSEN
SPINOZA *Edited by* DON GARRETT
THE STOICS *Edited by* BRAD INWOOD
TOCQUEVILLE *Edited by* CHERYL B. WELCH
WITTGENSTEIN *Edited by* HANS SLUGA and DAVID STERN