

# The Philosophy of Gottlob Frege

Richard L. Mendelsohn

CAMBRIDGE

CAMBRIDGE

more information - [www.cambridge.org/9780521836692](http://www.cambridge.org/9780521836692)

This page intentionally left blank

## The Philosophy of Gottlob Frege

This book is an analysis of Frege's views on language and metaphysics raised in "On Sense and Reference," arguably one of the most important philosophical essays of the past hundred years. It provides a thorough introduction to the function/argument analysis and applies Frege's technique to the central notions of predication, identity, existence, and truth. Of particular interest is the analysis of the Paradox of Identity and a discussion of three solutions: the little-known *Begriffsschrift* solution, the sense/reference solution, and Russell's "On Denoting" solution. Russell's views wend their way through the work, serving as a foil to Frege. Appendixes give the proofs of the first sixty-eight propositions of *Begriffsschrift* in modern notation.

This book will be of interest to students and professionals in philosophy and linguistics.

Richard L. Mendelsohn is Professor of Philosophy at Lehman College and the Graduate School, the City University of New York.



# The Philosophy of Gottlob Frege

RICHARD L. MENDELSON

*Lehman College and the Graduate School, CUNY*



**CAMBRIDGE**  
**UNIVERSITY PRESS**

CAMBRIDGE UNIVERSITY PRESS

Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo

Cambridge University Press

The Edinburgh Building, Cambridge CB2 2RU, UK

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

[www.cambridge.org](http://www.cambridge.org)

Information on this title: [www.cambridge.org/9780521836692](http://www.cambridge.org/9780521836692)

© Richard L. Mendelsohn 2005

This book is in copyright. Subject to statutory exception and to the provision of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published in print format 2005

ISBN-13 978-0-521-83669-2 hardback

ISBN-10 0-521-83669-7 hardback

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this book, and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

*For Marsha, Robin, and Josh*

*With Love*





# Contents

<i>List of Principles</i>	<i>page</i> xi
<i>Preface</i>	xv
1 Biography	1
2 Function and Argument	7
2.1 Introduction	7
2.2 What Is a Function?	8
2.3 Function and Argument	9
2.4 Extensions of the Notation	13
2.5 The Substitution Principle for Reference	19
2.6 Formal Mode and Material Mode	23
3 Sense and Reference	27
3.1 Introduction	27
3.2 The Paradox of Identity	28
3.3 The Sharpened Paradox	29
3.4 The Generalized Paradox	30
3.5 Three Solutions	30
3.6 Sense and Reference	33
4 Frege's <i>Begriffsschrift</i> Theory of Identity	41
4.1 Introduction	41
4.2 The <i>Begriffsschrift</i> Semantic Theory	42
4.3 Criticism: The Received View	49
4.4 Criticism: Church-Langford Considerations	52
4.5 Criticism: The Alleged Regress	58
4.6 Criticism: Use/Mention Confusion	60

5	Concept and Object	63
5.1	Introduction	63
5.2	Objects	65
5.3	The Combining Tie	67
5.4	Logical Grammar	69
5.5	Metaphors	71
5.6	The Puzzle of the Concept <i>Horse</i>	73
5.7	An Analysis of the Puzzle	74
5.8	A Solution to the Puzzle	77
5.9	Morals	81
6	Names and Descriptions	84
6.1	Introduction	84
6.2	Russell's Theory of Descriptions	85
6.3	The Scope Distinction	87
6.4	Russell's Three Puzzles	90
6.5	Frege and Russell on Definite Descriptions	95
7	Existence	100
7.1	Introduction	100
7.2	The Frege/Russell View About 'Existence'	102
7.3	Is 'Exists' a Predicate?	104
7.4	Russell's Machinery	105
7.5	Frege's Mistake	110
7.6	Nonreferring Singular Terms	118
7.7	Kant on Being	121
8	Thought, Truth Value, and Assertion	125
8.1	Introduction	125
8.2	The Frege Argument	125
8.3	A Sharpening of Frege's Argument	126
8.4	A Problematic Use of Frege's Argument	127
8.5	A Way out of Frege's Argument	130
8.6	Truth and Assertion	133
8.7	Is 'True' a Predicate?	135
8.8	The Correspondence Theory of Truth	137
9	Indirect Reference	140
9.1	Introduction	140
9.2	The Sense/Reference Story	144
9.3	Some Loose Ends	149
9.4	The Infinite Hierarchy	151
9.5	Collapsing the Hierarchy	154
9.6	Russell's Other Shoe	158
9.7	Reflections on the Argument	161

10	Through the Quotation Marks	163
10.1	Introduction	163
10.2	Quine: Structureless Names	166
10.3	Davidson: Demonstrative Names	170
10.4	Parsons: Fregean Names	173
10.5	A Formalism	180
10.6	Philosophical Remarks	181
	<i>Appendix A. Begriffsschrift in Modern Notation: (1) to (51)</i>	185
	<i>Appendix B. Begriffsschrift in Modern Notation: (52) to (68)</i>	198
	<i>Notes</i>	203
	<i>Bibliography</i>	217
	<i>Index</i>	223



## List of Principles

<b>PRINCIPLE 2.2.1 (FUNDAMENTAL PROPERTY OF FUNCTIONS)</b> <i>For any <math>x, y</math> in the domain of <math>f</math>, if <math>x = y</math>, then <math>f(x) = f(y)</math></i>	page 8
<b>PRINCIPLE 2.2.2 (GENERALIZED FUNDAMENTAL PROPERTY OF FUNCTIONS)</b> <i>If <math>x_1 = y_1, \dots, x_n = y_n</math>, then <math>g(x_1, \dots, x_n) = g(y_1, \dots, y_n)</math></i>	204
<b>PRINCIPLE 2.3.1 (COMPOSITIONALITY FOR REFERENCE)</b> <i>For any function-expression <math>\theta(\Omega)</math> and any name <math>\alpha</math>, <math>r(\theta(\alpha)) = r(\theta)[r(\alpha)]</math></i>	11
<b>PRINCIPLE 2.3.2 (INFORMAL COMPOSITIONALITY FOR REFERENCE)</b> <i>The reference of a complex is a function of the reference of its parts</i>	11
<b>PRINCIPLE 2.3.3 (EXTENSIONALITY FOR REFERENCE)</b> <i>For any function-expression <math>\theta(\Omega)</math> and any names <math>\alpha, \beta</math>, if <math>r(\alpha) = r(\beta)</math>, then <math>r(\theta(\alpha)) = r(\theta(\beta))</math></i>	12
<b>PRINCIPLE 2.3.4 (GENERALIZED COMPOSITIONALITY FOR REFERENCE)</b> <i>For any <math>n</math>-place function-expression <math>\theta(\Omega_1, \Omega_2, \dots, \Omega_n)</math> and any names <math>\alpha_1, \alpha_2, \dots, \alpha_n</math>, <math>r(\theta(\alpha_1, \alpha_2, \dots, \alpha_n)) = r(\theta)[r(\alpha_1), r(\alpha_2), \dots, r(\alpha_n)]</math></i>	204
<b>PRINCIPLE 2.3.5 (GENERALIZED EXTENSIONALITY FOR REFERENCE)</b> <i>For any <math>n</math>-place function-expression <math>\theta(\Omega_1, \Omega_2, \dots, \Omega_n)</math> and any names <math>\alpha_1, \alpha_2, \dots, \alpha_n, \beta_1, \beta_2, \dots, \beta_n</math>, if <math>r(\alpha_1) = r(\beta_1), \dots, r(\alpha_n) = r(\beta_n)</math>, then <math>r(\theta(\alpha_1, \alpha_2, \dots, \alpha_n)) = r(\theta(\beta_1, \beta_2, \dots, \beta_n))</math></i>	204
<b>PRINCIPLE 2.5.1 (SUBSTITUTION FOR REFERENCE)</b> <i>If <math>r(\alpha) = r(\beta)</math>, <math>S\alpha</math> and <math>S\alpha/\beta</math> have the same truth value</i>	19
<b>PRINCIPLE 2.5.2 (LEIBNIZ'S LAW)</b> $(\forall x)(\forall y)(x = y \supset (Fx \equiv Fy))$	19

PRINCIPLE 2.5.3 (ABOUTNESS)	<i>S<math>\alpha</math> is about <math>r(\alpha)</math></i>	21
PRINCIPLE 2.5.4 (CORRECTED SUBSTITUTION FOR REFERENCE)	<i>If <math>S\alpha</math> is about <math>r(\alpha)</math>, then if <math>r(\alpha) = r(\beta)</math>, then <math>S\alpha</math> and <math>S\alpha/\beta</math> have the same truth value</i>	23
PRINCIPLE 3.3.1 (BEGRIFFSSCHRIFT SUBSTITUTION)	<i>If <math>S\alpha</math> is about <math>r(\alpha)</math>, then if <math>r(\alpha) = r(\beta)</math>, <math>S\alpha</math> and <math>S\alpha/\beta</math> have the same cognitive value</i>	29
PRINCIPLE 3.6.1 (SENSE DETERMINES REFERENCE)	$r(\eta) = r(s(\eta))$	35
PRINCIPLE 3.6.2 (REFERENCE IS A FUNCTION)	<i>If <math>s(\eta) = s(\zeta)</math>, then <math>r(\eta) = r(\zeta)</math></i>	35
PRINCIPLE 3.6.3 (COMPOSITIONALITY FOR SENSE)	$s(\theta(\alpha)) = s(\theta)[s(\alpha)]$	37
PRINCIPLE 3.6.4 (EXTENSIONALITY FOR SENSE)	<i>If <math>s(\alpha) = s(\beta)</math>, then <math>s(\theta(\alpha)) = s(\theta(\alpha/\beta))</math></i>	37
PRINCIPLE 3.6.5 (SUBSTITUTION FOR SENSE)	<i>If <math>S\alpha</math> is about <math>r(\alpha)</math>, then if <math>s(\alpha) = s(\beta)</math>, then <math>S\alpha</math> and <math>S\alpha/\beta</math> have the same cognitive value</i>	39
PRINCIPLE 4.2.1 (BEGRIFFSSCHRIFT SUBSTITUTION)	<i>If <math>S\alpha</math> is about <math>r(\alpha)</math>, then if <math>r(\alpha) = r(\beta)</math>, then <math>S\alpha</math> has the same conceptual content as <math>S\alpha/\beta</math></i>	44
PRINCIPLE 4.4.1 (CHURCH-LANGFORD TRANSLATION)	<i>If <math>oR\omega</math>, then <math>T(o) R T(\omega)</math></i>	53
PRINCIPLE 4.4.2 (SINGLE-QUOTE TRANSLATION)	<i>Expressions inside single quotes are not to be translated</i>	53
PRINCIPLE 7.2.1 (FREGE/RUSSELL ON 'EXISTENCE')	<i>To assert that <math>Fs</math> exist is to say that there are <math>Fs</math>, and to deny that <math>Fs</math> exist is to say that there aren't any <math>Fs</math></i>	102
PRINCIPLE 7.2.2 (FREGE/RUSSELL ON EXISTENCE)	<i>(i) 'x exists' is not a first-order predicate; (ii) Existence is not a property of objects but of properties; and (iii) Existence is completely expressed by means of the quantifier 'There is'</i>	103
PRINCIPLE 7.3.1 (REDUNDANCY THEORY OF EXISTENCE)	$(\exists x) F(x) \equiv (\exists x)(\mathcal{E}x \wedge Fx)$ $\neg(\exists x) F(x) \equiv \neg(\exists x)(\mathcal{E}x \wedge Fx)$	105
PRINCIPLE 9.2.1 (INDIRECT REFERENCE)	$r_1(t) = r_o(\Theta(t))$	147
PRINCIPLE 9.2.2 (COMPOSITIONALITY FOR INDIRECT REFERENCE)	$r_1(\theta(\alpha)) = r_1(\theta)[r_1(\alpha)]$	148

- PRINCIPLE 9.2.3 (*THAT*)  $r_o(\Theta(\theta(\alpha))) = r_o(\Theta(\theta))r_o[(\Theta(\alpha))]$  148
- PRINCIPLE 9.2.4 (COMPOSITIONALITY FOR *THAT*)  $r_o(\theta(\Theta(t))) = r_o(\theta)[r_1(t)]$  148
- PRINCIPLE 9.2.5 (EXTENSIONALITY FOR *THAT*) *If  $r_1(\alpha) = r_1(\beta)$ , then  $r_o(\theta(\Theta(\alpha))) = r_o(\theta(\Theta(\beta)))$*  148
- PRINCIPLE 9.4.1 (NO SELF-REFERENCE)  $s_i(t) \neq r_i(t)$  153
- PRINCIPLE 9.6.1 ( $\Theta$  COLLAPSE)  $\Theta\Theta(\theta(\alpha)) = \Theta(\theta(\alpha))$  160
- PRINCIPLE 10.2.1 (QUINE NO SELF-REFERENCE) *A name must be distinct from the object it names* 167
- PRINCIPLE 10.5.1 (QUOTATION-NAME DENOTATION) *For any expression  $e \in \mathcal{V}^n$ ,  $\langle lq, e, rq \rangle$  denotes  $e$*  180





## Preface

Gottlob Frege is celebrated for his distinction between the *Sinn* and *Bedeutung* – the sense and reference – of a term. The distinction is readily understood. The reference of the name ‘Plato’ is the bearer of the name, that most famous and widely revered philosopher, who lived more than two thousand years ago in ancient Greece. The sense of the name ‘Plato’, on the other hand, corresponds to what we would ordinarily recognize as belonging to its meaning: what speakers and hearers understand by the word that enables them to identify what they are talking about and to use the word intelligently. Why is Frege celebrated for this distinction? After all, just a generation or two before, Mill (1843) expounded his distinction between the connotation and denotation of a name. In *The Port Royal Logic*, Arnauld (1662) drew a kindred distinction between an idea and its extension. In his *Summa Logicae*, William of Ockham (c. 1323) distinguished between the term in mental language associated with a word and what it supposits. Earlier still, in ancient times, the Stoic logicians distinguished between an utterance, its signification, and the name-bearer.<sup>1</sup> This is a very natural distinction, and we find variations on its theme reappearing throughout philosophical history. What makes *Frege’s* distinction so noteworthy? The answer lies with his compositionality principles, one for reference and the other for sense. These represent a genuine advance. Frege conceived of the semantic value of a complex construction in language as being determined by the simpler ones from which it is built in a mathematically rule-governed manner. These rules provided him with a framework within which rationally to connect and unify the semantic story posited for various linguistic entities. At the very same time, it generated an explanation for the creativity of language. This last insight, which

came into clearer focus only late in Frege's intellectual life, has proved compelling and invigorating to the logical, psychological, linguistic, and philosophical investigation of language in the twentieth and twenty-first centuries.

Although the rudiments of the function/argument analysis were in place in *Begriffsschrift*, the fundamental semantic notion of the content [*Inhalt*] of a sentence was unstable. Frege was assuming a classic philosophical picture of a level of thoughts and another level of a reality that was represented by these thoughts. But it was a picture that needed to be drawn more sharply in order to fit with the mathematical devices he had created. The *Begriffsschrift* notion of the content of a simple atomic sentence  $S\alpha$  combined two distinct semantic strands: the part corresponding to the singular term was the reference of the expression and the part corresponding to the predicate was the sense of the expression. Keeping his eye firmly focused on the function/argument structure, Frege was able to win through (although twelve years later) to his sense/reference distinction: this helped enormously to clarify the important connections between the various types of expressions set in place by the compositionality principles. But confusion remained, most clearly in the application of the distinction to predicate expressions, and, relatedly, in the way in which the function/argument structure was to apply at the level of sense. We will examine an important example of the former error, namely, his enormously influential treatment of *existence*: although the problem of accounting for the informativeness of existence statements is on a par with the problem of accounting for the informativeness of identity statements, Frege ignored the parallel and persisted in denying that existence was a property of objects. Frege (1892*c*) drew his sense/reference distinction to explain the informativeness of descriptions without, unlike Russell after him, also providing a logical mechanism for them. Russell accounted for the sense of a description via the inferential connections of the underlying predicate construction; but Frege regarded descriptions as individual constants, and it remains an open problem how his notion of sense engages with these predicate constructions. Russell's famous account of definite descriptions provides a powerful foil for probing Frege's semantic theory. Russellian views will wend their way through our discussion of Frege's semantics, leading us to an example of the second sort of problem mentioned above, namely, Frege's analysis of indirect contexts. It is widely believed that Frege's semantics of indirect contexts leads to an infinite hierarchy of semantic primitives, a problem actually set in motion by Russell's (1905) criticism of Frege's distinction. We will examine both

*oratio obliqua* and *oratio recta* contexts and show that neither leads to the absurdity charged. The critical distinction, as Dummett saw, is between customary sense and indirect sense; the differences in the levels of indirect sense pose no theoretical challenge to a rule-governed semantic story.

We will, in this book, be tracing some of the philosophical implications of what we take to be Frege's central innovation in philosophy of language, namely, the function/argument analysis. We do not pretend that this book is a comprehensive treatment of Frege's philosophy. We have little to offer on his important contributions to the foundations of mathematics. Even in our discussion of Frege's philosophy of language, there will be omissions: in particular, Frege's treatment of demonstratives – indeed, any in-depth analysis of Frege's notion of sense. These introduce a level of difficulty that we are not prepared to address. Our landscape is already sufficiently fraught with philosophical minefields, for we will be tackling some of the fundamental issues that exercised philosophers in the twentieth century, and we are pleased to have been able to advance as far as we have on them. Our goal here is, quite modestly, to illuminate Frege's central insight, which we take to be the function/argument analysis, at the level of reference, and to pursue this insight into the most difficult terrain of indirect contexts, hoping thereby to help clarify philosophical issues Frege grappled with.

On our reading, the sense/reference theory marked a sharp rejection of the view Frege had held earlier in *Begriffsschrift*, and which was later a standard of Russell and the early Wittgenstein, namely, the view that has come to be known as *direct reference*. Wittgenstein (1922) expressed the doctrine so:

3.203 A name means an object. The object is its meaning.

Although, as we just mentioned, Frege (1879) also upheld this principle, Frege (1892*c*) categorically rejected it. Frege (1892*c*) abandoned direct reference entirely, by contrast with Russell (1905), who, faced with the same puzzle, preserved direct reference for “genuine” proper names. The disagreement between the two is evident in the series of letters they exchanged.<sup>2</sup> In recent years, direct reference has once again become the focal point of philosophical controversy. Russellians accept the principle, while Fregeans reject it.

Within the context of the controversy, it is clearly inadvisable to translate Frege's *Bedeutung* into English as *meaning*. For on that suggested translation, Wittgenstein's words capture exactly the thought Frege (1892*c*)

sought to uphold, and the disagreement between the two disappears.<sup>3</sup> A number of Frege scholars, including those who have worked so hard to make his views available to the English-speaking world, have replaced earlier choices, like the classical Black and Geach (1952) rendering as *reference*, in favor of *meaning*. But the virtues of this replacement are quite theoretical and have yet to reveal themselves. Whatever they might be, they are thoroughly outweighed by the confusion and discomfort engendered in a philosophically literate English-language reader for whom the issue of the *meaning* of a proper name, not its *Bedeutung*, is salient. Black and Geach's (1952) original choice of *reference* for *Bedeutung*, and secondarily, expressions like *designation* and *denotation*, are most comfortable. These preserve the truth value of the German original, and, in addition, provide us with a means of stating Frege's view with reasonable clarity in English. Because Black and Geach (1952) is no longer readily available, we will use Beaney's (1997) translation as the primary source for our citations. (All quotations of Frege's writings are drawn from the translations identified in the Bibliography.) Beaney (1997: 44) admits that "[i]f forced to choose, I myself would use 'reference' . . .," but in the text he decided to leave the noun '*Bedeutung*' untranslated.

We will see in Chapter 1 that Frege's project was primarily technical. His *Logicist* program, as it has come to be called, involved (a) formalizing a logic sufficient to represent arithmetical reasoning, (b) providing definitions for arithmetical constants and operations, in purely logical terms, and (c) representing the definitionally expanded truths of arithmetic as truths of logic. Portions of this project were enormously successful, but others turned out to be disastrous. Russell located a contradiction in Frege's unrestricted comprehension schema for sets and communicated it to Frege just as the second volume of *Grundgesetze* was in press. Frege never found a solution to the problem and came to believe his program was in ruins. The Logicist program was dealt another severe setback years later when Gödel showed that not all the truths of arithmetic were provable. In any event, work on the foundations of mathematics and the philosophy of mathematics soon outstripped Frege's achievements, even his relevance. Frege's philosophy of language, however, remains intensely vital today. Not since medieval times has the connection between logic and language been so close.

Earlier versions of parts of this book have, over time, been published as separate essays. Portions of Chapters 2 and 8 are from "Frege and the Grammar of Truth," which appeared in *Grammar in Early Twentieth-Century*

*Philosophy*, ed. Richard Gaskin (Routledge, London, 2001), pp. 28–53. Portions of Chapters 3 and 4 are from “Frege’s *Begriffsschrift* Theory of Identity,” *Journal of the History of Philosophy* 20 (1982), 279–99. Portions of Chapter 5 are from “Frege on Predication,” *Midwest Studies in Philosophy* 6 (1981), 69–82. Portions of Chapter 9 are from “Frege’s Treatment of Indirect Reference,” in *Frege: Importance and Legacy*, ed. Matthias Schirn (Walter de Gruyter, Berlin, 1996), pp. 410–37. With respect to the first and last of these articles, however, we caution the reader that the position we adopt here is significantly different from the one we defended in those essays.

Our debt to the work of Michael Dummett should be evident throughout. Almost single-handedly, he brought Frege’s philosophy into mainstream consciousness. And although we disagree with W. V. O. Quine on many of these pages, our debt to his work is evident as well. Our original interest in Frege was piqued by the way in which Quine applied technical devices to philosophical problems. Finally, we are very grateful to F. Fritsche, who helped correct earlier drafts of the two appendixes.



## Biography

The known details of the personal side of Frege's life are few.<sup>1</sup> Friedrich Ludwig Gottlob Frege was born November 8, 1848, in Wismar, a town in Pomerania. His father, Karl Alexander (1809–1866), a theologian of some repute, together with his mother, Auguste (d. 1878), ran a school for girls there. Our knowledge of the remainder of Frege's personal life is similarly impoverished. He married Margarete Lieseberg (1856–1904) in 1887. They had several children together, all of whom died at very early ages. Frege adopted a child, Alfred, and raised him on his own.<sup>2</sup> Alfred, who became an engineer, died in 1945 in action during the Second World War.<sup>3</sup> Frege himself died July 26, 1925, at age seventy-seven.

We can say somewhat more about his intellectual life. Frege left home at age twenty-one to enter the University at Jena. He studied mathematics for two years at Jena, and then for two more at Göttingen, where he earned his doctorate in mathematics in December 1873 with a dissertation, supervised by Ernst Schering, in geometry. Although mathematics was clearly his primary study, Frege took a number of courses in physics and chemistry, and, most interestingly for us, philosophy. At Jena, he attended Kuno Fischer's course on Kant's Critical Philosophy, and in his first semester at Göttingen, he attended Hermann Lotze's course on the Philosophy of Religion. The influence and importance of Kant is evident throughout Frege's work, that of Lotze's work on logic is tangible but largely circumstantial.<sup>4</sup>

After completing his *Habilitationsschrift* on the theory of complex numbers, Frege returned to Jena in May of 1874 in the unsalaried position of lecturer [*Privatdozent*]. The position was secured for him by the mathematician Ernst Abbé, his guardian angel at Jena from the time he arrived

as a student to his ultimate honorary professorship.<sup>5</sup> Abbé controlled the Carl Zeiss foundation, which received almost half of all the profits from the Zeiss lens and camera factory (which Abbé had helped the Zeiss family establish). Frege's unsalaried honorary professorship at Jena was made possible because he received a stipend from the Zeiss foundation.

Frege taught mathematics at Jena and his first published writings were mainly reviews of books on the foundations of mathematics. In 1879, five years after returning to Jena, he published his *Begriffsschrift*. It was not well received. For one thing, the notation was extraordinarily cumbersome and difficult to penetrate. Also Frege failed to mention, and contrast with his own system, the celebrated advances in logic by Boole and Schröder, in which both classical truth-functional logic and the logic of categorical statements were incorporated into a single mathematical system. In his review of *Begriffsschrift*, Schröder ridiculed the idiosyncratic symbolism as incorporating ideas from Japanese, and as doing nothing better than Boole and many things worse. Schröder had not realized how far Frege had penetrated, and neither did many of his contemporaries.<sup>6</sup>

For three years, Frege worked hard to explain and defend his *Begriffsschrift*, though not with much success.<sup>7</sup> The fault lies in no small measure with Frege himself, for he failed to distinguish in importance the specifics of his notation (which has, thankfully, been totally abandoned) from the logical syntax and semantics it instantiated. What Frege had created, of course, was a formal language in which he axiomatized higher-order quantificational logic; derived many theorems of propositional logic, first-order logic, and second-order logic; and defined the ancestral relation. *Begriffsschrift* represents a milestone, not only in the history of logic and, thereby, in the history of philosophy, but also in the history of modern thought, for it was one of the first sparks in a hundred-year explosion of research into the foundations of mathematics, and into the application of mathematical representation to structures other than numbers and shapes.

Frege soon broke away from this engagement and returned to his creative project announced in *Begriffsschrift*:

[We] divide all truths that require justification into two kinds, those whose proof can be given purely logically and those whose proof must be grounded on empirical facts. . . . Now, in considering the question of to which of these two kinds arithmetical judgments belong, I first had to see how far one could get in arithmetic by inferences alone, supported only by the laws of thought that transcend all particulars. The course I took was first to seek to reduce the concept of ordering in a series to that of *logical* consequence, in order then to progress to the concept of number. . . . (Frege 1879: 48)



Having codified the notion of proof, of logical consequence, and of ordering in a sequence in *Begriffsschrift*, Frege pursued his investigation into the notion of *cardinal number*, publishing his philosophical strategy in 1884 in *Grundlagen*. Unlike his *Begriffsschrift*, *Grundlagen* is almost devoid of formal symbolism and is otherwise directly engaged with the main views current about arithmetic. His polemic against contemporary empiricist and naturalist views of the concept of number is devastating. It is not only the specifics of these views that Frege believes to be wrong, but also the methodology of seeking a foundation for mathematics by identifying referents for the number words, whether they be material objects, psychological ideas, or Kantian intuitions. This is the cash value of his injunction against looking for the meaning of number words *in isolation*. The numbers, along with sets and the truth values, are *logical objects*: their meaning is intimately bound up with our conceptualization of things. He codified this attitude in his famous *Context Principle* – never to look to the meaning of a word in isolation, but only in the context of a proposition. For Frege, the foundations of mathematics were to be found in the new logic he had created, the language of which was adequate to express all elementary arithmetic statements, so that the truths of logic could be seen to be, when spelled out, truths of logic. *Grundlagen* is widely regarded as a masterpiece written by a philosopher at the height of his powers: in the years from 1884 through the publication of *Grundgesetze*, in 1893, we see Frege at his creative height.

Frege's *Grundlagen*, although free from the symbolism of his more technical works, did not receive much notice, and the little it did receive was, as usual, full of misconceptions. It is not entirely clear why this is so. Perhaps Frege appeared too philosophical for the mathematicians who were working in related areas – he was ignored by Dedekind, roundly criticized by Cantor, and dismissed by Hilbert – and too technical for the philosophers. Only the direct interaction with Husserl – Frege (1894) demolished Husserl's early psychologism in a review – had a clear and immediate impact on active philosophers of his day. Husserl abandoned his psychologism shortly thereafter, but he was none too generous in later life when he recalled Frege to be a man of little note who never amounted to much.

Frege's own philosophical education and his knowledge of historical and contemporary philosophers is extremely problematic. When he quotes from some of the classical philosophers like Descartes, Hobbes, and Leibniz, it is frequently from a popular anthology put together by Baumann (1868) of writings on the philosophy of space and time. Kant gets a great many footnotes, though largely for his work on arithmetic

and geometry. It is never clear how much of a philosopher's work Frege was familiar with because he picked and chose discussions that were directly related to the problems he was working on. As with an autodidact, there appear to be immense holes in Frege's knowledge of the history of philosophy; this, plus the single-mindedness with which he approached issues, as if with blinders to what was irrelevant, just underscored his intellectual isolation.

*Grundlagen* could not, of course, represent the end of his project. Frege would never be satisfied until he demonstrated his position formally. And it was the effort to formalize his view that forced significant changes in the *Grundlagen* story. Frege had tried to make do earlier in *Begriffsschrift* without the notion of set; he had yet to convince himself that the notion was legitimate and that it belonged in logic. At any rate, with the publication of *Grundlagen*, Frege's course was clear: to fill in the logical details of the definition of number he there presented in the manner of his *Begriffsschrift*. What had been missing was a conception of a set; this Frege won through to. Along the way, a sharpening of his philosophical semantics led to the mature views in philosophy of language for which he has been justly celebrated. "Über Sinn und Bedeutung" was published in 1892, and its companion essays appeared in print about that same time.

*Grundgesetze* was published in 1893 by Hermann Pohle, in Jena. Frege had had difficulty finding a publisher for the book, after the poor reception given to his other works. Pohle agreed to publish the work in two parts: if the first volume was received well, he would publish the second one. Unfortunately it was not received well, to the extent that it was acknowledged by anyone at all. Pohle refused to publish the second volume, and Frege paid for its publication out of his own pocket some ten years later.

Just as Volume 2 of *Grundgesetze* was going to press in 1902, Russell communicated to Frege the famous contradiction he had discovered. Here is the beginning of the first letter to Frege, dated June 16, 1902:

Dear Colleague,

I have known 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. I find myself in full accord with you on all main points, especially in your rejection of any psychological element in logic and in the value you attach to a conceptual notation for the foundations of mathematics and of formal logic, which, incidentally, can hardly be distinguished. On many questions of detail, I find discussions, distinctions and definitions in your writings for which one looks in vain in other logicians. On functions in particular (sect. 9 of your *Conceptual*

*Notation*) I have been led independently to the same views even in detail. I have encountered a difficulty only on one point. You assert (p. 17) that a function could also constitute the indefinite element. This is what I used to believe, but this view now seems to me dubious because of the following contradiction: Let  $w$  be the predicate of being a predicate which cannot be predicated of itself. Can  $w$  be predicated of itself? From either answer follows its contradictory. We must therefore conclude that  $w$  is not a predicate. Likewise, there is no class (as a whole) of those classes which, as wholes, are not members of themselves. From this I conclude that under certain circumstances a definable set does not form a whole. (Frege 1980: 130–1)

From his Axiom 5,

$$\{x|Fx\} = \{x|Gx\} \equiv (\forall x)(Fx \equiv Gx),$$

which lays out the identity conditions for sets, Frege (1893) derives Proposition 91:

$$Fy \equiv y \in \{x|Fx\}.$$

Russell's contradiction is immediate when, in this proposition, the property  $F$  is taken to be *is not an element of itself* and the object  $y$  is taken to be *the set of all sets that are not elements of themselves*:<sup>8</sup>

$$\neg\{x|\neg x \in x\} \in \{x|\neg x \in x\} \equiv \{x|\neg x \in x\} \in \{x|\neg x \in x\}.$$

Unlike Peano, to whom Russell had also communicated the paradox, Frege acknowledged it with his deep intellectual integrity and attempted to deal with it in an appendix – but to no avail, as he himself acknowledged. He was deeply shaken by this contradiction, which emerged from an axiom about which he had, as he said, always been somewhat doubtful. His life's work in a shambles, Frege's creative energies withered. The foundational paradoxes became a source of immense intellectual stimulation (as Frege himself had surmised in a letter to Russell) and his achievements were soon surpassed by the work of Ernst Zermelo and others. By the time the young Ludwig Wittgenstein came to see him in 1911 to study foundations of mathematics, Frege referred him to Russell. There was a brief flurry of activity in 1918–19 when Frege published some work in philosophy of logic in an Idealist journal. They appear to represent the first chapters of a planned book on logic. These essays remain among the most influential writings of the twentieth century. But the foundations of arithmetic are a different story. We find him saying, in the early 1920s, that he doubts whether sets exist at all. And he is trying to see if the roots

of arithmetic are to be found in geometry, a complete turnaround from his earlier views.

That we know of Frege today is largely through his influence on the giants of modern analytic philosophy. Russell was the first to become aware of his work in the philosophy of language and logic. He included an appendix describing Frege's views in his *Philosophy of Mathematics* of 1903. Indeed, immediately afterward, Russell appears to have been most deeply preoccupied with working out Frege's sense/reference theory, an enterprise he abandoned because he thought there were insuperable difficulties with the view and also because he had an alternative in his theory of descriptions. Wittgenstein, too, had been deeply influenced by Frege's views, and many parts of the *Tractatus* are devoted to them. Finally, we mention Rudolf Carnap, who had attended Frege's lectures at Jena – he describes how Frege lectured into the blackboard so that the handful of students in the room could barely hear him – and whose book *Meaning and Necessity* resuscitated interest in Frege and formal semantics.

Frege retired from Jena in 1918. He had become increasingly involved with right-wing political organizations toward the latter part of his life, and the journal he kept in spring 1924<sup>9</sup> reveals a side of him that is not very appealing.

## Function and Argument

### 2.1 Introduction

*Begriffsschrift* was, as the subtitle announced, *a formula language of pure thought modeled upon the language of arithmetic*. Frege borrowed the notation for functions from arithmetic, and enlarged the realm of applicability of a function beyond the domain of numbers. Then, supplanting the subject/predicate division, which was characteristic of previous logical systems, by a function/argument division, he created a logical notation, a *Begriffsschrift* – literally, *Concept Writing* – which would serve to represent thoughts about any objects whatsoever. Like the language of arithmetic, his *Begriffsschrift* represented thoughts so that the inferential connections between them were molded in the representations themselves. The project was enormously successful. Not only did Frege create modern quantificational logic, but he also provided the theoretical framework for many subsequent philosophical developments in logic as well as in speculative philosophy. As Dummett (1981*a*) correctly remarked, Frege’s work shifted the central focus of philosophy from the epistemological issues raised by Descartes back to the metaphysical and ontological issues that were salient after Aristotle.

The function/argument analysis Frege (1879) presented was, however, flawed. There was a significant confusion in his operating semantic notion of the *content* [*Inhalt*] of a sentence. Frege came to recognize that repairs were needed, and after much hard philosophical work, the theory with which we are now familiar emerged in the early 1890s. It was announced first in Frege (1891), and then elaborated upon in Frege (1892*c*) and Frege (1892*a*). We will present Frege’s

mature function/argument analysis, and later on, when we discuss the sense/reference distinction, explain some of the changes he had to make in his earlier theory. There is no place we are aware of where a reader can find the function/argument structure spelled out in any detail, so we have taken the liberty of presenting it here. The reader for whom this material is too elementary can simply leap to the next chapter.

## 2.2 What Is a Function?

The modern notion of a function goes like this. For any nonempty sets,  $S$  and  $S'$  (not necessarily distinct), a function  $f$  from  $S$  to  $S'$  correlates elements of  $S$  (*the domain of  $f$* ) with elements of  $S'$  (*the range of  $f$* ). If  $x \in S$ , then  $f(x) \in S'$  and  $f(x)$  is *the value* of the function  $f$  for *the argument*  $x$ . We are justified in speaking of *the value* of the function for a given argument because of the following fundamental property of functions:

**PRINCIPLE 2.2.1 (FUNDAMENTAL PROPERTY OF FUNCTIONS)** *For any  $x, y$  in the domain of  $f$ , if  $x = y$ , then  $f(x) = f(y)$ .*

Hence,  $f$  associates each element of  $S$  with but a single element of  $S'$ .<sup>1</sup>

A function is a special type of a *relation*, one that associates each element of the domain with a unique element of the range. Of course, a given element in the domain might be associated with more than one element of the range. In that case, however, the association is a relation that is not a function. *Being the brother of*, for example, is a relation that is not a function: it associates an individual with his brother(s). *Being the square root of* is an arithmetic relation that is not a function: although we speak of *the square root of 4*, we speak misleadingly, for there are two square roots of 4, +2 and -2.

Set-theoretically, relations and functions are conceived of as  $n$ -tuples of elements. A two-place relation, for example, will be a subset of the Cartesian Product  $S \times S'$ , that is, the set of ordered pairs  $\langle x, y \rangle$ , with  $x \in S$  and  $y \in S'$  ( $S$  and  $S'$  not necessarily distinct). Principle 2.2.1 tells us that if  $y = z$  whenever  $\langle x, y \rangle$  and  $\langle x, z \rangle$  are both in the relation, then that relation is also a function.

Before continuing, a word of caution is in order. Frege did not identify a function with a set of ordered pairs. The set of ordered pairs corresponds rather to what he called the *Werthverlauf* – the value range or course of values – of the function. We will see in Chapter 5 that Frege maintained a fundamental ontological division between objects [*Gegenstände*] on the one hand and functions on the other (corresponding roughly – very

roughly – to the traditional distinction between objects and properties). The former, among which he counted *Werthverläufe*, are complete, self-subsistent entities; the latter are not self-subsistent, but, continuing Frege’s metaphors, are unsaturated and stand in need of completion. However, for Frege, functions are the same if they yield the same values for the same arguments. Since this is in accord with the extensional view we are used to, we can rely on our set-theoretic intuitions as heuristic whenever ontological considerations fade into the background.

Here are some examples of arithmetic functions. The square function  $f(x) = x^2$  is a singular function, that is, a function of one argument. It maps integers into integers, associating each integer with its square: it maps 1 into 1, 2 into 4, 3 into 9, and so on. Addition,  $f(x,y) = x + y$ , is a *binary* function. It maps a pair of integers into integers: it maps the pair  $\langle 1, 1 \rangle$  into 2, it maps the pair  $\langle 2, 3 \rangle$  into 5, and so on.

In speaking as we have of functions, we have said very little about how the association is to be set up, or how the function is to be evaluated for a given argument. The set-theoretic perspective bypasses this important feature of the algebraic character of functions, which is crucial to our intuitive understanding of the notion. For example, when we consider the square function, expressed algebraically as  $f(x) = x^2$ , we think of the function as a way of getting from one number (the argument) to another (the value). It is the well-known mathematical procedure associated with the algebraic formula that gives the sense that the association between the domain and range is *orderly*. It is actually a rather large leap to suppose that a set of ordered pairs satisfying Principle 2.2.1 is a function, even when no procedure is available for associating the elements of the domain with the elements of the range. We are not sure how Frege would stand on this issue. We are inclined to believe that without an algebraic formula he would not be so quick to accept the existence of a function, because he posited sets *only* as the extensions of concepts – without the concept to identify the elements of the set, one could not otherwise assume the existence of such a set. But we cannot be sure of this.

### 2.3 Function and Argument

We now rehearse the analysis of the function/argument notation in mathematics, drawing mainly from Frege (1891). The linear function

$$f(x) = (2 \cdot x) + 1 \tag{2.1}$$

maps integers into integers. For the arguments 1, 2, and 3, the function yields the values 3, 5, and 7, respectively. Frege observes that the arithmetic equation

$$'3 = (2 \cdot 1) + 1' \quad (2.2)$$

is an identity, in fact, a true identity.<sup>2</sup> (2.2) says that the number 3 is identical with the number which is obtained by adding 1 to the result of multiplying 2 by 1. Since  $'(2 \cdot 1) + 1'$  flanks the identity sign in (2.2), it serves as a name: it designates the number that is obtained by adding 1 to the result of multiplying 2 by 1, namely, the number 3.

Unlike the numeral '3', however, which is a *simple* referring expression,  $'(2 \cdot 1) + 1'$  is a *complex* referring expression: it contains numerals as proper parts along with the symbols for addition and multiplication. The complex expression  $'(2 \cdot 1) + 1'$  was constructed by replacing the variable 'x' in the right-hand side of the equation in (2.2) by the numeral '1'. Now,

$$'(2 \cdot x) + 1' \quad (2.3)$$

does not stand for a number, and it especially does not stand for a variable or indefinite number as some of Frege's contemporaries were inclined to suppose. To prevent just such an error, Frege preferred to leave the variable 'x' out entirely and enclose the remaining blank space in parentheses, so that (2.3) would become

$$'(2 \cdot ( )) + 1', \quad (2.4)$$

an evidently incomplete expression. Though preferred, this notation is deficient when we have a function of more than one argument because we lose the difference in the variables that shows when we must insert the same numeral and when we need not. Frege eventually compromised by using the lower case Greek  $\eta$  and  $\zeta$  instead of the blank spaces, writing (2.4) as

$$'(2 \cdot \eta) + 1'. \quad (2.5)$$

Again, (2.5) does not stand for a number, but rather, Frege suggested, for the linear function we started out with, namely,

$$(2 \cdot \eta) + 1. \quad (2.6)$$

Frege's suggestion captures the notation beautifully. For, inserting the numerals

$$'1', '2', '3' \quad (2.7)$$



in (2.5) yields the expressions

$$'(2 \cdot 1) + 1', '(2 \cdot 2) + 1', '(2 \cdot 3) + 1', \quad (2.8)$$

respectively. And evaluating the function (2.6) for the arguments

$$1, 2, 3 \quad (2.9)$$

yields the values

$$3, 5, 7, \quad (2.10)$$

respectively.<sup>3</sup> The numerals in (2.7) stand for the numbers in (2.9), and the expressions in (2.8) stand for the numbers in (2.10). These numbers in (2.10) are the values of the function (2.6) for the arguments (2.9).

Frege's analysis of the mathematical notation has thus far yielded the following results. When a number-name (that is, a numeral or complex referring expression which uniquely identifies a number) is inserted in a function-expression, the number-name and the function-expression combine to form a complex referring expression. Let  $\theta(\Omega)$  be a function-expression with one argument place marked by  $\Omega$ ; and let  $\alpha$  be a name. We combine the function-expression  $\theta(\Omega)$  with the name  $\alpha$  to form the complex expression  $\theta(\alpha)$ . What does this complex expression refer to? It refers to the value of the function that  $\theta(\Omega)$  refers to, evaluated for the argument  $\alpha$  refers to.

Let  $r(\eta)$  be *the reference of  $\eta$* . We express this principle governing the function/argument notation as

**PRINCIPLE 2.3.1 (COMPOSITIONALITY FOR REFERENCE)** *For any function-expression  $\theta(\Omega)$  and any name  $\alpha$ ,  $r(\theta(\alpha)) = r(\theta)[r(\alpha)]$ .*

Take, for example, the complex expression ' $3^2$ '. The reference of this complex expression –  $r('3^2')$  – is 9. It is the result of applying the function designated by the function-expression –  $r('(\eta)^2')$  – to the object designated by the argument-expression –  $r('3')$ . Compositionality is sometimes expressed like this:

**PRINCIPLE 2.3.2 (INFORMAL COMPOSITIONALITY FOR REFERENCE)** *The reference of a complex is a function of the reference of its parts.*

Since a function yields a unique value for a given argument, we obtain as a direct corollary to Principle 2.3.1 that a complex referring expression

formed in this manner has a *unique* reference,

PRINCIPLE 2.3.3 (EXTENSIONALITY FOR REFERENCE) *For any function-expression  $\theta(\Omega)$  and any names  $\alpha, \beta$ , if  $r(\alpha) = r(\beta)$ , then  $r(\theta(\alpha)) = r(\theta(\beta))$ .*<sup>4</sup>

Principles 2.3.1 and 2.3.3 are the key principles of the function/argument analysis. Principle 2.3.1 says, informally, that the reference of a complex expression is uniquely determined by the reference of its parts. Principle 2.3.3 says, informally, that the reference of the constituent expressions is the *only* feature of these expressions that counts towards determining the reference of the complex. It is evident from the examples given that Principle 2.3.3 is the relevant principle when it comes to the practical question of determining whether a given expression  $E(\eta)$ , containing the constituent expression  $\eta$ , is complex or not. The procedure is in two parts. Step One: we replace  $\eta$  by coreferential expressions, and if the reference of the whole remains invariant under these substitutions, then the likelihood is that the reference of  $E(\eta)$  depends upon the reference of the constituent  $\eta$ . Step Two: we replace  $\eta$  by expressions that stand for different objects, and we repeat the procedure from Step One for each of these expressions; if we find reference-invariance in each case, then we have evidence of the orderly connection between the reference of the part and the reference of the whole characteristic of a function, and thus we have evidence that  $E(\eta)$  is a function-expression.

If a complete expression contains a name as a proper part, and if this constituent name might be replaced by others, in each case to result in a sensible complete expression, then what remains of the complete expression after the constituent name is deleted is a function-expression *provided that* Principles 2.3.1 and 2.3.3 (or their generalizations) are satisfied. The reference of the complex expression is thus shown to be a function of the reference of the constituent name. Frege's clearest statement of the general function/argument analysis is from *Grundgesetze*, Volume 2, Section 66:

Any symbol or word can indeed be regarded as consisting of parts; but we do not deny its simplicity unless, given the general rules of grammar, or of the symbolism, the reference of the whole would follow from the reference of the parts, and these parts occur also in other combinations and are treated as independent signs with a reference of their own. (Black and Geach 1952: 171)

A *simple* expression is an expression that has no significant structure, that is, one that cannot be parsed into function-expression and argument-expression(s) such that the reference of the whole is a function of the

reference of the parts. A simple expression might be a single symbol, for example, the numeral ‘1’ or it might be a sequence of symbols, for example, the English number-name ‘seven’. A well-known thesis of Quine’s (1953*a*) is that “‘Cicero’” is a simple symbol: although it appears to contain ‘Cicero’ as a part, it does not really do so, because the reference of “‘Cicero’” is *not* a function of the reference of ‘Cicero’. We will have more to say about this example in Section 2.5, and we return to it again in Chapter 10. If an expression is not simple, then it is *complex*, and it admits of a parsing into function-expression and argument-expression(s). The arithmetic equation (2.2) was a special case in that on each of the proposed analyses the argument-expression was complete. This need not be so. Frege, for example, understands first-order quantifiers to stand for (second-level) functions that map (first-level) functions into truth values, and so a quantificational statement is parsed into a (second-level) function-expression and an argument-expression which is itself a (first-level) function-expression. We will have more to say about the complete/incomplete dichotomy in Chapter 5, but for now we note that the simple/complex distinction cuts across it. A complete expression might be simple or complex, and so too an incomplete expression.

## 2.4 Extensions of the Notation

The mathematical symbolism Frege analyzes is an artificial notation designed to facilitate mathematical reasoning, and it has been constructed with an eye toward maximizing perspicuity, brevity, and precision. The virtues of the symbolism are evident to anyone who tried to work with, say, the English expression ‘the number that is obtained by adding one to the result of multiplying two by one’ instead of ‘ $(2 \cdot 1) + 1$ ’. The English expression is just so unwieldy. Nevertheless, whatever can be expressed in this notation can be expressed in English, for we learn to use the notation by mastering a scheme for associating mathematical symbols with expressions in English. We can regard a mathematical expression and its natural language correlate as notational variants and so transfer Frege’s observations concerning the function/argument structure of the mathematical notation to the structure of the natural language correlates. For example, the English expressions ‘one’ and ‘two’, like the Arabic numerals ‘1’ and ‘2’ with which they are correlated, are simple expressions, no parts of which contribute towards determining the reference of the whole; and corresponding to ‘ $\eta \times \zeta$ ’ we have the English function-expression ‘ $\eta$  times  $\zeta$ ’ from which complex English number-names, like ‘two times one’, can

be constructed. The expression

‘The number which is obtained by adding one to the result  
of multiplying two by one’ (2.11)

is a complex designator of the number three constructed from (say) the function-expression

‘The number which is obtained by adding  $x$  to the result  
of multiplying two by one’ (2.12)

by inserting ‘one’ for the variable. Continuing in this manner, then, we see how that portion of a natural language which serves for discourse about numbers can be analyzed along function/argument lines.

We can go further still. Proper names like ‘Robert’, ‘Winston’, and ‘Paris’ are simple expressions. The name ‘Robert’, for example, contains the name ‘Bert’ as a proper part, but the reference of ‘Bert’ does not contribute toward determining the reference of ‘Robert’. On the other hand, an expression like the definite description

‘Abraham Lincoln’s wife’ (2.13)

is a complex expression. (2.13) stands for Mary Todd Lincoln, Abraham Lincoln’s wife. If we replace ‘Abraham Lincoln’ in (2.13) by ‘George Washington’, we get

‘George Washington’s wife’, (2.14)

which stands for Martha, George Washington’s wife. The descriptions in (2.13) and (2.14) each refer to a unique object (assuming a person can have but one wife). Moreover, the reference of each is dependent solely upon the reference of the constituent name, so that if in (2.13), say, we replace ‘Abraham Lincoln’ by any coreferential singular term, for example, ‘the President of the United States in 1862’, the resulting complex expression stands for the same person as does (2.13), namely, Mary Todd Lincoln. Hence, we can regard each of (2.13) and (2.14) as having been constructed from the function-expression

‘ $x$ ’s wife’ (2.15)

by inserting the appropriate name for  $x$ . We can therefore regard (2.15), then, as standing for

$x$ ’s wife, (2.16)

a function that maps a person into his wife. This is an example of our use of the procedure outlined in the previous section.

However, the most interesting extension of the analysis is to sentences themselves. Consider, again, equation (2.2). We note that the numeral '3' might be replaced by other number-names, and the expression resulting in each case will be a well-formed, senseful, arithmetic equation. For example, replacing '3' by '2 + 1' yields the equation

$$'2 + 1 = (2 \cdot 1) + 1', \quad (2.17)$$

and replacing it by '2' yields the equation

$$'2 = (2 \cdot 1) + 1'. \quad (2.18)$$

Each has a unique truth value: (2.17) is true and (2.18) is false. (2.17) and (2.18) share a common structure with (2.2). Each can be regarded as having been obtained from the expression

$$' \eta = (2 \cdot 1) + 1 ' \quad (2.19)$$

by inserting the appropriate number-name for ' $\eta$ '. However, although (2.19) appears to be an incomplete expression, it is not yet clear whether it is a function-expression. The decision turns on whether we can regard

$$\eta = (2 \cdot 1) + 1 \quad (2.20)$$

as a function; and the problem here is to delineate the range of (2.20), that is, to say what value (2.20) might yield for, say, the argument 2. Now, there is an orderly connection between the reference of the number-name inserted in (2.19) and the truth value of the resultant equation. First, inserting a number-name in (2.19) results in an equation with a unique truth value. Second, the truth value of the equation remains invariant when replacing that number-name by any other naming the same number:  $r('3') = r('2 + 1')$ , and (2.2) and (2.17) are both true. Hence, once again applying the procedure described, we find that (2.19) appears to conform to Principles 2.3.1 and 2.3.3, with (2.20) the designated function, mapping integers into truth values. Completing (2.19) by a number-name results in an equation which designates the value of the function (2.20) evaluated for the argument designated by the inserted number-name. Frege concludes that (2.19) is a function-expression and that (2.2) is therefore a complex referring expression that stands for its truth value. Frege calls the two truth values *the True* and *the False*. (We consider the argument in much greater detail in Chapter 8.)

Once again, Frege transfers the lessons learned about the mathematical notation to natural language. The declarative sentence

‘Abraham Lincoln has red hair’ (2.21)

is complex. The constituent name ‘Abraham Lincoln’ can be replaced by other singular terms, in each case to result in a sentence that has a unique truth value, true if the object named has red hair and false if the object named does not have red hair. The truth value of (2.21) depends solely upon the reference of the constituent name: if we replace ‘Abraham Lincoln’ by any coreferential singular term, the resulting sentence has the same truth value as (2.21). They are both false. So the sentence (2.21) is complex: it is constructed by inserting ‘Abraham Lincoln’ for ‘ $x$ ’ in the function-expression

‘ $x$  has red hair’; (2.22)

and (2.22), then, stands for

$x$  has red hair, (2.23)

a function that maps objects into truth values.

A singulary function like (2.23), whose value for any argument is a truth value, Frege calls a *concept* [*Begriff*]. A binary function whose value for any pair of arguments is a truth value he calls a *relation*.<sup>5</sup>

Of particular importance for the advancement of symbolic logic is Frege’s analysis of the truth-functional connectives and the quantifiers. The declarative sentence

‘It is not the case that Abraham Lincoln has red hair’, (2.24)

contains a whole declarative sentence, (2.21), as a proper part. Moreover, the truth value of (2.24) is uniquely determined by the truth value of (2.21). If we replace the constituent sentence by any other sentence having the same truth value, for example,

‘George Washington has blond hair’, (2.25)

the resulting sentence,

‘It is not the case that George Washington has blond hair’, (2.26)

has the same truth value as does (2.24). Hence, the occurrence of (2.21) in (2.24) is what Quine (1953*b*: 159) calls a *truth-functional* occurrence:

An occurrence of a statement as a part of a longer statement is called *truth-functional* if, whenever we supplant the contained statement by another statement

having the same truth value, the containing statement remains unchanged in truth value.

If we treat a declarative sentence in the way Frege does, namely, as a singular term that stands for its truth value, then Quine's characterization of a truth-functional occurrence of a sentence is one that conforms to Frege's Principle of Extensionality for Reference 2.3.3 where both the argument-expression and the complex expression containing it are sentences. That is, from Frege's perspective, truth-functionality is just a special case of extensionality.

The reference of (2.24) is uniquely determined by the reference of the constituent sentence (2.21), and so we regard (2.24) as a complex constructed by inserting the appropriate sentence for ' $\eta$ ' in the function-expression

‘It is not the case that  $\eta$ ’. (2.27)

The function referred to by (2.27) is

It is not the case that  $\eta$ , (2.28)

a function that maps truth values into truth values. For the argument *true*, (2.28) yields *false* as value, and for the argument *false*, (2.30) yields *true* as value.

The declarative sentence

‘Something has red hair’, (2.29)

though superficially similar to (2.21), receives a quite different analysis.<sup>6</sup> ‘Something’ occupies a position in (2.29) that can be filled by a singular term, and there is thus a temptation to suppose that ‘something’ functions in (2.29) much like ‘Abraham Lincoln’ functions in (2.21). But the absurdity of this suggestion becomes apparent, as Frege showed, if we extended this analogy to ‘nothing’ as it occurs in

‘Nothing has red hair’. (2.30)

For we would then have to say that ‘nothing’ in (2.30) stands for something, namely, nothing. This, among other reasons, led Frege to adopt the view that (2.29) and (2.30) should be understood rather along the lines, respectively, of

‘There is at least one argument for which the function  
 $\eta$  has red hair yields the value *true*’ (2.31)

and

‘There is no argument for which the function  $\eta$  has  
red hair yields the value true’. (2.32)

In Frege’s terminology, a *first-level function* is a function that takes objects as arguments; a *second-level function* is a function that takes first-level functions as arguments.<sup>7</sup> On Frege’s analysis of (2.29) and (2.30), ‘something’ and ‘nothing’ stand for second-level functions that map first-level functions into truth values, and since in each case the value is invariably a truth value, Frege calls these functions *second-level concepts*.

Three points should be noted here. First, Frege was able to handle a problem that remained recalcitrant for those working in the Aristotelian tradition of the categorical statements, namely, the logical analysis of statements involving relations. His seamless treatment of concepts and relations enabled formal logical analysis of statements like ‘Every number is greater than some number’, statements of *multiple generality*, an essential first step to any formalization of arithmetic proofs. Second, Frege provided an account of the quantifiers that also seamlessly generalized from first order, to second order, to  $n$  order. In point of fact, Frege (1879) attached little importance to the distinction between first-order and higher-order quantification.<sup>8</sup> His own characterization of the axioms and rules governing the German gothic letters he used for bound variables, although expressed using individual bound variables, was intended to serve for variables of any kind.<sup>9</sup> As a result, the rules Frege (1879) stated for first-order quantifiers were intended to be entirely general and serve for quantifiers of any order. Third, Frege (1879) took (2.30) to be the proper analysis of

‘A red-haired thing exists’. (2.33)

On Frege’s view, *existence* is a second-level concept. We will examine this very influential view in Chapter 7.

We have not nearly exhausted the results Frege obtained by means of this function/argument analysis of language. To continue recounting them, however, would be unnecessary, for enough has been presented for the reader to appreciate the powerful tool it represents for the investigation of language. Much of the power and generality of this function/argument analysis arises from Frege’s inclusion of declarative sentences along with proper names and definite descriptions in the category of complete expression. In a sense, there is nothing new in construing a sentence as a kind of name which stands for something. The idea is



a familiar one in philosophy, only it has been more common to suggest states of affairs, thoughts, or propositions as the sort of item named. Hence, what is so surprising in Frege's theory is not that sentences stand for anything, but that they stand for truth values. Nevertheless, Frege has supplied a rather strong argument for supposing that a sentence stands for a truth value if it refers at all; we have given the argument in the analysis of arithmetic equations, and we shall state it more sharply again when we have occasion to examine it closely. Moreover, Frege has also provided reasons for supposing that a sentence does stand for something, not the least of which is the unified theory of language that emerges on this supposition. For example, supposing that a sentence stands for its truth value, we can regard a property like *having red hair* as a function that maps objects into truth values.<sup>10</sup>

## 2.5 The Substitution Principle for Reference

Let us standardize some notation. Where  $S\alpha$  is a sentence containing the singular term  $\alpha$ ,  $S\alpha/\beta$  results upon replacing  $\alpha$  at one or more of its occurrences in  $S\alpha$  by the singular term  $\beta$ ;<sup>11</sup> and where  $\eta$  is any referring expression, we continue to employ  $r(\eta)$  for the entity it refers to.

As a special case of Principle 2.3.3 when the complex expression is a sentence, Frege's function/argument analysis of sentences yields the following well-known substitution principle:

**PRINCIPLE 2.5.1 (SUBSTITUTION FOR REFERENCE)** *If  $r(\alpha) = r(\beta)$ ,  $S\alpha$  and  $S\alpha/\beta$  have the same truth value.*

Underlying this substitution principle is a fundamental logical law: if an object  $x$  is identical with an object  $y$ , then any property of  $x$  is equally a property of  $y$  and conversely. This law, which is sometimes picturesquely referred to as *The Indiscernibility of Identicals*, is expressed in first-order logic as follows:

**PRINCIPLE 2.5.2 (LEIBNIZ'S LAW)**  $(\forall x)(\forall y)(x = y \supset (Fx \equiv Fy))$ .

Principle 2.2.1 (*The Fundamental Property of Functions*) bears a striking resemblance to Principle 2.5.2 (*Leibniz's Law*). Frege's insightful suggestion was to supplant the traditional property/object ontology underpinning Principle 2.5.2 with the function/object ontology of Principle 2.2.1. Mathematically, the fit is stunning. Overwhelmingly so. Philosophically, these abstract functions like *having red hair* posed something of a challenge in his day, but they certainly introduced an exciting new

perspective on a classical topic that has become commonplace in the current environment in which we talk of sets of objects.

Principle 2.5.1 and Principle 2.5.2 are both commonly referred to in the literature as *Leibniz's Law*. But Principle 2.5.1 and Principle 2.5.2 are distinct principles, and only Principle 2.5.2 merits being called a *law*. Principle 2.5.2 cannot plausibly be denied. For, intuitively, to suppose that  $x$  is one and the same thing as  $y$ , and yet to suppose, further, that  $x$  has a property  $y$  lacks (or conversely), is, in effect, to suppose that one and the same thing both has and lacks the given property. Principle 2.5.2, then, is on the same footing as the venerable *Law of Noncontradiction*,  $(\forall x) \neg (Fx \wedge \neg Fx)$ . Principle 2.5.1, on the other hand, is not so well grounded.

Here is a counterexample to Principle 2.5.1 that Quine (1951) made famous.<sup>12</sup>

$$\text{Cicero} = \text{Tully}, \quad (2.34)$$

and

$$\text{'Cicero' has six letters}, \quad (2.35)$$

are both true. Hence,  $r(\text{'Cicero'}) = r(\text{'Tully'})$ . We replace 'Cicero' by 'Tully' in (2.35), and the result, by Principle 2.5.1, ought to be true. But,

$$\text{'Tully' has six letters} \quad (2.36)$$

is false. 'Tully' only has five letters.

Quine's case is a counterexample to Principle 2.5.1 but not to Principle 2.5.2. It would be a counterexample to Principle 2.5.2 only if (2.35) and (2.36) both ascribed the same property to the same object, one truly and the other falsely. But on the most natural interpretation, these two sentences ascribe the same property to different objects. For, while (2.35) says that the name 'Cicero' has six letters, (2.36) says that a different name, 'Tully', has six letters. Now a man is not, in general, identical with his name (or names). Although (2.34) is true,

$$\text{'Cicero'} = \text{'Tully'} \quad (2.37)$$

is false. Were (2.37) true, that is, were the name 'Cicero' identical with the name 'Tully', then, according to Principle 2.5.2, any property of the one would have to be a property of the other. But, (2.37) is not true, and certainly Principle 2.5.2 does not require that distinct objects should have all their properties in common. Since the name 'Cicero' is not the same object as the name 'Tully', there is no conflict with Principle 2.5.2

arising from the fact that ‘Cicero’ has a property, namely, *having six letters*, which ‘Tully’ lacks.

We have supplied an interpretation on which Quine’s case is seen to be consistent with Principle 2.5.2, but we have not thereby shown that there is no interpretation on which it would conflict with it. Further assurance might be requested. But what reason remains for persisting in this direction? Only, as far as can be determined, the occurrence of the name ‘Cicero’ in (2.35) and the name ‘Tully’ in (2.36). For one might be inclined to suppose that (2.35) must therefore be about Cicero and (2.36) about Tully. The operative assumption here is that if a sentence contains a singular term, then that sentence must be *about* that which the singular term stands for, either ascribing a property to it or expressing that it is one of the relata in some  $n$ -ary relation. In short, the assumption is

PRINCIPLE 2.5.3 (ABOUTNESS)  $S\alpha$  is about  $r(\alpha)$ .

Assuming Principle 2.5.3, sentence (2.35) receives the following analysis: the name ‘Cicero’ stands for the man Cicero, and the remainder of the sentence, namely,

‘ ’ has six letters, (2.38)

stands for the property ascribed to Cicero. Similarly for (2.36): the property (2.28) refers to is ascribed to the man Tully. So, since (2.34) is true, we conclude that (2.35) and (2.36) both ascribe the same property to the same object.

Whether we have a counterexample to Principle 2.5.2, however, is still unclear, for what has yet to be shown is that (2.35) and (2.36) differ in truth value on this interpretation. We had originally agreed that (2.35) was true and (2.36) false, but that was based on the original interpretation and it cannot, without argument, be assumed here. Until we are told what property (2.38) is alleged to stand for, then, we must suspend judgment about truth value. But this is a minor point. The real interest in the case attaches to Principle 2.5.3, because Principle 2.5.3 is the link between Principle 2.5.1 and Principle 2.5.2: if Principle 2.5.2 and Principle 2.5.3 are both true, then Principle 2.5.1 must also be true. And, since we have established that Principle 2.5.1 is false, either Principle 2.5.2 or Principle 2.5.3 – possibly both – must be false. Now, what has been suggested in the last paragraph is that Principle 2.5.3 is true, and thus Principle 2.5.2 is false. But Principle 2.5.2, as we saw earlier, is among the most fundamental of logical principles. Hence, if we are to reject Principle 2.5.2 in favor of Principle 2.5.3, it would seem that Principle 2.5.3 would have to be at

least as basic. Yet it is doubtful that Principle 2.5.3 is even true. The term ‘cicerone’ contains the name ‘Cicero’ as a proper part; nevertheless, we should hardly suppose that the sentence

She tipped the cicerone 100 lire (2.39)

is about the great roman orator.

The notation for natural language is not so uniform that a given sequence of letters must always serve the same function in every context in which it occurs – not even that the different functions served by a given sequence of letters need be related one to the other. Thus the sequence of letters which forms an initial segment of the term ‘cicerone’ serves in other contexts (for example, in (2.34)) to stand for the Roman; but that is accidental to its occurrence in ‘cicerone’. Again, the name ‘Cicero’ surely occurs in (2.35), but it does not stand for the man in that context any more than it does in (2.39). Rather, it is exhibited, as the single quote’s indicate, and (2.35) is about the name, not the man. The only reasonable course, then, is to reject the unrestricted Principle 2.5.3 and the purported counterexample to Principle 2.5.2. (Other familiar counterexamples to Principle 2.5.1 of the sort where a singular term occurs in a clause governed by a verb of propositional attitude, or by a modal operator, can, by parallel reasoning, also be shown not to violate Principle 2.5.2.)

We see from this example that a singular term occurring in a sentence need not be serving in that sentence simply to stand for what it ordinarily designates. If so, we ought not to expect that the sentence is about that which the singular term ordinarily designates. If  $S\alpha$  is not about  $r(\alpha)$ , then we cannot expect the truth value of  $S\alpha$  to depend simply on  $r(\alpha)$ . We therefore cannot expect  $S\alpha/\beta$  to have the same truth value as  $S\alpha$  even though  $r(\alpha) = r(\beta)$ . This analysis of the purported counterexample to Principle 2.5.2, which is essentially due to Frege, does, however, have a constructive side. For, in exposing the assumption Principle 2.5.3, the nature of the connection between Principle 2.5.1 and Principle 2.5.2 has been revealed, and thus the needed correction for Principle 2.5.1 becomes apparent. If  $S\alpha$  is about  $r(\alpha)$ , then it would seem that  $S\alpha$  and  $S\alpha/\beta$  must agree in truth value whenever  $r(\alpha) = r(\beta)$ . As Quine (1953a: 140) remarks, “whatever can be affirmed about [an] object remains true when we refer to the object by any other name.” So we replace Principle 2.5.1 with

PRINCIPLE 2.5.4 (CORRECTED SUBSTITUTION FOR REFERENCE) *If  $S\alpha$  is about  $r(\alpha)$ , then if  $r(\alpha) = r(\beta)$ , then  $S\alpha$  and  $S\alpha/\beta$  have the same truth value.*

Thus substitutivity *salva veritate* is a necessary condition for  $S\alpha$  to be about  $r(\alpha)$ .<sup>13</sup>

## 2.6 Formal Mode and Material Mode

Principle 2.5.2 (*Leibniz's Law*) is a fundamental logical law governing identity and, as we have just seen, admits of no counterexamples. If someone purports to have a case where an object  $x$  is identical with an object  $y$  but  $x$  has a property  $y$  lacks, we reply that this person is mistaken, either in supposing that  $x$  is identical with  $y$  or in supposing that the property  $x$  is said to possess is the same property as the property  $y$  is said to lack. We do not admit counterexamples to Principle 2.5.2 because Principle 2.5.2 is, in effect, our guide in speaking of same object and same property. There are many examples in the philosophical literature of such use of Principle 2.5.2. We cite here two.

- I. A penny, when viewed from one perspective, appears to be circular, but, when viewed from another perspective, does not appear to be circular. By Principle 2.5.2, it cannot be that the same thing was viewed each time, for that which was viewed the first time appeared to be circular, while that which was viewed the second time did not appear to be circular. Since that which was viewed the first time is not identical with that which was viewed the second time, then, by elementary properties of identity, it follows that on at least one of these two occasions, that which was viewed was not identical with the penny. This argument is often invoked by Sense-datum Theorists.
- II. John utters a sentence,  $\sigma$ , and what he says is true, while Tom utters the very same sentence  $\sigma$  and what he says is not true. Since that which John said is true and that which Tom said is not true, then by Principle 2.5.2, it cannot be the case that that which John said is the same thing as that which Tom said. And, since each uttered the very same sentence  $\sigma$ , it follows, again, by elementary properties of identity, that on at least one of these two occasions, that which was said was not the sentence  $\sigma$ , and so, on at least one of these two occasions, it was something other than the sentence  $\sigma$  that was said either to be true or to be false. This argument is commonly invoked by Propositionalists.

It is important to note, however, that in each of the above arguments it was assumed that the property  $x$  was said to possess was the same property as the property  $y$  was said to lack. All Principle 2.5.2 establishes is a certain connection between objects and properties so that in example I, for example, it cannot be *both* that the same thing was viewed each time *and also* that the property the thing viewed the first time was said to possess was the same property as the property the thing viewed the second time was said to lack. Again, in example II, it cannot be *both* that John and Tom said the same thing *and also* that what John said was true and what Tom said was not true. Principle 2.5.2 is indifferent to which alternative is rejected in each case. It requires that one (at least) be rejected, but the justification for rejecting one rather than the other must be sought elsewhere. So, a Naive Realist might attempt to deal with the phenomenon described in example I and remain consistent with Principle 2.5.2 by urging that the same thing was viewed each time, namely, the penny, but that the property said to be possessed by the penny when viewed the first time is not the same property as the property said to be lacked by the penny when viewed the second time. He might go on to claim that the property ascribed to the penny the first time was not *appearing circular* but rather *appearing circular from perspective  $p_1$* , and the property the penny was said to lack the second time was not *appearing circular* but *appearing circular from perspective  $p_2$* . Similarly, a Nominalist might respond to the case described in example II by urging that John and Tom did say the same thing, namely the sentence  $\sigma$ , but that truth must be relativized to a language, a time, a speaker, and a context.

The corrected Substitution Principle 2.5.4 is understood to be the formal mode analogue to Principle 2.5.2. Singular terms stand for objects and predicates stand for properties, so if  $S\alpha$  expresses that  $r(\alpha)$  has a given property, then if  $r(\alpha) = r(\beta)$ , then  $S\alpha/\beta$  expresses that  $r(\beta)$ , that is,  $r(\alpha)$ , has the given property. Now, just as Principle 2.5.2 lays down a connection between objects and properties that grounds our talk of same object and same property, so Principle 2.5.4 is intended to lay down a connection between singular terms and predicates to ground our talk of same singular term and same predicate. Purported counterexamples to Principle 2.5.4 are therefore regarded as exhibiting ambiguity, as, for example, in the Tully/Cicero case earlier where we saw that ‘Cicero’ stood for the man in (2.34) but not in (2.35). On any adequate symbolization, different symbols would be introduced to record the difference in function. Moreover, the indifference of Principle 2.5.2 to the competing Sense-datum and Naive Realist accounts of the phenomenon described in

example I is reproduced by Principle 2.5.4 at the formal-mode level. For, consider the two sentences

That which is viewed from perspective  $p_1$  appears circular      (2.40)

and

That which is viewed from perspective  $p_2$  appears circular.      (2.41)

As the case was presented in example I, (2.40) was true and (2.41) was false. Now, the Sense-datum theorist holds that (2.40) and (2.41) ascribe the same property to different objects. At the formal-mode level he might be characterized as splitting (2.40) into singular term and predicate as follows:

— appears circular  
that which is viewed from perspective  $p_1$       (2.42)

so that (2.41) would be the result of replacing the singular term ‘that which is viewed from perspective  $p_1$ ’ in (2.40) by the singular term ‘that which is viewed from perspective  $p_2$ ’. And since (2.40) is true and (2.41) is false, he concludes, by Principle 2.5.4, that

$r$ (‘that which is viewed from perspective  $p_1$ ’)  $\neq$   $r$ (‘that  
which is viewed from perspective  $p_2$ ’).      (2.43)

The Naive Realist, on the other hand, wishes to hold that the same thing was viewed each time, namely, the penny. To remain consistent with Principle 2.5.2, he holds that the property ascribed to the penny in (2.40) is not the same property as that ascribed to the penny in (2.41). Therefore, at the formal mode level he might be characterized as splitting (2.40) into singular term and predicate in the following way:

— from perspective  $p_1$  appears circular  
that which is viewed      (2.44)

so that (2.41) is not related to (2.40) as  $S\alpha/\beta$  to  $S\alpha$ . The Sense-datum theorist, then, takes (2.40) and (2.41) to contain the same predicate but different singular terms, whereas the Naive Realist takes (2.40) and (2.41) to contain the same singular term but different predicates.

The picture we get is that Principle 2.5.4 is a device for parsing a sentence into singular term and predicate so that the structure of the sentence mirrors the structure of the world – or, more accurately, given the indeterminacy just exhibited, so that it mirrors a coherent structuring of the world. The division in language therefore has ontological significance:

the connection between the function/argument structure and Leibniz's Law provides the basis for identifying objects and properties. But the division in language also has semantic significance: the function/argument structure enables a computation of the truth value of a complex sentence given the reference of its parts, and this provides a basis for capturing the generative truth conditions for the sentence. The function/argument structure therefore gave new vigor to this heady combination of ontology and semantics that lies at the core of the most speculative metaphysical programs: determining what there must be in order to account for human understanding. Still, all of this machinery operates at the level of reference. The sense of an expression is not determined by its reference. The relationship, if any, is precisely in the opposite direction. But telling this part of the story will occupy us in future chapters.

Frege's analysis of language is part of a large movement in Western intellectual history that occupied some of the greatest mathematical minds of the nineteenth century. This was to abstract from and universalize the procedures of arithmetic and geometry and apply them to areas that were not usually thought of as being amenable to such types of analysis. It was this movement that broadened the study of mathematics beyond arithmetic and geometry to abstract structures of any kind. Language is the structure that consumed Frege's interests. His genius was to see in natural language a formal abstract syntax whose grammar inherits its logic from the function/argument structure by which it represents reality. It is no accident that Noam Chomsky (1957) should hark back to Frege's (1923) paeon to the creativity of the devices of language in motivating his generative approach to grammar.



## Sense and Reference

### 3.1 Introduction

Few texts are as well known to modern philosophers as Frege's (1892c: 151–2) opening paragraph:

Equality gives rise to challenging questions which are not altogether easy to answer. Is it a relation? A relation between objects, or between names or signs of objects? In my *Begriffsschrift*, I assumed the latter. The reasons which seem to favour this are the following:  $a = a$  and  $a = b$  are obviously statements of differing cognitive value. . . . Now if we were to regard equality as a relation between that which the names ' $a$ ' and ' $b$ ' designate [*bedeuten*], it would seem that  $a = b$  could not differ from  $a = a$  (i.e. provided  $a = b$  is true). A relation would thereby be expressed of a thing to itself but to no other thing. What is intended to be said by  $a = b$  seems to be that the signs or names ' $a$ ' and ' $b$ ' designate [*bedeuten*] the same thing, so that those signs themselves would be under discussion; a relation between them would be asserted. But this relation would hold between the names or signs only in so far as they named or designated something. It would be mediated by the connexion of each of the two signs with the same designated thing. But this is arbitrary. Nobody can be forbidden to use any arbitrarily producible event or object as a sign for something. 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. If the sign ' $a$ ' is distinguished from the sign ' $b$ ' only as object (here, by means of its shape), not as sign (i.e., not by the manner in which it designates something), the cognitive value of  $a = a$  becomes essentially equal to that of  $a = b$ , provided  $a = b$  is true. A difference can arise only if the difference between the signs corresponds to a difference in the mode of presentation [*Art des Gegebenseins*] of that which is designated.

Yet this passage is easily misunderstood. The fault lies in part with Frege, who failed to distinguish for his readers (1) the error he locates in the

paradox (namely, the underlying substitution principle), and (2) the error (if any) in taking identity to relate expressions. Two misconceptions, in particular, need to be corrected:

- It is thought that Frege held the view that “identity relates expressions” because he had not identified a notion of sense. This is false. Frege had a very clear notion of sense in *Begriffsschrift* for singular terms. What he lacked was a clear understanding of the structural role sense was to play in his semantic story.
- It is thought that there is something logically wrong with the view that “identity relates expressions.” This is false. There is nothing logically wrong with that view. What is wrong is the view that “identity relates expressions *and not objects*.”

Our discussion of this paradox will be in two parts. In this chapter, we will examine the paradox itself, discuss and contrast the solutions offered by Frege and Russell, and then outline the general features of the sense/reference theory. In the next chapter, we will revisit Frege’s (1879) account of identity, examining why he was led to hold it and what is wrong with it.

### 3.2 The Paradox of Identity

Identity statements differ in “cognitive value” [*Erkenntniswerth*]. Here is a simple example. ‘Mark Twain = Mark Twain’ is a mere truism, but ‘Mark Twain = Samuel Clemens’ says something of considerable historical significance. How does this fact challenge the standard view, on which  $\alpha = \beta$  is understood to express that the relation *being one and the same thing as* holds between the objects designated by  $\alpha$  and  $\beta$ ? Since the relation is supposed to hold between the objects themselves, all that  $\alpha = \beta$  expresses – the cognitive content of the sentence – is that the objects stand in the given relation.  $\alpha = \beta$  and  $\gamma = \delta$  ( $\alpha, \beta, \gamma, \delta$  not necessarily distinct) all say the same thing – have the same cognitive content – if  $\alpha, \beta, \gamma, \delta$  all stand for the same object. For the same relation is said to hold between the same objects. So, ironically, on the view that  $\alpha = \beta$  is about the objects designated by  $\alpha$  and  $\beta$ , the identity, if true, appears less a significant remark about the designated objects(s) than a trivial rehearsal of the Law of Identity. This is Frege’s *Paradox of Identity*.

## 3.3 The Sharpened Paradox

We are still using the notation introduced in Section 2.5. Where  $\eta$  is any referring expression,  $r(\eta)$  is the *reference of*  $\eta$ ; and where  $S\alpha$  is a sentence containing the singular term  $\alpha$ ,  $S\alpha/\beta$  is a sentence that we obtain by replacing  $\alpha$  at one or more of its occurrences in  $S\alpha$  by  $\beta$ . In the argument above, Frege assumes that if we understand  $\alpha = \beta$  to express that a relation holds between  $r(\alpha)$  and  $r(\beta)$ , then the way in which the objects are specified is irrelevant to the cognitive content of  $\alpha = \beta$ . If we replace  $\alpha$  (or  $\beta$ ) by a coreferential singular term, the resulting sentence must have the same cognitive value as the original. The assumption here is not peculiar to identities, but reflects a general view about the relation between the reference of a term and the cognitive content of the sentence containing it. The assumption goes like this: if  $S\alpha$  is genuinely about  $r(\alpha)$ , that is, if  $S\alpha$  ascribes a property to  $r(\alpha)$  or if  $S\alpha$  expresses that  $r(\alpha)$  stands in a particular relation, then the cognitive content of  $S\alpha$  simply is  $r(\alpha)$ 's having the given property or  $r(\alpha)$ 's standing in the given relation. Only the object  $r(\alpha)$ , not the term  $\alpha$ , is part of the content of  $S\alpha$ . So, if we replace  $\alpha$  at one or more of its occurrences in  $S\alpha$  by any coreferential singular term, the resulting sentence must have the same cognitive value as  $S\alpha$ . This is the generalized substitution principle:

**PRINCIPLE 3.3.1 (BEGRIFFSSCHRIFT SUBSTITUTION)** *If  $S\alpha$  is about  $r(\alpha)$ , then if  $r(\alpha) = r(\beta)$ ,  $S\alpha$  and  $S\alpha/\beta$  have the same cognitive value.*

With Principle 3.3.1 in hand, Frege's argument is easily shown to be valid. Suppose that  $\alpha = \beta$  is about  $r(\alpha)$  and  $r(\beta)$ , and consider our two identities,

$$\text{Mark Twain} = \text{Samuel Clemens} \quad (3.1)$$

$$\text{Mark Twain} = \text{Mark Twain.} \quad (3.2)$$

Sentence (3.1) is true, so  $r(\text{'Mark Twain'}) = r(\text{'Samuel Clemens'})$ . Since we obtain (3.2) from (3.1) by replacing 'Samuel Clemens' by 'Mark Twain', by the *Begriffsschrift* Substitution Principle 3.3.1, the two sentences (3.1) and (3.2) must have the same cognitive value. This argument does not depend upon any characteristic of the particular names chosen, so we have quite generally that  $\alpha = \alpha$  and true  $\alpha = \beta$  cannot differ in cognitive value.

### 3.4 The Generalized Paradox

The paradox is not, however, restricted to identity. Similar arguments are easily devised for other commonplace properties and relations.<sup>1</sup> The sentence

Mark Twain wrote *Innocents Abroad*, (3.3)

for example, ascribes the property of having written *Innocents Abroad* to Mark Twain. By Principle 3.3.1, if we replace 'Mark Twain' in (3.3) by any coreferential singular term, the cognitive value of the sentence should remain unchanged. Now, 'the person who wrote *Innocents Abroad*' is such a coreferential singular term, for not only is it true that Mark Twain wrote *Innocents Abroad*, but he was the only person to have done so. Yet

The person who wrote *Innocents Abroad* wrote *Innocents Abroad*, (3.4)

if not a truism, is so nearly so as to clearly differ in cognitive value from (3.3). In this example, we chose the substituted singular term carefully in order to parallel Frege's argument, wherein he had transformed an informative sentence into a trivial one. But there is no need to cleave to this format. For the same intuitions that led us to regard (3.1) as an informative identity should also lead us to deny that

Samuel Clemens wrote *Innocents Abroad* (3.5)

has the same cognitive value as (3.3). The nontrivial character of (3.1) goes hand in hand with (3.5)'s telling us something that (3.3) does not, and conversely.<sup>2</sup>

While identity does, as Frege says, challenge reflection, there is nothing peculiar about identity unearthed here. The root of the paradox lies with an erroneous conception about the cognitive content of a sentence, which just happens to show its virulence in the case of identity.

### 3.5 Three Solutions

There are two assumptions in the argument:

- The *Begriffsschrift* Substitution Principle 3.3.1, and
- The view that Identity Relates Objects.

These two lead to the conclusion that there are no informative identities.

Frege proposed two distinct solutions to the paradox at different points in his career. Each was based on the belief that identities like (3.1) and (3.2) differ in cognitive value. Each therefore regarded the paradox

as a *reductio* of one or the other of the above assumptions. The early Frege (1879) regarded the paradox as a *reductio* of the second assumption, namely, that “identity relates objects.” The later Frege (1892*c*), however, regarded it as a *reductio* of the first assumption, namely, the Substitution Principle 3.3.1, according to which cognitive value is held to remain invariant under substitution of coreferential singular terms.

Russell (1905) proposed a third way out of the paradox. Like the early Frege (1879), he was committed to the Substitution Principle 3.3.1. Like the later Frege (1892*c*), he was committed to the idea that identity related objects. How then did he block the unwanted conclusion of the paradox that there are no informative identities? Unlike Frege, early or late, he sharply differentiated expressions that had otherwise been lumped together under the category of proper names. In particular, Russell (1905) transferred definite descriptions – expressions that he and Frege had earlier considered to be operating logically as full-fledged individual constant terms – out of the category of proper names and into the category of quantified expressions. It is these definite descriptions, he believed, that generated the informative identities. Russell (1905) denied that when identities like (3.1) and (3.2) differed in cognitive value, they were of the form  $S\alpha$ ; accordingly, the one cannot be derived from the other by substitution using Principle 3.3.1. Genuine proper names *are*, however, subject to Principle 3.3.1, and in their case the conclusion of the paradox is not blocked: there are no informative identities involving proper names. We will talk about Russell’s theory in much greater detail in Chapters 6 and 7.

We will canvass these three solutions very briefly.

#### *The Begriffsschrift Solution*

Frege (1879) did not explain *why* he chose to take identity as a relation between expressions. We are left in the dark until Frege (1892*c*). There Frege tells us that reflection on the paradox had convinced him that the information conveyed by an identity could not be about the objects themselves, for then each true identity would reduce to conveying the trivial information that the designated object is self-identical. If a true identity of the form  $\alpha = \beta$  is to be informative, its informativeness must reside in the fact that the different expressions  $\alpha$  and  $\beta$  turn out to stand for the same thing:

What we apparently want to state by  $a = b$  is that the signs or names ‘*a*’ and ‘*b*’ designate [*bedeuten*] the same thing, so that those signs themselves

would be under discussion; a relation between them would be asserted. (Frege 1892c: 151)

For these reasons, Frege says, he had in *Begriffsschrift* rejected the view that  $\alpha = \beta$  expresses that the relation *being one and the same thing as* holds between the objects designated by  $\alpha$  and  $\beta$ , and maintained instead that  $\alpha = \beta$  expresses that a relation holds between the expressions  $\alpha$  and  $\beta$  themselves, namely, the equivalence relation *designating one and the same thing as*.<sup>3</sup>

The *Begriffsschrift* solution to the paradox is a bad solution because it singles out identity as the source of the difficulty when, as we have seen, the problem is far wider than that. The solution for identity must either be generalized (all properties and relations are really about words) or dropped. There remain two questions of interest about the solution that we will look at closely in the next chapter. First, if Frege was moved by the paradox to adopt the view he did, then he must have had a view about conceptual content that entailed a substitution principle like Principle 3.3.1: What was it? Second, there have been rumblings in the literature that the view that identity relates expressions is incoherent: Is it?

#### *The Sinn/Bedeutung Solution*

Although  $S\alpha$  is (apparently) about  $r(\alpha)$ , and although  $r(\alpha) = r(\beta)$ , it need not be that  $S\alpha$  and  $S\alpha/\beta$  have the same cognitive value. The moral Frege (1892c) drew, as we all know, was to abandon the semantic theory of his *Begriffsschrift* days and the Substitution Principle that generated the paradox. The *Begriffsschrift* Substitution Principle 3.3.1 represents, for the mature Frege, a hybrid: the *reference* of the parts is connected with the *sense* of the complex. This is now split up into two distinct substitution principles. Substitution of coreferential singular terms preserves *truth value*, not, as had been previously thought, *cognitive value*; this, that is, cognitive value, is preserved under substitution of singular terms having the same sense. We present the main outlines of the theory later in this chapter.

#### *Russell's 1905 Solution*

Along with the author of *Begriffsschrift*, Russell (1905) holds firm to the Substitution Principle 3.3.1. Russell is explicit that a predicate-expression designates a *propositional function*, not, as Frege (1892c) would have it, a truth function. This is the disagreement he clearly enunciates in the exchange of letters with Frege. The moral Russell draws from the fact

that  $S\alpha$  and  $S\alpha/\beta$  differ in cognitive value is that  $\alpha$  and  $\beta$  cannot both be genuine names of the same object! Russell denies that true  $\alpha = \beta$  can be informative when the expressions flanking the identity symbol are *logically proper names*. If a sentence is genuinely of the form  $S\alpha$ , the object itself is a constituent of the *singular proposition* it expresses. The meaning of a logically proper name is the object it stands for, so no other name of the same object can possibly induce any change in meaning.

When one or both sides of the identity sign are flanked by definite descriptions, however, the sentence can, on Russell's view, be informative. For the meaning of a definite description is not exhausted by the object it denotes. This is what Russell means when he says that a definite description has no meaning "in isolation:" a sentence containing a definite description is not of the form  $S\alpha$ , the proposition expressed by a sentence containing a definite description is *not* a singular proposition, and there is no object denoted by the definite description that is a constituent of the proposition. The contribution a definite description makes to the proposition it is used to express is rather whatever property the descriptive predicate stands for. So the Substitution Principle 3.3.1 is preserved: substitution of coreferential names conserves cognitive content. Definite descriptions, however, are only apparently names and they are not governed by this substitution principle.<sup>4</sup> Russell's, quite clearly, is a very different semantics from Frege's. Frege's two-level theory separates out the incomplete character of the predicate from its informativeness; Russell's single-level theory runs the two together. The failure to observe this distinction has been a consistent barrier to the clear understanding of Frege's views.<sup>5</sup> We will turn to Russell's theory in Chapter 6.

### 3.6 Sense and Reference

Here is a brief overview of the sense/reference distinction and, especially, of the notion of sense. On the matter of sense, we can only be brief because Frege did little to develop the notion. In his logical work, he was primarily interested in the reference of expressions, and, for him, the main virtue of clarifying this distinction was to assure that the notion of *reference* was not compromised by matters that properly belonged elsewhere.

The reference of a singular term is the object for which it stands, either by having been assigned to that object or by uniquely describing it. 'The Stagirite', 'Aristotle', 'Plato's greatest pupil', and 'the teacher of Alexander the Great' all refer to the same object, namely, Aristotle. But they do not all have the same sense. The sense of a singular term is

that “wherein the mode of presentation is contained,” and it thus carries the burden of introducing, presenting, or picking out the referent. ‘The Stagirite’, for example, picks out Aristotle as having been Stagira’s most famous native son, while ‘Plato’s greatest pupil’ picks him out as having studied with Plato and as having been his finest student. Both of these terms stand for Aristotle, but since they pick Aristotle out in different ways, they do not have the same sense. To take another example, compare ‘the number which is obtained by adding 2 four times’ with ‘the number which is obtained by adding 4 twice’. Both terms stand for the number 8, but the first instructs us to take 2 four times, that is,

X X  
X X  
X X  
X X

while the second instructs us to take 4 twice, that is,

X X X X  
X X X X

and so they do not have the same sense.

It is important to Frege’s notion that any picking out of an object is from a perspective, but it need not be that this perspective be that of some definite description. The famous footnote in Frege (1892c: 153) certainly indicates that the sense can, in certain circumstances, be that of a definite description:

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

But Frege was also clear that a thing might be given immediately in experience [*Anschaung*] without its being rendered completely transparent to us, for example, a geometric shape.<sup>6</sup> An identity can be informative even when one of the expressions presents its referent immediately like this.<sup>7</sup>

Where  $\eta$  is an expression, let  $s(\eta)$  be *the sense of*  $\eta$ . Then we can identify the following principles governing the relation between an expression’s



sense and its reference. First, Frege clearly believes that sense determines reference:<sup>8</sup>

PRINCIPLE 3.6.1 (SENSE DETERMINES REFERENCE)  $r(\eta) = r(s(\eta))$ .

Second, he clearly believes that the reference determined in this way is unique, that is, that any two terms having the same sense refer to the same object:

PRINCIPLE 3.6.2 (REFERENCE IS A FUNCTION) *If*  $s(\eta) = s(\zeta)$ , *then*  $r(\eta) = r(\zeta)$ .

The sense of an expression is that which is communicated or conveyed by the expression, the information it contains. The sense of an expression is not material, nor is it perceptible, but it is an objective entity nonetheless that exists independent of any individual's consciousness. "For one can hardly deny that mankind has a common store of thoughts which is transmitted from one generation to another" (Frege 1892c: 154); but "if every thought requires an owner and belongs to the contents of his consciousness, then the thought has this owner alone; and there is no science common to many on which many could work. . . ." (Frege 1918: 336). He is especially anxious to distinguish senses, which are objective, from ideas [*Vorstellungen*], which are private to each individual:

It is so much of the essence of any one of my ideas to be a content of my consciousness, that any idea someone else has is, just as such, different from mine. (Frege 1918: 335)

For communication would be impossible:

If every man designated something different by the name "moon", namely, one of his own ideas, . . . an argument about the properties of the moon would be pointless: one person could perfectly well assert of his moon the opposite of what the other person, with equal right, said of his. If we could not grasp anything but what was within our own selves, then a conflict of opinions [based on] a mutual understanding would be impossible, because a common ground would be lacking, and no idea in the psychological sense can afford us such a ground. (Frege 1893: 17)

And he would be unable to protect mathematics from the error of psychologism that ever threatens to engulf it:

[F]or me there is a domain of what is objective, which is distinct from that of what is actual, whereas the psychological logicians without ado take what is not actual to be subjective. . . . Because the psychological logicians fail to recognize the

possibility of there being something objective that is not actual, they take concepts to be ideas and thereby consign them to psychology. (Frege 1893: 15–16)

When a person grasps [*fassen*] a sense, “there must be something in his consciousness that is aimed at [it]” (Frege 1918: 342). Nevertheless, “he does not create it but only comes to stand in a certain relation to what already existed – a different relation from seeing a thing or having an idea” (Frege 1918: 337). Senses, then, belong neither in the outer world of material entities nor in the many private inner worlds of psychological entities, but in a specially designated Third Realm [*dritte Reich*].<sup>9</sup>

The sense/reference distinction is partly an ontological distinction, but it is also a distinction between the ways in which entities are related to signs: an entity might be referred to by a sign, or it might be expressed by the sign, or it might be associated with the sign. ‘Idea’ and ‘sense’ serve to label particular types of entities, and it appears that ideas are the only sort of things that can be associated with an expression and that sense are the only sorts of things that can be expressed by an expression. ‘Reference’, however, carries no such implications: the reference of an expression is simply that which we use the expression to talk about, and insofar as we talk about ideas and senses as well as all the usual things, we use expressions that refer to them. A’s idea of the moon is not the reference of the expression ‘the moon’, but presumably that which A associates with the expression; on the other hand, A’s idea of the moon is the reference of the expression ‘A’s idea of the moon’, and yet another idea might be associated with this expression. Again, the sense of the expression ‘the Morning Star’ is the sense of ‘the Morning Star’, but it is the reference of ‘the sense of “the Morning Star”’, and with this expression yet another sense would be expressed. However, although Frege does not explicitly say so, it is usually supposed that the sense, the reference, and the associated idea of a given expression must be distinct, and that no expression can both refer to and express one and the same entity.<sup>10</sup>

A sequence of noises or of marks on paper, if it is to be a word or phrase – what Frege calls a “sign” – must have a sense; but it need not follow that any reference corresponds to it. For example, ‘the celestial body most distant from the Earth’ and ‘the least rapidly converging series’, Frege (1892c: 153) says, both have a definite sense, but it is doubtful whether the first has any reference, and it is demonstrable that the second lacks one. Many of the names found in fiction and myth fall here: ‘Medusa’, ‘Zeus’, ‘Santa Claus’. Thus Frege answers the vexing

question: How is it that a name that refers to nothing at all can still be a meaningful sign? Nondesignating singular terms have a sense but no reference: corresponding to such an expression is a criterion for recognizing whether a given object is the reference, and although there is no object satisfying the conditions laid down, the ‘way of recognizing’ gives the term the stability in discourse necessary for communication.

As we mentioned earlier, Frege’s distinction between sense and reference is intended to disambiguate the *Begriffsschrift* notion of content; and just as in *Begriffsschrift* Frege held both complex expressions as well as their parts have content, so in “On Sense and Reference” he extends the sense/reference distinction to complexes as well as to parts. As before, let  $\theta(\Omega)$  be a function-expression with one argument place marked by  $\Omega$ . Our two principles of the last chapter, the Compositionality Principle for Reference 2.3.1 and the Extensionality Principle for Reference 2.3.3, have their analogues for sense,<sup>11</sup> respectively

PRINCIPLE 3.6.3 (COMPOSITIONALITY FOR SENSE)  $s(\theta(\alpha)) = s(\theta)[s(\alpha)]$

and

PRINCIPLE 3.6.4 (EXTENSIONALITY FOR SENSE) *If*  $s(\alpha) = s(\beta)$ , *then*  $s(\theta(\alpha)) = s(\theta(\alpha/\beta))$ .

In *Begriffsschrift*, we may recall, Frege (1879) thought of a sentence as a complex name constituted of a function-expression and argument-expressions in such a manner that the contents of the parts of the sentence were part of the content of the whole sentence. The same idea is at work here, except that Frege (1892c) takes the reference of the sentence to be composed out of the reference of the parts of the sentence; and he takes the sense of the sentence to be composed out of the senses of the parts of the sentence. By distinguishing sense from reference, Frege now avoids the awkward feature of the *Begriffsschrift* theory that had caused him problems, namely, the fact that objects themselves were parts of contents. For, now, if a sentence contains a singular term, then the reference of that singular term, that is, the object for which it stands, is part of the reference of the sentence. It is not a part of the thought expressed by the sentence, because the thought, which Frege identifies with the sense of the sentence, contains the sense of the singular term as a part, not its reference.

We must distinguish between saying, on the one hand, that the reference (sense) of a complex name is a *function* of the reference (sense) of

the parts of the name, and saying, on the other hand, that the reference (sense) of a complex name *contains* the reference (sense) of the parts of the name. Frege noted that we must be careful about transferring our talk of parts and wholes with regard to linguistic expressions into the realm of sense and reference:

However, I have here used the word 'part' in a special sense. I have in fact transferred the relation between the parts and the whole of the sentence to its *Bedeutung*, by calling the *Bedeutung* of a word part of the *Bedeutung* of the sentence, if the word itself is a part of the sentence. This way of speaking can certainly be attacked, because the whole *Bedeutung* and one part of it do not suffice to determine the remainder, and because the word 'part' is already used of bodies in another sense. A special term would need to be invented. (Frege 1892c: 159)

At that time, Frege did not seem to be aware of the depth of the error. Eventually he had to abandon the claim that the reference of a part of a complex name is part of the reference of the complex name. For 'the Queen of England' contains the name 'England' as a proper part, but the Queen of England does not contain the country England as a proper part. On the other hand, Frege never dropped the part/whole metaphor for senses. We discuss the implications of this in Chapter 9.

The sense of a declarative sentence is the thought it expresses, and the reference of a declarative sentence is its truth value; and, as was the case with singular terms, a sentence might have a sense but lack a reference. A sentence is, for Frege, a complex name, and the relation between the reference of a sentence and the reference of the parts of the sentence is given by the Substitution Principle for Reference (2.5.1). But what about the relation between the sense of a sentence and the sense of the parts of the sentence? That we must recognize the sense of the sentence, the thought it expresses, as consisting of parts, is given in this very famous passage from Frege (1923: 537–8).

It is astonishing what language can do. With a few syllables it can express an incalculable number of thoughts, so that even a thought grasped by a human being for the very first time can be put into a form of words which will be understood by someone to whom the thought is entirely new. This would be impossible, were we not able to distinguish parts in the thought corresponding to the parts of a sentence, so that the structure of the sentence serves as an image of the structure of the thought. . . . If, then, we look upon thoughts as composed of simple parts, and take these, in turn, to correspond to the simple parts of sentences, we can understand how a few parts of sentences can go to make up a great multitude of sentences, to which, in turn, there correspond a great multitude of thoughts.

In addition, Frege (1892c: 157) notes:

If it were a question only of the sense of the sentence, the thought, it would be needless to bother with the *Bedeutung* of a part of the sentence; only the sense, not the *Bedeutung*, of the part is relevant to the sense of the whole sentence.

This suggests a substitution principle for sense analogous to that for reference. If the sense, and only the sense, of a given name contributes to the sense of a sentence containing that name, then were we to replace that name by any other having the very same sense, the resulting sentence ought to express the same thought as the original. Frege does not actually state such a principle in his published writings, but he often comes close, as, for example, in this passage from Frege (1893: 90):

The names, whether simple or themselves composite, of which the name of a truth-value consists, contribute to the expression of the thought, and this contribution of the individual [component] is its *sense*. If a name is part of the name of a truth-value, then the sense of the former name is part of the thought expressed by the latter name.

We thus have a substitution principle for sense comparable to the substitution principle for reference discussed earlier:

PRINCIPLE 3.6.5 (SUBSTITUTION FOR SENSE) *If  $S\alpha$  is about  $r(\alpha)$ , then if  $s(\alpha) = s(\beta)$ , then  $S\alpha$  and  $S\alpha/\beta$  have the same cognitive value.*

Finally, we should mention Frege's treatment of 'that' clauses – Quine has called these “opaque contexts” – contexts in which substitution of coreferential singular terms fails to preserve truth value. From

The Evening Star = the Morning Star (3.6)

and

John knows that the Evening Star is a body illuminated  
by the sun, (3.7)

we cannot validly conclude

John knows that the Morning Star is a body illuminated  
by the sun. (3.8)

Instead of taking it as a counterexample to the Substitution Principle for Reference 2.5.1, Frege regards the invalidity of this inference as evidence that (3.7) is not about the Evening Star, and so, 'the Evening Star' is not serving in (3.7) to stand for the Evening Star. Insofar as it is serving

to refer to something in (3.7), however, it will, as we mentioned earlier in discussing the notion of reference, stand for that which we intend to speak about in (3.7). And Frege says that what we are speaking about in this context is the sense of the expression. Thus Frege distinguishes between the *customary reference* of an expression and its *indirect reference*. In (3.7), 'the Evening Star' does not have its customary reference, the Evening Star, but its indirect reference; for we are using the term in that context to speak about the sense of the expression, that is the indirect reference. Similarly, Frege distinguishes between the *customary sense* of an expression and its *indirect sense*.

This machinery provides us with a test for determining whether two expressions have the same sense. If  $\alpha$  and  $\beta$  have the same sense, then a sentence involving a 'that' clause, like *John knows that  $S\alpha$*  should have the same truth value as *John knows that  $S\alpha/\beta$* . This test requires a considerable amount of fine-tuning. In particular, it faces a significant problem originally set by Mates (1952). Consider two expressions that have the same meaning, for example, 'fortnight' and 'two weeks'. It is quite possible that an individual might not know that the two expressions have the same meaning. So, that individual might know that John will be back in two weeks, and yet not know that John will be back in a fortnight. Since he or she knows the one proposition to be true but not the other, they must be different propositions, and so, by Frege's lights, this means that the sense attached to 'fortnight' is different from the sense attached to 'two weeks'. Reasoning in this way, it would turn out that no two distinct expressions can differ in sense, even if, on commonsense grounds, they have the same meaning. Church (1954) responded to the challenge by urging that any competent language user who knows the sense of each of these expressions must know that they have the same sense. But his argument relies on an appeal to the Church-Langford Translation Test, which we raise serious doubts about in Section 4.4. We discuss the issue in Chapter 9.

Frege's treatment of these cases is controversial. What is important and correct, however, is his recognition that they do not stand as counterexamples to the Substitution Principle for Reference 2.5.1 but rather show that the terms in those contexts do not serve purely to stand for what they ordinarily denote.<sup>12</sup>

## Frege's *Begriffsschrift* Theory of Identity

### 4.1 Introduction

Many commentators have been content to accept Frege's (1892*c*) account of his *Begriffsschrift* theory of identity, resulting in a somewhat distorted picture of the sense/reference theory. What little criticism there has been of the *Begriffsschrift* view can be grouped into the following three charges:

1. It has been alleged that the information contained in an identity statement, when interpreted in the manner of *Begriffsschrift*, can only be the trivial information that the linguistic community has adopted such-and-such conventions, not the substantial information embodied in a genuine discovery about the world. (This is derived from Frege's (1892*c*) own criticism of the *Begriffsschrift* theory.) See Linsky (1967), Kneale and Kneale (1962).
2. It has been alleged that the *Begriffsschrift* theory is circular or that it involves a vicious infinite regress. See Russell (1903*b*), Wiggins (1965), Kneale and Kneale (1962).
3. It has been alleged that the *Begriffsschrift* theory is flawed by use/mention confusion. See Church (1951), Furth's introduction to Frege (1893).

Not one of these adequately reflects the subtlety of Frege's *Begriffsschrift* view. We will examine and evaluate what he says in *Begriffsschrift* in Section 4.2, and then turn to these three criticisms: we treat the first in Sections 4.3 and 4.4, the second in Section 4.5, and the third in Section 4.6.

4.2 The *Begriffsschrift* Semantic Theory

In Chapter 8 of *Begriffsschrift*, Frege (1879: 64) defines “identity of content [*Inhaltsgleichheit*]” as follows: ‘ $\vdash A \equiv B$ ’ means that “the symbol  $A$  and the symbol  $B$  have the same conceptual content, so that  $A$  can always be replaced by  $B$  and conversely.” The symbol for identity of content ‘ $\equiv$ ’ was part of the object language, and since it, unlike the symbols for negation and material implication, represented a relation that holds between expressions instead of their contents, it required a special convention:

Whilst elsewhere symbols simply represent their contents, so that each combination into which they enter merely expresses a relation between their contents, they at once stand for themselves as soon as they are combined by the symbol for identity of content; for this signifies [*bezeichnet*] the circumstance that two names have the same content. (Frege 1879: 64)

Hence names in *Begriffsschrift* were systematically ambiguous. They stood for the objects they customarily denoted everywhere save when they occurred at either end of the symbol for identity of content, where they stood for themselves.

Although Frege notes that identity of content alone among the logical constants relates expressions, he does not reveal why this is so. Rather, his primary concern is to justify including such a relation in his *Begriffsschrift*. Our working hypothesis, however, is that Frege’s (1892*c*) reconstruction was accurate and that he chose to take identity of content as a relation between expressions in order to deal with the Paradox of Identity. So he must have held a substitution principle like Principle 2.3.1 in *Begriffsschrift*. Our first task is to confirm this. Our second is to examine how, within such a framework, Frege hoped to account for the cognitive content of identities.

There are two axioms governing identity of content in *Begriffsschrift*. These are Proposition 52,

$$x \equiv y \supset (f(x) \supset f(y)), \quad (4.1)$$

and Proposition 54,

$$x \equiv x. \quad (4.2)$$

These two propositions are sufficient to characterize identity in a first-order theory. Syntactically ‘ $\equiv$ ’ is indistinguishable from ‘ $=$ ’. Informally, too, identity of content appears to be no different from identity in that Frege will use ‘ $x \equiv y$ ’ to express, say, that  $x$  is the same number as  $y$ . There is, however, a peculiarity about ‘ $\equiv$ ’ that Frege does not mention in



Chapter 8. Sentences as well as singular terms are said to have conceptual content, and Frege uses ‘ $\equiv$ ’ also to express the circumstance that two sentences have the same conceptual content. He allows the constants in (4.1) and (4.2) to be replaced either by singular terms or by sentences, so that

$$x \equiv y \supset (f(x) \equiv f(y)) \quad (4.3)$$

is easily derived in his system. Now (4.3) says that if  $\alpha$  and  $\beta$  have the same conceptual content,  $S\alpha$  and  $S\alpha/\beta$  have the same conceptual content. (This might well have been what Frege had in mind when, in his informal definition of identity of content quoted above, he said that ‘ $\vdash A \equiv B$ ’ means that “A can always be replaced by B and conversely,” that is, the replacement preserves conceptual content. As such, the replacement would certainly preserve truth value, justifying (4.1) above.) When  $\alpha$  and  $\beta$  are both singular terms,  $\alpha \equiv \beta$  appears to have the same truth conditions as  $\alpha = \beta$ , that is, it will be true when, and only when,  $\alpha$  and  $\beta$  stand for the same object. But we have yet to clarify what sameness of conceptual content amounts to when  $\alpha$  and  $\beta$  are both sentences.

We turn, then, to the semantic theory Frege held in *Begriffsschrift*. On that theory, a sentence stands for its content [*Inhalt*]. This might be understood to be a thought, but only if we are careful not to assimilate this use to Frege’s more familiar technical notion of thought [*Gedanke*] from the sense/reference theory. In *Begriffsschrift*, it seems to be more like a state of affairs or a circumstance, that is, something that could obtain. The sentence stands for its content – the sentence is, in Frege’s words, a *Vertreter* (proxy or substitute) for its content – and the parts of the sentence in turn stand for corresponding parts of the content of the sentence. In accordance with the (primitive) function/argument analysis Frege had introduced in *Begriffsschrift*, were we to replace any part of the sentence by another having the same content, the resulting sentence should have the same content as the original.

Only that portion of the content of a sentence that counted for inference was of any interest to Frege, and this he called the “conceptual content [*begrifflichen Inhalt*]” of the sentence:

[N]ote that the contents of two judgments can differ in two ways: either the conclusions that can be drawn from one when combined with certain others also always follow from the second when combined with the same judgments, or else this is not the case. The two propositions ‘At Plataea the Greeks defeated the Persians’ and ‘At Plataea the Persians were defeated by the Greeks’ differ in the first way. Even if a slight difference in sense can be discerned, the agreement

predominates. Now I call that part of the content that is the *same* in both the *conceptual content*. Since *only this* has significance for the *Begriffsschrift*, no distinction is needed between propositions that have the same conceptual content. (Frege 1879: 53)

Since each sentence is inferable from itself, with or without any additional premises, Frege's condition for sameness of conceptual content comes to this: two sentences have the same conceptual content if, and only if, they are mutually inferable.

The quoted condition for sameness of conceptual content, however, applies only to a "possible content of judgment [*beurtheilbar Inhalt*]," that is, to the content of a declarative sentence, and nowhere in *Begriffsschrift* had Frege explicitly stipulated the conditions under which two singular terms were to have the same conceptual content. Presumably there would be no need to spell this out if identity of content was functioning just like identity. Frege's practice, as we have mentioned, was to regard two singular terms as having the same conceptual content if, and only if, they designated the same object. So he apparently identified the conceptual content of a singular term with the object it denoted. But, in keeping with the overall function/argument structure of *Begriffsschrift*, the indicated course would be to take two singular terms as having the same conceptual content if, and only if, replacing one by the other in a given sentence results in another sentence having the same conceptual content as the original. Putting these together yields the following substitution principle:

PRINCIPLE 4.2.1 (*BEGRIFFSSCHRIFT* SUBSTITUTION) *If  $S\alpha$  is about  $r(\alpha)$ , then if  $r(\alpha) = r(\beta)$ , then  $S\alpha$  has the same conceptual content as  $S\alpha/\beta$ .*

Principle 4.2.1 is our *Begriffsschrift* analogue to Principle 3.3.1. The only difference is that in Principle 4.2.1 Frege speaks of "conceptual content" where in Principle 3.3.1 he uses "cognitive value," and this, we now see, is virtually a terminology difference. Frege had supposed that the conceptual content of a complex was a function of the conceptual content(s) of the part(s). Further, this function/argument structure was construed ontologically as a part/whole relation, so that the conceptual content of the whole contained the conceptual content of the part(s). For singular terms, the conceptual content was the object designated, but for sentences, the conceptual content, which was intended to be something like a *Gedanke*, was more like a Russellian singular proposition. It was the author of *Begriffsschrift* who held the view Frege later scorned, that

is, of having objects themselves as parts of thoughts. Frege had apparently grafted this fairly common and intuitive notion of the content of a sentence onto a function/argument structure that was bound to reject it. For, as Frege (1892c) showed, substitution of coreferential singular terms preserves nothing more than truth value – not *Sinn* (an expression's sense), not even *beurtheilbar Inhalt*, Frege's *Begriffsschrift* hybrid of *Sinn* and *Bedeutung*. As such, the difficulty with Frege's *Begriffsschrift* semantics was bound to surface at some point when substituting coreferential singular terms resulted in a sentence that did not have the same conceptual content as the original.

That point came with identities. Frege discovered that while the logically true  $\alpha = \alpha$  and the contingently true  $\alpha = \beta$  appeared to have the same conceptual content, because both say of the same object(s) that they stand in the same relation, they could not have the same conceptual content. They could not have the same conceptual content because, since one is logically true and the other contingently true, they could not be mutually inferable. He did not, however, abandon Principle 4.2.1 at this point. Blinded no doubt by the use/mention sloppiness that infected his *Begriffsschrift* thinking, Frege saw this problem as purely local to identity, one that could be taken care of by reinterpreting identity as a relation between expressions. So, since  $r(\alpha) = r(\beta)$ , the difference in conceptual content between the identities led, via Principle 4.2.1, to the conclusion that  $\alpha = \beta$  could not express a relation holding between the objects  $r(\alpha)$  and  $r(\beta)$  themselves. Taking identity to relate the expressions, on the other hand, ensured that the content of the logically true identity differed from that of the contingently true identity: where the content of  $\alpha = \beta$  would have  $\beta$  as a constituent, the content of  $\alpha = \alpha$  would have  $\alpha$  as a constituent. With this device, the semantic framework was patched up.

But this solution to his problem required some explanation. Immediately after introducing '≡' and explaining the unusual convention that the expressions flanking '≡' were to stand for themselves, Frege (1879: 64) noted: "This makes it appear at first as if it were here a matter of what pertains to the *expression* alone, *not to the thought*, and as if there were no need at all for different symbols for the same content and hence for a symbol for identity of content either." Frege was anxious lest the reader believe that identity of content was only incidentally, if at all, connected with thought; for, were this so, there would be little need for a *symbol* for identity of content in what was supposed to be, as the subtitle announced, a *formula language of pure thought*.

Frege had good reason to be apprehensive on this score, for a strong case can be developed against including a symbol for identity of content in his *Begriffsschrift*. Consider: if  $\alpha$  and  $\beta$  have the same conceptual content, then  $S\alpha$  and  $S\alpha/\beta$  also have the same conceptual content, excepting, of course, when  $S\alpha$  is a sentence expressing identity of content. It seems, then, to be of no logical consequence whether we use the one term or the other in a given sentence, excepting, again, a sentence expressing identity of content. But  $\alpha = \alpha$  and  $\alpha = \beta$  differ in conceptual content (if  $\alpha$  and  $\beta$  are distinct signs) only because of what now appears to be the purely ad hoc device of having  $\alpha$  and  $\beta$  stand for themselves. The information thus obtained is of very limited applicability – limited, that is, to sentences expressing identity of content. Were we to excise ‘ $\equiv$ ’ from the notation, the remainder of the logical symbolism would be indifferent to  $\alpha$  and coreferential  $\beta$ . A comparison with Frege’s treatment of the active/passive distinction in *Begriffsschrift* is quite telling. Since each one of an active/passive pair of sentences has the same conceptual content as the other, Frege ignored this grammatical distinction in his *Begriffsschrift* and symbolized each in the same way. Why did he not adhere to this practice for sentences that differ only in that where the one contains  $\alpha$  the other contains coreferential  $\beta$ ? The difference in singular terms fails to reflect a difference in conceptual content, so the indicated course would be to ignore the trivial difference in formulation, symbolize each in the same way, and thus eliminate the need for the symbol for identity of content altogether.

The assumption apparently underlying this objection is that names are meaningless marks, arbitrarily chosen labels or tags that simply stand for objects but otherwise carry no meaning; and it is just this assumption that Frege rejected in *Begriffsschrift*. Instead, Frege (1879: 65) urged that “different names for the same content are not always just a trivial matter of formulation, but touch the very heart of the matter if they are connected with different modes of determination.” Frege illustrated this as follows. Fix a point, A, lying on the circumference of a given circle and pass a straight line through A, extending the line so that it intersects with the circle. This point of intersection, which we will call ‘B’, obviously depends upon the position of the straight line, so that as the line is rotated about A, B varies accordingly:

It can now be asked: what point is yielded when the line is perpendicular to the diameter? The answer will be: The point A. The name B has therefore in this case

the same content as the name A; and yet just one name could not have been used from the beginning, since the justification for doing so is only provided by the answer. The same point is determined in two ways:

1. immediately through intuition [*Anschauung*];
2. as the point B when the line is perpendicular to the diameter. (Frege 1879: 64–5)

Unlike in the active/passive case where sameness of conceptual content is immediate, then, it is not always obvious whether two singular terms happen to have the same conceptual content. This nonobviousness was not, however, to be attributed to the unpredictable creativity of a language-using community that can generate names for a given object as whim dictates, these names differing only in formulation. This case would be essentially indistinguishable from the active/passive case. Rather, Frege's point is that a singular term can carry something with it, a mode of determination [*Bestimmungsweise*], that enables us to figure out whether, given a particular object, the term stands for it, and so to figure out whether the term stands for the same object as does some other singular term that determines it in a different manner. The knowledge gained as a result would reflect substantial information about the object. This, then, provides the link between names and thought: some identities express synthetic truths yielding solid information about their subject matter. In the example, the same point is determined in two different ways, and the nontrivial fact that the same point is determined in each of the two different ways is expressed by 'A  $\equiv$  B' but not by 'A  $\equiv$  A'.

Finally, Frege (1879: 65) sums up the connection between names, *Bestimmungsweisen*, and identity of content:

The need for a symbol for identity of content thus rests on the following: the same content can be fully determined in different ways; but *that*, in a particular case, *the same content* is actually given by *two modes of determination* is the content of a *judgment*. Before this judgment can be made, two different names corresponding to the two modes of determination must be provided for that that is thereby determined. But the judgment requires for its expression a symbol for identity of content to combine the two names.

(Another reason Frege offered for including ' $\equiv$ ' in his *Begriffsschrift* was that it would be needed to introduce definitions, but this was only a "superficial reason.")

The reader will immediately recognize the resemblance between the *Begriffsschrift* notion of *Bestimmungsweise* and the sense/reference notion

of an *Art des Gegebenseins* or *Darstellungsweise*, that is, the way in which an object is given or presented, this latter being the active ingredient of an expression's sense [*Sinn*]. Indeed, Frege's example in "On Sense and Reference" of an object's being presented in different ways could easily have served in *Begriffsschrift* as an example of an object's being determined in different ways:

Let  $a$ ,  $b$ ,  $c$  be the lines connecting the vertices of a triangle with the midpoints of the opposite sides. The point of intersection of  $a$  and  $b$  is then the same as the point of intersection of  $b$  and  $c$ . So we have different designations for the same point, and these names ('point of intersection of  $a$  and  $b$ ', 'point of intersection of  $b$  and  $c$ ') likewise indicate the mode of presentation; and hence the statement contains actual knowledge. (Frege 1892c: 152)

This similarity with the sense/reference theory can be pushed still further. Thus, where Frege (1892c) distinguished between that which a term stands for, its reference [*Bedeutung*], and that which the term expresses [*Ausdrucken*], its sense, so Frege (1879) distinguished between that which a term stands for, its content, and that which the term goes along with [*Entsprechenden*], a *Bestimmungsweise*. Again, where Frege (1892c) distinguished between that which identity relates and that wherein the information conveyed by an identity resides, so Frege (1879) held that although identity is to relate the terms flanking the identity sign, the information is to be that the same content is given by the two ways of determining it.

But although Frege (1879) had in effect drawn a sense/reference distinction for singular terms, he had only *beurtheilbar Inhalt* for sentences. So, after all has been said, he was still committed to Principle 4.2.1: substitution of coreferential singular terms – whether they stand for their ordinary contents or, in the context of the symbol for identity of content, themselves – preserved conceptual content. As a result, Frege had nowhere to trace the *Bestimmungsweise* associated with a singular term save in the conceptual content of a sentence. This placed him on the horns of a dilemma. Where  $r(\alpha) = r(\beta)$ , but the *Bestimmungsweise* associated with  $\alpha$  is different from the *Bestimmungsweise* associated with  $\beta$ , either  $S\alpha$  and  $S\alpha/\beta$  have the same conceptual content or they do not. If  $S\alpha$  and  $S\alpha/\beta$  have the same conceptual content, then, so far as inference is concerned – and this is, recall, the sole interest of *Begriffsschrift* – it makes no difference whether the *Bestimmungsweisen*  $\alpha$  and  $\beta$  correspond to are the same or not. On the other hand, if  $S\alpha$  and  $S\alpha/\beta$  differ in conceptual content, then the Substitution Principle 4.2.1 would be violated. Anxious to avoid the first alternative, which would have rendered '≡' an idle excrement

of the logical apparatus, Frege did not realize he was being impaled by the second. For, by taking identity to relate expressions instead of their contents, he was committed to there being a logically significant difference in the symbols that reflected no logically significant difference in the contents. This, of course, was in direct contravention to what he had said when he explained the notion of conceptual content. Formally, this meant that he was holding

$$\alpha \equiv \beta \tag{4.4}$$

to be true and

$$(\alpha \equiv \alpha) \equiv (\alpha \equiv \beta) \tag{4.5}$$

to be false, contradicting (4.3).

### 4.3 Criticism: The Received View

It is commonly believed that Frege (1879) had attributed the informativeness of an identity in *Begriffsschrift* to the difference in the terms flanking the identity sign; that he became disenchanted with the *Begriffsschrift* account because on close analysis it turned out that the information thus conveyed could be only of the arbitrary conventions of the language-using community; that he subsequently discovered, some years after *Begriffsschrift*, that a term has, besides its reference, a sense whose connection with the reference is not a matter of convention; and that this discovery allowed Frege (1892c) to locate the “actual” or “proper” knowledge conveyed by an identity in the difference in the senses of the terms flanking the identity sign. Obviously this story cannot be correct. For, as we have seen, Frege already had a notion in *Begriffsschrift*, that of a *Bestimmungsweise*, corresponding roughly to the sense/reference notion of sense, and he had attributed the informativeness of an identity in the difference in the *Bestimmungsweisen* associated with the expressions flanking the identity sign, not merely to the difference in the expressions themselves; indeed, he had introduced *Bestimmungsweisen* precisely to counter the charge that he had trivialized identities by construing identity as identity of content. What he did lack in *Begriffsschrift* was the sense/reference distinction for sentences,<sup>1</sup> and this attended his clarification of the function/argument analysis of language.

This criticism of Frege’s *Begriffsschrift* account of identity is derived from Frege’s (1892c) comments in the opening paragraphs, which have been interpreted by a number of commentators as a precursor of a

more contemporary dispute between those who would require full-blooded propositions in semantics and those who would attempt to reduce semantic facts to facts about words. Under this recreation of Frege's intellectual history, the move from the *Begriffsschrift* account to the sense/reference account is viewed as a rejection of the reductionist attitude and an adoption of the propositionalist attitude.

Typical examples of this received view are (from Linsky and the Kneales):

Nor is Frege able to accept the other of the two alternatives, that identity is a relation between names or signs of objects. Then ' $a = b$ ' would just say that the name ' $a$ ' and the name ' $b$ ' are names for the same thing. This analysis cannot be correct, Frege argues, because the fact that ' $a$ ' is a name for  $a$  and that ' $b$ ' is also a name for  $a$  results from a purely arbitrary agreement governing the use of these marks (or sounds). Furthermore, when I say that Venus is the morning star I am conveying information about the heavens, not about our arbitrary use of signs. (Linsky 1967: 22)

As we have seen, Frege suggested that a statement of identity must really be about the expressions appearing on the two sides of the identity sign, and he tried to make this clear by saying that ' $\equiv$ ' was to be understood as a symbol for identity of content between expressions. But he came to see later that this was not a satisfactory solution of the puzzle. For he realized that if the original statement [that the morning star is identical with the evening star] was not really about the planet Venus, but about the contents of certain phrases, it would belong to philology rather than to astronomy, which is obviously not the case, since the discovery of the identity of the morning star and the evening star was made by observation and calculation, not by reflection on the use of words. (Kneale and Kneale 1962: 494)

This criticism of the *Begriffsschrift* theory goes, roughly: if a sentence expresses a relation holding between expressions rather than between what those expressions stand for, then the sentence conveys information, not about what those expressions stand for, but only about the expressions themselves; and since 'Venus = the Morning Star' conveys information about Venus and the Morning Star, '=' cannot relate the expressions 'Venus' and 'the Morning Star'. Yet surely we are sufficiently sophisticated in the techniques of modern philosophy – for example, shifting between material and formal mode – to doubt the premise of the criticism. The relation *denotes the same thing as* clearly relates expressions, not what expressions stand for, and yet the sentence

'Venus' denotes the same thing as 'the Morning Star' (4.6)



serves in certain circumstances to convey information about the heavens, and in other circumstances to convey information about words. Again, Linsky's claim that "the fact that 'a' is a name for *a* and that 'b' is also a name for *a* results from a purely arbitrary agreement concerning the use of these marks" is false. It might be arbitrarily agreed that 'a' names *a* and arbitrarily agreed that 'b' names *b*, yet although 'a' and 'b' name the same thing, it would be a flagrant intensional fallacy to conclude that it was arbitrarily agreed that 'a' and 'b' name the same thing. Linsky saddles Frege with this argument, but it is hardly credible that Frege would commit just this kind of error in, of all places, "On Sense and Reference."

The Linsky-Kneale criticism of Frege's *Begriffsschrift* theory of identity is, as we have mentioned, derived from Frege's (1892*c*) opening discussion of the matter. What could Frege have had in mind by bringing up the fact that words can be arbitrarily assigned to things, that we would not be able to convey proper knowledge if we regarded signs only as objects, not as symbols? These all point to the received interpretation, which is that senses and propositions are needed for semantics. But we cannot accept this interpretation. There is always the possibility that Frege was genuinely confused about the issue and his reasoning was just bad here. But we prefer a more positive story, if available, and we offer the following for consideration.

It is fairly clear, from what we have been arguing, that Frege's major criticism of the *Begriffsschrift* theory was of the underlying semantic theory that yielded Principle 4.2.1. Correcting his *Begriffsschrift* errors involved explaining the sense/reference distinction, the function/argument distinction, and the concept/object distinction, a lengthy and complicated process that occupied the whole of "On Sense and Reference" and spilled over into its companion essays. In the quoted passage, however, Frege chose to criticize his *Begriffsschrift* theory of identity on the very narrow grounds that it failed in its stated objective. That is, because he had taken identity to relate expressions, there was no way he could distinguish those identities where the names differed trivially in formulation (which would be essentially like the active/passive difference) and those cases where the names differed in *Bestimmungsweisen*. Both turn out to be given equal weight, so that he had not, by this device, succeeded in capturing the informative character of identities. Linsky and the Kneales miss this point: unaware that Frege had distinguished *Bestimmungsweisen* in *Begriffsschrift*, they took Frege to be urging that it was the lack of such a distinction that caused the theory's downfall.

But we still need to return to these criticisms. For the particular line of argument put forward by Linsky and the Kneales is intimately connected with a powerful argument we associate with Church (1950) and Langford, and which has become something of a staple in the arsenal of modern analytic philosophy. In rejecting the Linsky-Kneale arguments, we are rejecting, in part, the efficacy of the Church-Langford Translation Test. We turn to this now.

#### 4.4 Criticism: Church-Langford Considerations

It is characteristic of analytic philosophy to recast problems about objects in terms of problems about words. This is most evident in Carnap's (1937) material mode/formal mode distinction and Quine's (1960) technique of semantic ascent. There are, no doubt, differences between the Object Language and the Word Language, or else philosophers would not have preferred one to the other. But these differences do not adversely distort the problem analyzed, or so it is claimed. This last is the moot point, the target of the Church-Langford Translation Test (hereafter the Test).

Here is an example of how the Test is used. One analysis of *believes that* takes it to be a relation between an individual and a sentence. Since this is a philosophical analysis, these two English sentences should say the same thing:

John believes that the moon is round. (4.7)

John believes 'the moon is round'. (4.8)

Do they say the same thing? They appear to, but, to be sure, we apply the Test. Translating into German, we get, respectively,

John glaubt, dass der Mond rund ist. (4.9)

John glaubt 'the moon is round'. (4.10)

These two German sentences clearly express different propositions, so the two English sentences must, despite appearances, also express different propositions. The Test shows the analysis to be wrong.

Let  $\omega$  and  $\omega$  abbreviate the two sentences in the home language; let  $T(\eta)$  be the translation of  $\eta$  into a particular foreign language; and let  $R$  represent a semantic relation between two sentences (for example, synonymy, implication). The Church-Langford Translation Test works as follows. Suppose that

$$\omega R \omega. \quad (4.11)$$

According to what is claimed to be a generally accepted translation principle,

PRINCIPLE 4.4.1 (CHURCH-LANGFORD TRANSLATION) *If  $oR\omega$ , then  $T(o) R T(\omega)$ ,*

for a sentence and its translation express the same proposition. But, applying

PRINCIPLE 4.4.2 (SINGLE-QUOTE TRANSLATION) *Expressions inside single quotes are not to be translated,*

the translation exaggerates a difference between  $o$  and  $\omega$ . The home speaker, upon inspecting  $\omega$ , reads inside the single quotes and understands the words mentioned. The foreign-language speaker, on the other hand, upon inspecting  $T(\omega)$ , does not understand the words inside the single quotes since they are in an unknown (to him or her) language. It is obvious that

$$T(o) R T(\omega) \tag{4.12}$$

is false. So, by Modus Tollens, (4.11) must be false.

Whether an  $o$  sentence can express the same proposition as its  $\omega$  counterpart is not our concern here. What we are concerned with is the widely believed claim that Frege's rejection of the *Begriffsschrift* account of identity is prompted by the same sort of concerns that prompt the Church-Langford Translation Test. Linsky (1983: 7) frames the issue:

In "On Sense and Reference" Frege tells why he abandoned his early view: "Nobody can be forbidden to use any arbitrarily producible event or object as a sign for something. 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." I think that Frege here is alluding to the same considerations as those involved in the so-called "Church-Langford translation test."

Linsky claims that Frege's *Begriffsschrift* theory of identity fails because it violates the Church-Langford Translation Test, and that Frege abandoned it for essentially similar reasons. Now, although the *Begriffsschrift* theory does fail, it does not fail for anything like the reasons Linsky puts forward. Our argument here will be in two parts. First, we will argue that the Test itself is questionable because Principle 4.4.2 is questionable. Second, whatever the outcome of the Test, we will argue that it cannot yield the conclusion Linsky seeks to draw from it.

Here is Linsky's (1983: 7–8) argument:<sup>2</sup>

Consider the statement

Venus is the Morning Star (4.13)

According to the *Begriffsschrift* theory, which Frege now rejects, (4.13) says the same thing as

'Venus' and 'the Morning Star' denote the same thing (in English). (4.14)

If this supposition is correct, the translations of these two sentences into Italian must also convey the same information; that is

*Venere è la Stella Mattutina* (4.15)

must convey the same information as

'Venus' e 'the Morning Star' significano la stessa cosa (in Inglese). (4.16)

In translating (4.14) into (4.16), it is necessary to leave the quoted material intact, for it is the names 'Venus' and 'the Morning Star' that are said to denote the same thing (in English). Any translation of (4.14) must be about these very names and not about their Italian translations, 'Venere' and 'la Stella Mattutina'. Suppose now that you are an Italian who does not understand any English at all. Then it is clear that (4.15) and (4.16) convey different information to you. (4.15) tells you an important astronomical fact, but (4.16) tells you nothing about the heavens at all. It tells you, rather, a fact about the English language. It is clear that (4.15) and (4.16) have different cognitive content. Then (4.13) and (4.14) must also convey entirely different information, and for the same reason. (4.13) is about the heavens and (4.14) is about English words. Thus Frege is forced to abandon his earlier view.

Linsky concludes: "(4.13) is about the heavens and (4.14) is about English words." This claim has no bite unless it means "(4.13) is about the heavens *and not about English words*; (4.14) is about English words *and not about the heavens*." This is far from obvious. The second half of the claim, for example, is simply false. (4.14) is as much about the heavens as it is about words. A philosophical naif would have no qualms expressing what he seeks to express using either (4.13) or (4.14). We do find single quotes around 'Venus', and again around 'the Morning Star', but even the philosophical sophisticate would not immediately infer that we are not thereby talking about the heavens. For we find 'denotes', a *disquotation operator*, right around the corner.

Linsky's argument turns on Principle 4.4.2. One cannot – must not – translate inside the single quotes; the words there are mentioned, not used. But should we accept Principle 4.4.2? It is worth noting that translating inside quotation marks – ordinary quotation marks, that is – is a

perfectly reasonable procedure. In a murder mystery, for example, one character talks about the suspect's exact words: the Italian translation has the character quoting the exact Italian words used to translate the suspect's remarks. Single quotation marks are a philosopher's invention, however, and normal practice does not automatically apply to them. But does philosophical practice demand Principle 4.4.2? Indenting is a variant of quoting. A sentence that is numbered and indented, as in a logic book, is mentioned. Principle 4.4.2 bars us from translating indented sentences; it therefore effectively bars us from translating a logic text into a foreign language. Philosophical practice, however, admits of such translations. As a matter of fact, Black and Geach's (1952) translation of "Über Sinn und Bedeutung" standardly translates the words inside single quotes.<sup>3</sup> Linsky is quite comfortable with this translation. He relies on it throughout his commentary, and never remarks upon this blatant authoritative disregard for his translation principle. The mere fact that one is mentioning an expression, even enclosing it within single quotes, is insufficient to justify not translating it.

Of course, this is not to say that mentioned expressions should always be translated. In a book about English grammar, where the examples of English usage are crucial, the mentioned examples should not be translated. This seems to be the situation Linsky sees here. His view is that (4.14) speaks about the words used in (4.13), those very words; one would do terrible violence to the thought if one used the Italian equivalents. That is why he will not translate inside the quotation marks.

But this is not right, for the Italian translation of Linsky's book would render (4.14) as

‘Venere’ e ‘la Stella Mattutina’ significano la stessa  
cosa (in Italiano), (4.17)

not (4.16), as Linsky claims above. Not only would one translate inside the quotes, but one would even change "in English" into "in Italiano." The effect of invoking Principle (4.4.2) is to keep reference constant in translation. This poses something of a problem when we have a case of self-reference: for, in order to preserve self-reference, reference has to be shifted. An article on the Liar Paradox contains the example

This sentence is false. (4.18)

A French translation of the article will surely render (4.18) as

Cette phrase est fausse. (4.19)

Reference has been changed so that the self-reference necessary to generate the paradox is preserved. Even though it is those very English words that are being referred to in (4.18) (as with the quoted names), we have shifted the reference in (4.19) in order to provide the obviously correct translation. Indeed, were we unable to shift reference, the English sentence (4.18) would be untranslatable into another language. In any event, Principle (4.4.2) appears to be far from a universal truth. And it is even a possibility that (4.17) should be the translation of (4.14), not (4.16).

Do (4.18) and (4.19) express the same proposition? That is a hard question. Thank goodness, we do not have to answer it. For that is something Church must argue for; we carry no such theoretical baggage.

Consider the following argument against Linsky:

A competent English speaker who hears (4.14) knows, because of the disquotation operator, that (4.13) follows from (4.14). The competent Italian speaker who hears (4.16) cannot make the analogous inference to (4.15). Therefore, (4.15) and (4.16) cannot be adequate translations of (4.13) and (4.14).

This shows that Linsky's proposed Italian translations of (4.13) and (4.14) are incorrect. The argument is not that (4.15) and (4.16) are incorrect translations of (4.13) and (4.14), respectively, *because* (4.15) and (4.16) convey different information. That would beg the question. Granted that (4.15) and (4.16) convey different information. Nonetheless, there are other relations that should be maintained, and they are not. The same inferential relations holding between (4.13) and (4.14) should hold between (4.15) and (4.16), and they do not. From (4.14), along with the two additional premises,

‘Venus’ denotes Venus, (4.20)

and

‘the Morning Star’ denotes the Morning Star, (4.21)

we can derive (4.13). But from (4.16), along with (4.20) and (4.21), we cannot derive (4.15), so these are not correct translations of (4.13) and (4.14).

Linsky does have a response. To be sure, (4.15) does not follow from (4.16), along with (4.20) and (4.21). One must translate (4.20) and (4.21) into Italian to validly derive (4.15). From (4.16), along with

‘Venus’ significa Venere, (4.22)

and

‘the Morning Star’ significa la Stella Mattutina, (4.23)

(4.15) follows.

This response makes clear one very common understanding of why the Church-Langford Translation Test works. The home speaker mistakenly supposes (4.13) and (4.14) to express the same proposition because he fails to recognize the empirical claims about the meanings of English words, (4.20) and (4.21), that underpin the identification. But this response raises perhaps more problems than it solves. For the claim that (4.22) translates (4.20) must entail that they express the same proposition. Do they? If (4.20) and (4.22) express the same proposition, then they should either both be trivial or both be substantial, and (4.22) seems to give much more substantial information than (4.20). So Linsky, in defending his translation (4.16), appears forced into the same position as one who maintains that  $\alpha = \alpha$  and true  $\alpha = \beta$  express the same proposition. On the one hand, he takes ‘Venus’ and ‘Venere’ to be trivial notational variants of the same word that can be replaced one for the other in any sentence preserving the proposition expressed. On the other hand, he takes “Venus” and “Venere” to be significantly different words that differ in both sense and reference. But what we have just seen is that, because of the disquotation principle, whatever attitude he adopts for ‘Venus’ and ‘Venere’ must also be adopted for “Venus” and “Venere”.

What to say about (4.20). and (4.21) is genuinely puzzling. Kaplan (1969) claims they are analytic. But the Church-Langford Test would seem to suggest otherwise; moreover, it is not clear how it can be analytic that this word refers to that object. But they do appear to be trivial; the knowledge that ‘Venus’ denotes Venus, for example, results from our understanding of the disquotation operator *denotes*. Other things being equal, if we fill the blank spaces in

‘\_\_\_\_\_’ denotes \_\_\_\_\_ (4.24)

by the same noun phrase, the result expresses a truth. We will return to this in Chapter 10.

Whatever our view about this issue, the matter of whether we translate inside the single quotes is entirely irrelevant to the coherence of the *Begriffsschrift* theory of identity. Here is an argument to show that (4.14) is as much about the heavens as it is about words. Our reasoning follows

the analysis to be given in the next section. Let us first restate (4.14) as

The denotation of 'Venus' is the same as the denotation  
of 'the Morning Star'. (4.25)

Now, as we will see shortly, there are two ways of parsing (4.25) so that it has the form  $R(\alpha, \beta)$ , namely, in the manner of (4.28) and in the manner of (4.29). Taken the first way, it expresses that the object Venus and the object the Morning Star are identical; taken the second way, it expresses that the equivalence relation *denoting the same thing as* holds between the two expressions, 'Venus' and 'the Morning Star'. Moreover, the Italian translation of (4.25) will preserve the two logical analyses, and this will be so whether or not we translate inside the quotation marks. The claim that one is about words and the other about the world is unsupportable, even if we accept the result of the Church-Langford Translation Test.

#### 4.5 Criticism: The Alleged Regress

Let us recall the first clause of Frege's definition of ' $\vdash A \equiv B$ ', which was, with single quotes inserted appropriately:

'A' has the same conceptual content as the symbol 'B'. (4.26)

Russell (1903*b*) remarked that, if taken as a definition of identity, (4.26), "verbally at least, suffers from circularity." Although Russell did not enlarge upon this observation, his brief, and slightly hedged, criticism has been fleshed out by David Wiggins, who argues that (4.26) generates a vicious infinite regress:

Asking for the sense of ' $a = b$ ' I am told ' $a$ ' and ' $b$ ' have the same content, or designate only one thing. Unless something is said to justify calling a halt here, the explanation generates a new statement of the same form as the original explicandum – 'The content or designatum of " $a$ " = the content or designatum of " $b$ ".' Applying the same explanation to this we get 'The content or designatum of "the content or designatum of " $a$ " = the content or designatum of "the content or designatum of " $b$ ".' But evidently we can never reach in this way what seems to be needed to carry the explanation through, a statement only about signs. (Wiggins 1965: 51)

Wiggins apparently believes that the *Begriffsschrift* analysis of an identity never actually yields a sentence about signs; rather, each reinterpretation in accordance with the *Begriffsschrift* formula only succeeds in yielding another sentence about objects.



It is our view, however, that Wiggins has simply misinterpreted the *Begriffsschrift* formula. Let us rephrase (4.26) to read

The conceptual content of the symbol 'A' is the same as the  
conceptual content of the symbol 'B'. (4.27)

Now, there are two logical analyses of (4.27) worth considering, each assigning it the form  $R(\alpha, \beta)$ :

$R$ : \_\_\_\_ is the same as \_\_\_\_  
 $\alpha$  : the conceptual content of the symbol 'A'  
 $\beta$  : the conceptual content of the symbol 'B' (4.28)

and

$R$ : The conceptual content of \_\_\_\_ is the same as the  
conceptual content of \_\_\_\_ .  
 $\alpha$  : the symbol 'A'  
 $\beta$  : the symbol 'B' (4.29)

If we analyze (4.27) in the manner of (4.28), then (4.27) is understood to express a relation that holds between the conceptual content of the symbol 'A' and the conceptual content of the symbol 'B', that is, between A and B, namely, identity. If, on the other hand, we analyze (4.27) in the manner of (4.29), then (4.27) is understood to express a relation that holds between the symbol 'A' and the symbol 'B', namely, the equivalence relation *having the same conceptual content*. It is clear that (4.29) is the analysis Frege wanted. Wiggins, however, requires (4.28) in order to get his regress off the ground, and the regress is only a frill since, on his interpretation, Frege's would simply be a fancy way of expressing old-fashioned identity.

Although we are sympathetic to the sort of criticism Wiggins (like Russell) wishes to make of the *Begriffsschrift* theory, we shall not pursue their particular avenue of attack. For the problem with the *Begriffsschrift* view is evident as soon as we compare the two analyses of (4.27). It is well known that an equivalence relation can be restated in terms of the identity relation. As a matter of fact, Frege's (1884b: Section 64) discussion of the matter is the locus classicus for such a procedure. But the heart of the *Begriffsschrift* view is that the equivalence relation *having the same conceptual content* cannot be eliminated in favor of the identity of conceptual contents. This is the exact point on which the *Begriffsschrift* view is vulnerable, as we will show in the next section.

## 4.6 Criticism: Use/Mention Confusion

Frege had not, in *Begriffsschrift*, yet come to appreciate the need for careful observation of the distinction between sign and thing signified: a cursory reading reveals, for example, that he unknowingly conflates the objectual and substitutional interpretation of the quantifiers, and also that, on his definition of a function, it turns out to be a certain sort of sign. There is no doubt that this use/mention sloppiness was a significant contributory factor in his decision to treat identity as a “relation between expressions.” Indeed, Frege (1893: 6) indicates as much:

The primitive signs used in *Begriffsschrift* occur here also, with one exception. Instead of the three parallel lines I have adopted the ordinary sign of equality, since I have persuaded myself that it has in arithmetic precisely the meaning that I wish to symbolize. . . . The opposition that may arise against this will very likely rest on an inadequate distinction between sign and thing signified.

Our own view, however, is that while use/mention confusion was a contributory factor, it was Frege's confused notion of *beurtheilbar Inhalt* that was salient in his rejection of the standard notion of identity from arithmetic.

Use/mention error is clearly evident in his positive view that “identity relates expressions.” Here, Furth's observation is pertinent:

[The *Begriffsschrift* theory of identity] has the merit of accounting for the interest of true ‘ $A = B$ ’ as against the uninformativeness of ‘ $A = A$ ’. But the price is exorbitantly high, for the device renders it practically impossible to integrate the theory of identity into the formalized object-language itself; e.g., to state generally such a law as that if  $F(a)$  and  $a = b$  then  $F(b)$ . (Frege 1893: xix)

Furth is too generous in conceding that the *Begriffsschrift* theory accounts for the difference in cognitive value between  $\alpha = \alpha$  and true  $\alpha = \beta$ ; Frege (1892c) thought that it could not. The second part of Furth's criticism is also problematic. Although the first observation is disputable, the second is hardly controversial. Frege held that expressions flanking the symbol for identity of content stood for themselves, although in all other contexts they stood for their ordinary contents; combining two such constructions within the scope of a quantifier entailed quantifying into what was essentially an opaque context. There is no need to go over these difficulties; they are well known. But it is not at all clear what Furth's criticism is here.

It might be thought that the error in *Begriffsschrift* was to have the expressions flanking ‘ $\equiv$ ’ stand for themselves, but there is nothing inherently wrong in this. There is no logical error in having the constants of a first-order theory denote themselves. In fact, a common mathematical

modeling of first-order logic takes the domain of the theory to be the set of constants of that theory, each constant denoting itself.

Again, it might be thought that the error in *Begriffsschrift* was to take the expressions flanking ‘ $\equiv$ ’ to stand for themselves while allowing them to stand for their ordinary contents in all other contexts. But there does not seem to be anything incoherent in this either. Let  $r_o(\eta)$  be the usual reference of a singular term  $\eta$ , and let  $r_i(\eta)$  be its indirect reference (that is, its reference within the context of the symbol for identity of content, namely, itself). We might hold, then, that in  $\alpha \equiv \beta$ ,  $\alpha$  and  $\beta$  stand for themselves, but that

$$\alpha \equiv \beta \text{ is true if, and only if, } r_o(r_i(\alpha)) = r_o(r_i(\beta)). \quad (4.30)$$

There are two ways of understanding the right-hand side of (4.30), corresponding to the two ways of analyzing (4.27). Let us follow (4.28) as our model, so that ‘=’ is understood to express a “relation between objects.” It does, of course. Introducing a two-place predicate into a first-order theory with the usual axioms governing identity and with the usual semantic interpretation characterized by the view that “identity relates objects” is standard operating procedure. Let us reserve ‘=’ for this relation and ‘ $\equiv$ ’ for that which is supposed to “relate expressions.” Anyone who wishes to hold that “identity relates expressions” would most likely also want to preserve the intuitively valid inference patterns involving identity. Frege certainly wanted to: identity of content was supposed to behave syntactically like identity, with only the semantics redone to conform with Principle 4.2.1. What would be the difference between ‘=’ and ‘ $\equiv$ ’? On the substitutional interpretation, none would show up; and this is to be expected because on the substitutional interpretation the usual word/world semantics is bypassed. Syntactically, then, ‘=’ and ‘ $\equiv$ ’ are notational variants. Semantically, a distinction might, perhaps, be made out, though it would be evanescent, for the point of (4.30) is to undo the opacity-inducing effect of ‘ $\equiv$ ’, taking away with the right hand what one has given with the left. Surely there is no compelling reason to suppose that our ordinary notion of identity is better captured by ‘ $\equiv$ ’ than by ‘=’. The exercise might be pointless, but it is not incoherent.

What does appear to be incoherent, however, is the view that “identity relates expressions but not objects,” precisely the view Frege was holding in *Begriffsschrift*. He was preserving the syntax of ‘=’ while, at the same time, denying the possibility of the standard semantics. He was arguing that the right-hand side of (4.30) had to be modelled on (4.29), not (4.28), which is, of course, false; as such,  $r_o(r_i(\alpha))$  could not be

construed as a composite function, and the relational analysis modeled on (4.29) could not be carried through. That is, denying the possibility of the standard semantics of '=' meant that, ultimately, Frege's syntax for '≡' had to differ. This is the result Furth correctly observed.

The failure of Frege's *Begriffsschrift* theory of identity ought by now to be apparent. Frege believed, in *Begriffsschrift*, that substitution of coreferential singular terms preserved conceptual content; this was codified in his symbolism by (4.3), which marks the significant difference in the syntax of '=' and '≡'. But the thrust of Frege's *Begriffsschrift* solution to the Paradox of Identity was to abandon the *Begriffsschrift* Substitution Principle 4.2.1, for he was allowing that  $S\alpha$  and  $S\alpha/\beta$  could differ in conceptual content when  $r(\alpha) = r(\beta)$ : he was arguing that (4.4) was true and (4.5) false, in effect, then, repudiating (4.3). As such, he was bringing the syntax of '≡' back into line with '='. Once he had done so, the standard semantics of '=' would be readily available and there would no longer be any good reason to deny that "identity relates objects." His view that "identity relates expressions and not objects," therefore, either had to turn out to be vacuous (in the sense explained two paragraphs back), or, if nonvacuous, then either he no longer had identity, or, if he did, the view is false.

## Concept and Object

### 5.1 Introduction

It is natural, now, to think of there being connected with a sign (name, combination of words, written mark), besides that which the sign designates, which may be called the *Bedeutung* of the sign, also what I should like to call the *sense* of the sign, wherein the mode of presentation is contained.

Thus Frege (1892c: 152) introduces the sense/reference distinction. A sign, he says, expresses [*Ausdrucken*] its sense and stands for, refers to, denotes [*Bedeutun*] or designates [*Bezeichnen*] its reference. Frege (1892c) confines his discussion to proper names [*Eigennamen*]; but he intended the sense/reference distinction to apply as well to concept words [*Begriffswörter*], and to function-expressions generally. A careful reading of Frege's later writings confirms this, but the decisive evidence is to be found in the unpublished manuscript which the editors entitled "Ausführungen über Sinn und Bedeutung":

In an essay ("On Sense and Reference") I have primarily distinguished between sense and reference only for proper names (or, if one prefers, singular terms). The same distinction can also be drawn for concept words. Now, a confusion can easily develop here, in that one so mixes up the division between concept and object with the distinction between sense and reference, that one runs together sense and concept on the one side, and reference and object, on the other. To each concept word or proper name, there corresponds, as a rule, a sense and a reference, as I am using these words. (Hermes, Kambartel and Kaulbach 1969: 128)

The sense/reference distinction therefore cuts across the function/object distinction, but Frege was not as clear as his words here would

indicate about how the two were related. The prevalent confusion about Frege's semantics for function-expressions is not solely the fault of Frege's commentators. It is due, in part, to Frege's own apparently conflicting demands upon his two central distinctions

We can identify two points at which this tension is most keenly felt.

First, the mature function/object distinction Frege (1893) presents is an amalgamation of two earlier distinctions. On the one hand, there is the concept/object distinction Frege (1884*b*) had drawn in the course of analyzing the notion of cardinal number: *Number*, Frege had insisted, is a property of concepts, not of objects. On the other hand, there is the later distinction Frege (1891) had drawn that was keyed to underpinning the construction of linguistic expressions: here the crucial feature of the distinction was the "unsaturatedness" of functions and the "saturatedness" of objects. Now, Frege (1884*b*) had not distinguished between sense and reference, and so the question naturally arises whether the *Grundlagen* concepts are to be identified with the referents or the senses of function-expressions. Frege (1893), in working out the technical details of his definition of number, identifies concepts with the referents of predicate-expressions; but the arguments for the concept/object distinction in Frege (1884*b*) would lead one to suppose that concepts ought to be identified with the senses of function-expressions. If the sole difference between *Werthverläufe*, which are objects on Frege's view, and concepts, understood as referents of function-expressions, is that the former are saturated and the latter unsaturated, then one is hard put to defend Frege's view of the logical priority of concepts to *Werthverläufe* such that *number* must be viewed primarily as a property of concepts, and only secondarily of *Werthverläufe*.<sup>1</sup> There is something irredeemably epistemological about Frege's (1884*b*) distinction. But epistemology is irrelevant to the function/object distinction. It is relevant rather to the sense/reference distinction, wherein epistemology is relegated to the level of sense. Hence the tendency on the part of Frege commentators to identify concepts with the senses of function-expressions.

A second source of confusion derives from the fact that the function/object distinction is to hold at the level of reference, and also at the level of sense. There is nothing intrinsically wrong with this. But at the level of reference it is interpreted as a function/argument structure, while at the level of sense it is interpreted as a part/whole structure. It is not clear how these two very different instantiations of the function/object structure are supposed to be integrated.<sup>2</sup> Furthermore, because Frege

himself confused levels and sometimes spoke of part/whole at the level of reference, and sometimes spoke of function/argument at the level of sense, the lessons supposed to be learned by his reader are hopelessly obscured.

However, we shall not deal with these difficulties here. Supposing that the function/object distinction does hold at the level of reference, we shall limit ourselves to becoming clear about the role this distinction plays at that level.

## 5.2 Objects

An object [*Gegenstand*] is that sort of entity which is referred to by an *Eigenname*. Not every object need actually have a name, but were any expression to stand for an object it would have to be an *Eigenname*. What is an *Eigenname*? “I call anything a proper name [*Eigenname*] if it is a sign for an object” (Frege 1892*b*: 185). Again, there need not actually be an object for which the *Eigenname* stands – ‘Odysseus’ is an *Eigenname* even though there is no such person as Odysseus. All that is required is that the expression purport to refer to an object. It is not clear to which of these two, *Eigennamen* or *Gegenstände*, Frege assigns priority. He often gives the impression that it rests with *Gegenstände*, as though this were a fundamental category of reality which it is incumbent upon language to reproduce; but we are inclined to agree with Dummett (1981*a*: 55–8) that it must reside with *Eigennamen*. Whichever way out of this circle, however, we remain with a notion that, according to Frege, is logically primitive and indefinable:

[T]he question arises what it is that we are here calling an object. I regard a regular definition as impossible, since we have here something too simple to admit of logical analysis. It is only possible to indicate what is meant [*gemeint*]. Here I can only say briefly: an object is anything that is not a function, so that an expression for it does not contain any empty place. (Frege 1891: 140)

Take the proposition “Two is a prime number.” Linguistically, we distinguish here between a subject, “two,” and a predicative constituent, “is a prime number.” . . . The first constituent, “two,” is a proper name of a certain number; it designates an object, a whole that no longer requires completion. The predicative constituent “is a prime number,” on the other hand, does require completion. I also call the first constituent saturated; the second, unsaturated. To this difference in the signs there of course corresponds an analogous one in the realm of references: to the proper name there corresponds the object; to the predicative part, something I call a concept. This is not supposed to be a definition; for the

decomposition into a saturated and an unsaturated part must be considered a logically primitive phenomenon which must simply be accepted and cannot be reduced to something simpler. (Frege 1971: 32–3)

We are left, therefore, with hints and metaphors, certainly not the most philosophically satisfying position to be in, but since the crude directions Frege provides for sorting expressions (and entities) appear to be sufficient for us to sort expressions (and entities) in the way he intended, we have enough to work with. An *Eigennamen*, he says, is a complete or saturated expression, one that has no empty places so that it can stand alone. Proper names (ordinarily so-called) and singular definite descriptions are *Eigennamen* (though not invariably), as well as declarative sentences [*Behauptungssätze*]. *Gegenstände*, similarly, are characterized as self-subsistent [*selbständig*], saturated [*gesättigt*], complete wholes [*vollständige Ganzen*]. All of the following are *Eigennamen*, and so refer to objects if they refer at all: ‘the moon’, ‘the Equator’, ‘the shape of a glass’, ‘the square of four’, ‘snow is white’, ‘the extension of the concept *horse*’, ‘the first man to have landed on the moon’, ‘Gödel’s first incompleteness theorem’, ‘Babe Ruth’s batting stance’, ‘the sinking of the *Lusitania*’, ‘the whereabouts of the Prime Minister’, ‘John’s having gone to jail’, and, of course, the notorious ‘the concept *horse*’.

The ontological classification of entities into objects and functions is exclusive: it is impossible that there should be anything that is both an object and a function. The distinction, it appears, is exhaustive as well, for Frege says, in one of the passages quoted above, that “an object is anything that is not a function.” Unlike objects, functions are unsaturated [*ungesättigt*], incomplete [*unvollständig*], in need of supplementation [*ergänzungsbedürftig*], unable to stand alone. The sign for a function must have one or more empty places, and so an expression that stands for a function must be what we called in the last section an *incomplete expression* or *function-expression*. A function, then, is that sort of entity that is referred to by an incomplete expression, and only an incomplete expression can stand for a function. As was the case with objects, it need not be supposed that to every function there corresponds an expression that stands for it, but only that were any expression to stand for a function, it would have to be an incomplete expression; and, again, it need not be supposed that there actually is a function that is denoted by an incomplete expression – just as there are *Eigennamen* which purport to refer but fail because there exist no objects to which they refer, so too there might



be function-expressions that purport to refer but fail because there exist no such functions.

### 5.3 The Combining Tie

The distinction between *Eigennamen* and function-expressions, like the distinction between objects and functions, is exclusive. It is not, however, exhaustive, for there are expressions that fail even to purport to refer, namely the syncategorematic expressions – parentheses, variables, and the assertion sign. *Eigennamen* and function-expressions are both, according to Frege, names – Frege (1893) calls function-expressions *Function names* [*Funktionsnamen*] – and they are names because they refer, or, better, purport to refer, to entities. The quoted passage with which we opened this chapter clearly indicates that Frege intended the relation between an *Eigennamen* and the object for which it stands to be the same as the relation between a *Funktionsname* and the function for which it stands: each *bedeutet* its *Bedeutung*.<sup>3</sup> The difference between *Eigennamen* and *Funktionsnamen* is not to be found in the relation they bear to entities.<sup>4</sup> The natural alternative to Frege’s taking the two expressions to refer to different things would have been to take them as referring differently to things.<sup>5</sup> But no, the difference between *Eigennamen* and *Funktionsnamen* is to be found in the kind of entity each refers to.

The difference between objects and functions, analogously, is to be found in the kind of expressions that stand for them. But this way of marking the distinction is not very informative (and, moreover, as we shall soon see, Frege could not coherently state the distinction in this way). The correct place, we think, to locate the difference, both at the formal mode level between *Eigennamen* and *Funktionsnamen*, and at the material mode level between objects and functions, is in their combining properties, for it is here that Frege justifies the distinction:

[I]t may perhaps be made a little clearer why these parts must be different. An object, e.g. the number 2, cannot logically adhere to another object, e.g. Julius Caesar, without some means of connection. This, in turn, cannot be an object but rather must be unsaturated. A logical connection into a whole can come about only through this, that an unsaturated part is saturated or completed by one or more parts. Something like this is the case when we complete “the capital of” by “Germany” or “Sweden”; or when we complete “one-half of” by “6.”

Now, it follows from the fundamental difference of objects from concepts that an object can never occur predicatively or unsaturatedly; and that logically, a

concept can never substitute for an object. One could express it metaphorically like this: There are different logical places; in some only objects can stand and not concepts, in others only concepts and not objects. (Frege 1971: 33–4)

So, for example, among functions Frege marks distinctions which he claims are as fundamental as the distinction between object and function:

Now just as functions are fundamentally different from objects, so also functions whose arguments are and must be functions are fundamentally different from functions whose arguments are objects and cannot be anything else. (Frege 1891: 146)

*Functions of two arguments* are just as fundamentally different from *functions of one argument* as the latter are from *objects*. For whereas objects are wholly *saturated*, functions of two arguments are saturated to a lesser degree than functions of one argument, which now are already *unsaturated*. (Frege 1893: 73)

On the other hand, no such distinctions are made for objects, because all objects have the same combining properties: a function which takes for its argument any object accepts every object as argument.

The distinction between *Eigennamen* and the various sorts of *Funktionsnamen* appears, then, to be directed at capturing the fact that certain expressions can combine to form unified whole expressions while others cannot, that certain sequences of signs can be regarded as complex signs while other sequences of signs are regarded as mere complexes of signs. A string of proper names, for example,

John Harry Tom, (5.1)

is merely a complex of signs; but a proper name followed by a (first-level) predicative expression,

John is happy, (5.2)

combine to form a unified whole expression. Frege accounts for these syntactic regularities by assigning to each *Funktionsname* so many blank spaces and specifying for each blank space what kind of expression can fill it – an *Eigenname*, or a one-place, first-level *Funktionsname*, etc. This classification of expressions amounts, in effect, to a rudimentary Adjukiewicz-type *categorial grammar* in which we have but one basic category,  $\mathcal{E}$  (for *Eigenname*). The various kinds of *Funktionsnamen* are all derived categories. In a categorial grammar, derived categories are represented by ‘fractions’: where  $C_1, \dots, C_n, C_{n+1}$  are categories, basic or derived,

$C_{n+1}/C_1, \dots, C_n$  (5.3)

is the category of those expressions which combine with expressions of categories  $C_1, \dots, C_n$  respectively, to constitute an expression of category  $C_{n+1}$ . So a first-level *Funktionsname* with one argument place which combines with an *Eigenname* to form an *Eigenname* would belong to the category  $\mathcal{E}/\mathcal{E}$ . A first-level *Funktionsname* with two argument places that combines with an *Eigenname* to form an *Eigenname* would belong to the category  $\mathcal{E}/\mathcal{E}\mathcal{E}$ . A second-level *Funktionsname* with one argument place that combines with a first-level *Funktionsname*, that is, an expression that belongs to the category  $\mathcal{E}/\mathcal{E}$ , to form an *Eigenname* would belong to the category  $\mathcal{E}/\mathcal{E}/\mathcal{E}$ , and so on. These categories are pairwise disjoint: no expression belongs to more than one category.

The relevant mode of combination here is, of course, *logical* combination. Frege's is a logical grammar. He has not attempted to characterize the grammatical sequences of some natural language, say English or German. This is evident from his purposeful ignoring of many of the traditional grammatical features of natural language: case systems, verb conjugations, active/passive forms, and the like; and most striking in this regard is his lumping together of singular terms and declarative sentences in the same syntactic category. At any rate, this grammar would certainly be inadequate to the task, for, as Chomsky (1957) has forcefully argued, such a grammar (categorial and phrase-structure grammars are notational variants) cannot capture the grammatical resources of natural language. One might, perhaps, in the current fashion, identify Frege's syntax with deep structure. We strongly doubt this particular identification; but we have neither the desire nor the competence nor the need to enter these waters.

#### 5.4 Logical Grammar

Our reason for stressing that Frege's is a logical syntax is that the admissible combinations are determined on logical rather than on grammatical grounds, and so we cannot simply transfer our grammatical intuitions about whether a given sequence of expressions forms a coherent whole to Frege's syntax. For example,

The present Queen of England is wise is a prime number (5.4)

would not ordinarily be considered a coherent grammatical English sentence. It is, however, perfectly proper on Frege's view, because 'the present Queen of England is wise' is an *Eigenname* and 'η is a prime number' is a first-level *Funktionsname*. Hence, although Frege has provided us with

a syntax, there is a sense in which he has not yet accounted for the fact that certain sequences of expressions combine to form complexes while others do not. At best, we have a systematic taxonomy of expressions. We lack an explanation of the logical role of the categories of expressions that will provide the relevant sense of combining. What is it about these types of expressions that accounts for their syntactic behavior? Indeed, why need there even be different categories of expressions at all? Here, we take it, is where the ontological distinction comes in. The different categories of expressions correspond to different categories of entities, and the combining properties of expressions is explained in terms of the combining properties of the entities represented.

Thus Frege projects these differences among expressions onto the world. "To this difference in the sign," he says in a passage quoted earlier, "there of course corresponds an analogous one in the realm of reference." The same adjectives – "unsaturated," "in need of supplementation," and so on – are applied to functions as well as to *Funktionsnamen*, and the grammatical categories are mirrored one-for-one at the level of ontology. To the category  $\mathcal{E}$  of expressions there corresponds the ontological category of objects. To the category  $\mathcal{E}/\mathcal{E}$  of expressions there corresponds the ontological category of first-level singular functions, and so on. As with the linguistic categories, the ontological categories are pairwise disjoint. This amounts, in effect, to a simple theory of types of entities. The complete entities, objects, which correspond to the linguistic category  $\mathcal{E}$  of expressions, we shall say are of type  $e$ . An incomplete entity belongs to the type  $\langle \tau_1, \dots, \tau_n \rangle$  if, and only if, it may be completed by (and only by) entities of type  $\tau_1, \dots, \tau_n$ , taken in that order. So, for example, the entities corresponding to the linguistic category  $\mathcal{E}/\mathcal{E}\mathcal{E}$  are of type  $\langle e, e \rangle$ ; the entities corresponding to the linguistic category  $\mathcal{E}/\mathcal{E}/\mathcal{E}$  are of type  $\langle e, \langle e \rangle \rangle$ ; and so on.

Nor, incidentally, are these ontological distinctions limited to the realm of references. The whole menagerie reappears in the realm of sense:

The whole [thought] owes its unity to the fact that the thought saturates the unsaturated part, or, as we can also say, completes the part needing completion. And it is natural to suppose that, for logic in general, combination into a whole always comes about through the saturation of something unsaturated. . . . By "compound thought" I shall understand a thought consisting of thoughts but not of thoughts alone. For a thought is complete and saturated, and needs no completion in order to exist. For this reason, thoughts do not cleave to one another unless they

are connected together by something that is not a thought, and it may be taken that this ‘connective’ is unsaturated.<sup>6</sup> (Frege 1923: 538)

Just as is the case for reference, then, we find that the sense of a part of a complex *Eigenname* is part of the sense of the whole *Eigenname*, and at least one part of the *Eigenname* must express an unsaturated entity, that is, a function, in order for the sense of the *Eigenname* to be complete. (This, by the way, is the sort of evidence mentioned at the beginning of this section, which confirms that Frege did intend the sense/reference distinction to apply to function-expressions.)

## 5.5 Metaphors

At the level of language, Frege’s talk of parts and wholes, of saturated and unsaturated entities, and so on, can be cashed in for syntactical rules. But at the level of ontology, this talk is metaphorical and treacherously misleading. As we noted earlier in Section 3.6, Frege himself eventually abandoned the part/whole metaphor at the level of reference.

Much has been made of the obscurity of Frege’s metaphors in the secondary literature – see especially Black (1954) and Marshall (1953). But the moral to be drawn here is either that Frege’s commentators have failed to heed his plea to meet him halfway, or else that the metaphors are genuinely unhelpful – not yet, as has been suggested, that the function/object distinction is itself incoherent. Nor is it fair to conclude with Marshall (1953: 267) that “Frege has taken a linguistic difference to be a rift in nature.” As anyone acquainted with the most elementary parts of mathematics will attest, numbers are very different from functions. Frege was trained as a mathematician, and we are sure that this difference was impressed upon him long before he took up logic. The ontological distinction between numbers and functions was arrived at independent of linguistic considerations, and at the earliest stages of his career Frege would no doubt have understood expressions like ‘the sine function’ and ‘the square function’ to stand for functions. Of course, in Frege’s day the notion of a function had not yet been firmly tied down, and his contemporaries were saying very foolish things about functions, for example, that they were variable numbers, or that they were mental operations, or that they were expressions of a special sort. It was in his attempt to clarify the notion of a function, on which he had based his *Begriffsschrift*, that Frege came to connect the ontological distinction between numbers

and functions with the logico-linguistic distinction between (complete) number-names and (incomplete) arithmetic function-expressions. This analysis was given in Section 5.4, and if we look back we will see that the rift in nature, to use Marshall's phrase, was not derived from his analysis of complex number-names, but rather it was assumed at the outset.

The crux of Frege's analysis is precisely here in the link-up. For the connection between numbers and functions, on the one hand, and complete and incomplete arithmetic expressions, on the other, dovetailed neatly with the naive view of representation Frege had held at least as early as *Begriffsschrift*, namely, that language represents by mimicry, that the structure of language mirrors the structure of the entities represented. On such a view, the different categories of expressions would correspond to different categories of entities in such a way that the combining properties of the expressions would mimic analogous properties of the entities represented. The entities would have to have analogous properties, and for want of a more perspicuous idiom for describing these properties one might (as Frege did) simply transfer the mechanistic vocabulary appropriate for expressions to the entities represented. So a complete expression stands for a complete entity, an incomplete expression stands for an incomplete entity, and in this way the syntactic coherence (incoherence) of a given sequence of expressions is explained by the ontological coherence (incoherence) of the entities represented by the expressions in the sequence. Now this outline is very efficient and very appealing but still, it is only an outline: lacking are independent grounds for supposing that there are entities with the desired properties. Here, then, is where the number/function distinction comes in. For in numbers and functions Frege had the entities whose existence and difference were arrived at independent of the needs of the representation scheme to complete the picture for the language of arithmetic. The generalization of Frege's analysis, as we rehearsed it in Chapter 2, was largely a matter of extending the notion of a function beyond arithmetic, so that the picture could be completed for the whole of language.

The generalized function/object distinction therefore does double duty for Frege: (1) it is an ontological distinction, an extension of the number/function distinction; (2) it is a semantical distinction designed to ground the difference in the logical behavior of *Eigennamen* and *Funktionsnamen*. It is important to keep these two aspects of the function/object distinction separate, for the well-known difficulties that beset the function/object distinction stem largely from the semantical role it is

intended to play. These difficulties surface in the notorious case of the concept *horse*, to which we shall now turn.

### 5.6 The Puzzle of the Concept *Horse*

The background, briefly, is as follows. A contemporary, Benno Kerry, had challenged Frege's claim that the concept/object distinction was exclusive. Kerry had argued that since, on Frege's view, 'the concept *horse*' is an *Eigenname*, it must stand for an object. So there would appear to be at least one concept, the concept *horse*, which is also an object. In reply, Frege conceded that 'the concept *horse*' is an *Eigenname*, and also that it must therefore stand for an object, but he rejected Kerry's assumption that the concept *horse* is a concept.

So, Frege had committed himself to the truth of an exceedingly paradoxical sentence,

The concept *horse* is not a concept. (5.5)

Some explanation was called for, and this was the reason for his essay "On Concept and Object."

Frege's strategy in that essay was to defend the semantic theory which entailed (5.5). As for the apparent contradiction, well, since theory demanded that (5.5) be true, Frege asked his readers simply to accept it – to "meet him half-way" and "not begrudge a pinch of salt." But this appeal is much too dogmatic: a reasonable doubt about the theory remains so long as Frege is unable to dispel the puzzle.

Frege did make some small efforts to tackle (5.5) directly, but they prove unsatisfactory. He attributed the paradoxical character of the sentence to an awkwardness of language, and he cited, as another example of such awkwardness, the following:

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. (Frege 1892*b*: 185)

But, contrary to what Frege says, we do not, simply by calling a given expression a grammatical predicate, deprive it of the property of being a grammatical predicate. The sentence

The grammatical predicate 'is red' is not  
a grammatical predicate (5.6)

is simply false. To be sure, Frege's own example,

“The grammatical predicate “is red” is not  
a grammatical predicate, (5.7)

is true. But, in the first place, it is not clear why we should regard (5.7) rather than (5.6) as the proper analogue to (5.5). In the second place, there is no awkwardness about (5.6), and whatever awkwardness accrues to (5.7) is readily explained. In the third place, since (5.6) is false, we are able to speak about grammatical predicates, whereas, on Frege's view, we do not seem to be able to say anything intelligible about concepts. Also, it is worth mentioning that we have a counterexample to Frege's view that the combining properties of expressions must mirror the combining properties of the items the expressions stand for: ‘is red’ is a grammatical predicate, but ‘the grammatical predicate “is red”’ is not a grammatical predicate, yet the latter expression obviously stands for the former. Finally, setting the analogy aside, Frege's claim that the problems with (5.5) are the result of a mere awkwardness of language just does not seem to capture the depths of his difficulties. In explaining his semantic theory, Frege is constantly using such expressions as ‘the concept *horse*’, intending to speak about concepts. Now one might suppose that this is just a loose way of speaking, and that the awkwardness of referring to a concept by means of an *Eigenname* (or of failing to refer to a concept at all) would be eliminated in some acceptable paraphrase. Frege, however, offers no paraphrase, and he even indicates that none is possible:

I admit that there is a quite peculiar obstacle in the way of an understanding with my reader. By a kind of necessity of language, 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 half-way – who does not begrudge a pinch of salt. (Frege 1892*b*: 192)

Inasmuch, then, as Frege was unable to shake off (5.5), the consensus of the commentators has been that Kerry had hit upon something important, though, we might add, there has been no similar consensus as to what it was that Kerry had hit upon. We shall try, now, to track down the error.

### 5.7 An Analysis of the Puzzle

Let us begin by reconstructing Frege's argument for (5.5). (5.5) is an immediate consequence of Frege's fundamental principle

$\eta$  is a predicate if, and only if,  $\eta$  stands for a concept. (5.8)



(Actually, (5.8) is not quite right:  $\eta$  might be a predicate without there being anything that it stands for – all that is required is that it purports to refer to a concept. But, since Frege himself often overlooked this point, we shall let (5.8) stand as is.) Given the obvious truth

‘The concept horse’ is not a predicate, (5.9)

then, by (5.8), we obtain

‘The concept horse’ does not stand for a concept. (5.10)

Paraphrasing (5.10) as

That which ‘the concept *horse*’ stands for is not a concept, (5.11)

and assuming the uncontroversial identity

That which ‘the concept *horse*’ stands for = the concept *horse*, (5.12)

we get (5.5) by substituting ‘the concept *horse*’ for ‘that which “the concept *horse*” stands for’ in (5.11).

However, there is something wrong with this argument. For, by parity of reasoning, we can show that there are no predicates at all; and surely, however perverse Frege might be in hanging on to (5.5), *this* conclusion is utterly unacceptable. Given the obvious truth,

‘horse’ is a predicate, (5.13)

then, by (5.8), we have

‘horse’ stands for a concept; (5.14)

and paraphrasing (5.14), we have

That which ‘horse’ stands for is a concept. (5.15)

But it is obviously true that

‘That which “horse” stands for’ is not a predicate, (5.16)

and so, by (5.8), we have

‘That which “horse” stands for’ does not stand for a concept. (5.17)

Paraphrasing (5.17) as

That which ‘that which “horse” stands for’  
stands for is not a concept, (5.18)

and assuming the uncontroversial identity

That which ‘that which “horse” stands for’ stands for = that which  
‘horse’ stands for, (5.19)

we obtain,

That which ‘horse’ stands for is not a concept, (5.20)

by substituting ‘that which “horse” stands for’ for ‘that which “that which ‘horse’ stands for” stands for’ in (5.16). Since we have been able to derive contradictory statements – (5.13) and (5.20) – then, by *reductio*, if (5.8) is true, (5.13) must be false. And since the argument does not depend upon any particular characteristic of the predicate ‘horse’, it is iterable for all other predicates.

There is one assumption in the argument we have not yet discussed, namely, the paraphrase

$\eta$  stands for a concept if, and only if, that which  
 $\eta$  stands for is a concept; (5.21)

and we have not mentioned the paraphrase largely because it would appear to be so obvious as not to require any comment. Nevertheless, Frege might object to (5.21). For, in every instantiation, the right-hand side of the ‘if and only if’ must be false: ‘that which  $\eta$  stands for’ is never a predicate, so by (5.8) it can never stand for a concept. Hence, in order for (5.21) to be true, there can be no expression which stands for a concept. It is clear that Frege would not be happy with this result. It ought be noted, however, that (5.21) was also assumed in Frege’s own argument – or, at least in our reconstruction of Frege’s argument – to derive (5.5) from (5.8); and if (5.21) is objectionable, it is objectionable in both arguments. That is, if we reject (5.21) to avoid the unwelcome consequence that there are no predicates, then we shall also have blocked the consequence Frege desired to draw from his argument, namely, that the concept *horse* is not a concept. Alternatively, if Frege does wish to derive (5.5) from (5.8), some such paraphrase is needed in order to link up what an expression denotes with its role in a given sentence. Yet it is difficult to see what other paraphrase would do the job: (5.21) is so intuitively obvious that if we are forced to reject (5.21) in order to keep (5.8), then something must be wrong with (5.8).

Indeed, there is something very wrong with (5.8). Assuming that an expression refers uniquely, we can symbolize (5.21), with the help of

Russell's *iota*-operator,<sup>7</sup> as follows:

$$(\exists!x)(r(\varphi) = x \wedge \text{Concept}(x)) \equiv \text{Concept}[(\iota x)(r(\varphi) = x)]. \quad (5.22)$$

On Russell's view, (5.22) is logically true and this, of course, supports our contention about (5.21). But the important point to note in our symbolization is that we have assumed 'η is a concept' to be a first-level *Funktionsname*, that is, to denote a function that only takes objects for arguments. Now it is fairly clear that Frege, too, is making this same assumption, for he claims that the sentence

$$\text{The concept } \textit{horse} \text{ is a concept} \quad (5.23)$$

is well formed, and also that 'the concept *horse*' is an *Eigenname*. But, if 'η is a concept' is a first-level *Funktionsname*, then it can only be sensefully completed by an *Eigenname*, and, according to Frege, any such completion must result in a false sentence. By the same token, 'η denotes a concept' must always be false. 'Concept' is a *Begriffswort* – it only makes sense to say of an object that it is or is not a concept – and since no object is a concept, there can be nothing of which it is true both that a given expression denotes it and also that it is a concept. Looking back to (5.8), now, we see that the right-hand side of the 'if and only if' will never be true; but since there are predicates, the left-hand side will sometimes be true. So, on the assumption that 'η is a concept' is a first-level *Funktionsname*, (5.8) must be false.

If, then, Frege assumes 'η is a concept' to be a first-level *Funktionsname* – and he clearly does – then the derivation of (5.5) from (5.8) appears to be valid; but since (5.8) turns out to be false on this interpretation, Frege has failed to establish (5.5), and so he has failed satisfactorily to answer Kerry. It is doubtful whether he would be any more successful at establishing (5.5) were he to reject the assumption that 'η is a concept' is a first-level *Funktionsname*, for, as we have seen, (5.21) would then be suspect, and no alternative paraphrase comes to mind that would do the job. Nevertheless, Frege must reject this assumption if his semantic theory is even to get off the ground.

## 5.8 A Solution to the Puzzle

What could have led Frege to suppose that 'η is a concept' is a first-level *Funktionsname*? Certainly there would seem to be good reason for him to think otherwise. He obviously believes that there are concepts, and, no doubt, he would also like to assert that there are concepts; but

by supposing ‘ $\eta$  is a concept’ to be a first-level *Funktionsname*, he has debarred himself – so he acknowledges – from ever truthfully saying so. What could have persuaded him to adopt this self-defeating line? The only reason he offers, as far as we can tell, is that ‘ $\eta$  is a concept’ can be sensefully completed by an *Eigenname*. This, surely, is insufficient. For, although ‘ $\eta$  exists’ can be sensefully completed by an *Eigenname*, Frege is quite clear that a sentence like

The number 2 exists (5.24)

is not to be construed as being of the form *Fa*, with ‘the number 2’ an *Eigenname* and ‘ $\eta$  exists’ a first-level *Funktionsname*. On the contrary, he claims, (5.24) is not about any particular object at all. Rather it is to be understood as expressing that something falls under the concept *being the number 2*, and the proper symbolization of (5.24) would thus be

$(\exists x)(x = \text{the number } 2)$ . (5.25)

Indeed, it is rather surprising that Frege (1892*b*) should so staunchly maintain that the singular definite article invariably signals an *Eigenname*, for, as we see, he had adopted a more flexible attitude elsewhere, and, moreover, his general rule of thumb was that the superficial grammar of a sentence was not always an accurate reflection of its logical structure. In order to establish that ‘ $\eta$  is a concept’ is a first-level *Funktionsname*, then, it is not enough merely to point to the fact that it can sensefully be completed by an *Eigenname* – especially when the interpretation creates such enormous logical difficulties. One must also show that the expression filling the blank space is operating as an *Eigenname* in that context. Since Frege has failed to show this, and since there is no other compelling reason to think that ‘ $\eta$  is a concept’ is a first-level *Funktionsname*, Frege can safely drop the assumption.

But, then, to which syntactic category does ‘ $\eta$  is a concept’ belong? We have mentioned that Frege believes that there are concepts, and also that he would prefer to be able truthfully to assert that there are concepts. Now, as Furth<sup>8</sup> points out, it appears that in his *Begriffsschrift* Frege actually has the expressive power to do so. If we want to assert the existence of a given concept, say, the concept *horse*, we might do so in (modern) symbolic notation as follows:

$(\exists f)(\forall x)(f(x) = \text{Horse}(x))$ . (5.26)

(5.26) appears to have the desired properties. For one thing, it is fairly clear that Frege would grant that (5.26) is true, and for another, we

have avoided the difficulties with the English sentence (5.26) symbolizes, namely,

The concept *horse* exists, (5.27)

where we apparently refer to a concept by means of an *Eigenname* – or, perhaps, fail to refer to a concept at all and only assert the existence of an object. Since no *Eigenname* occurs in (5.27), we can see that the difficulties with (5.26) are only apparent. Logically, ‘the concept *horse*’ is no more serving as an *Eigennamen* in (5.27) than is ‘the number 2’ in (5.26). The use of the singular term in each case is merely a linguistic device to satisfy the demands of English grammar.

Let us pursue this point further. Frege proposed, as a paraphrase of a sentence like

The president is a politician, (5.28)

the sentence

The president falls under the concept *politician*. (5.29)

The syntactic distinction in (5.28) between the singular term and the function-expression is captured in (5.29) with two different types of singular terms, ‘the president’ and ‘the concept *politician*’. Frege appears to be much too dogmatic in insisting that all singular terms belong to the same syntactic category. Had he marked a distinction between singular terms like those in (5.29), then, not only could he have accounted for the truth of

The concept *horse* is a concept, (5.30)

but a sharp distinction among singular terms might very well have made him more alert to a distinction among ‘objects’ that was ignored when he considered objects and *extensions* of concepts to belong to the same logical type.<sup>9</sup>

An interesting consequence of this analysis is that the two sentences (5.24) and (5.27) do not receive the same logical analysis. Although (5.24) is symbolized correctly as (5.25), (5.27) cannot be similarly symbolized as

$(\exists x)(x = \text{the concept } \textit{horse}),$  (5.31)

for (5.31), if it is meaningful at all, expresses the existence of an object, not of a concept. So ‘exists’ turns out to be ambiguous: in some cases it is to be understood to be playing the role of the first-order quantifier

(as, for example, in (5.24)), and in other cases it is to be understood to be playing the role of the second-order quantifier (as, for example, in (5.27)). It is natural, then, to suggest that ‘exists’ has the force of *being an object* in the former case, and of *being a (first-level) concept* in the latter case. If it does, we could then construe ‘ $\eta$  is an object’ as a second-level *Funktionsname*, denoting the first-order quantifier, and we could construe ‘ $\eta$  is a concept’ as a third-level *Funktionsname*, denoting the second-order quantifier.

On this interpretation, (5.23), which Frege had originally claimed to be false, would be symbolized as (5.26), and thus be true; and the paradoxical (5.5), which Frege had claimed to be true, would be symbolized as

$$\neg(\exists f)(\forall x)(f(x) = \text{Horse}(x)), \quad (5.32)$$

and (5.32) would be false. However, in acknowledging that the concept *horse* is a concept, we have not yet conceded victory to Kerry. Kerry’s sentence

$$\text{The concept } \textit{horse} \text{ is an object,} \quad (5.33)$$

if it is meaningful at all, would be symbolized as (5.31), and, so understood, it expresses the existence of an object; while (5.23), understood as (5.26), expresses the existence of a concept. Given the radically different logical treatments of ‘the concept *horse*’ in (5.23) and (5.33), Kerry would have to show that that which is said to be an object (speaking loosely) in (5.33) is one and the same entity as that which is said to be a concept (speaking loosely) in (5.23), in order to complete his argument. But he cannot do this. For somewhere along the line Kerry would have to claim that there is an object  $x$  and a function  $f$  such that  $x = f$ ; and, on Frege’s view, placing the identity sign thus between an object variable and a function variable is incoherent. Actually, it would seem best in this situation simply to deny that (5.33) is meaningful at all. It is entirely consistent with our interpretation that we adopt a suggestion of Geach’s that ‘the concept *horse*’ is never serving as an *Eigenname* in the sense that ‘the president’ does. If a sentence containing such an expression is genuinely about a concept, then this would be captured in the symbolization by our use of *Funktionsnamen* and second-order quantifiers, and the singular term would thus be eliminated. Alternatively, if the term is ineliminable, the sentence containing it would be deemed to be nonsense.

5.9 *Morals*

Certainly, we have here a much more plausible response to Kerry than the one Frege had originally given. But, although we have parried Kerry's counterexample, we have done so (as a little reflection would show) at the cost of rendering Frege's principle "No concept is an object" meaningless: there is just no way of coherently expressing this principle in the symbolism. Now it could be that our interpretation of 'η is a concept' is incorrect, and there is some other which would form the basis of a defensible response to Kerry, but we doubt it. For, if Frege wishes to maintain – as he clearly does – both that objects and concepts are of different logical types, and also that we cannot speak of identity across logical types, then it would appear that he is committed to the fact that "No concept is an object" is nonsense.

Dummett (1955: 269) has remarked:

[I]f Frege had confined himself to talking about these various types of *expression*, instead of that for which they stood, the appearance of paradox, the awkwardness of phrasing, the resort to metaphor, which pervade his writings, would all have been avoided. Frege was quite wrong in pretending that the same ills affect the formal mode of speech.

But Dummett is the one who is wrong: the same ills do affect the formal mode of speech. This is an immediate consequence of Frege's general view that the structure of language mirrors the structure of the world; for, on this view, predicates and concepts must have analogous properties. Let us consider what Frege (1971: 34) has to say about concepts:

It is clear that we cannot present a concept as independent, like an object; rather, it can only occur in connection. One may say that it can be distinguished within, but that it cannot be separated from the context in which it occurs. All apparent contradictions that one may encounter here derive from the fact that we are tempted to treat a concept like an object, contrary to its unsaturated nature. This is sometimes forced upon us by the nature of our language. Nevertheless, it is merely a linguistic necessity.

Now, if predicates and concepts are to have analogous properties, then just as a given concept cannot occur outside some connection, with an object, say, so too the predicate that stands for it cannot occur outside a connection, with, say, an *Eigennamen*. Hence one cannot extract a predicate from some complex term which contains it; a predicate does not form a separable unit. (By the same token, one cannot extract an *Eigennamen* from some complex name which contains it, for this would leave the remaining *Funktionsnamen* in isolation.) But, then, one cannot say of a given

expression that it is a *Funktionsname*, or that it stands for a concept; for in order to do so one would have to consider the expression in isolation – and, in the latter case, one would have to consider the concept in isolation as well: one would have to say, for example, that the expression ‘ $\eta$  is a horse’ denotes the concept  $\eta$  is a horse. One cannot do this, for to denote a concept is to act predicatively.<sup>10</sup> Frege has simply left no room for something’s being a predicate without it acting predicatively. So, on yet another score, (5.8) turns out to be problematic. Frege attributes these difficulties with (5.8) to some “linguistic necessity.” But it is not clear what this amounts to. It is doubtful that Frege is abandoning the view that the structure of language mirrors the structure of the world. This, after all, is his guiding idea about the way in which language represents the world. It is also doubtful that Frege is pointing to a superficial feature of the mode of expression chosen, something that could be eliminated by paraphrase. For he speaks of it as a *necessity*, and, furthermore, he offers no paraphrase. A closely related possibility is that the difficulties arise only for natural language and they would be circumvented in a formalized language like *Begriffsschrift*. It is not clear whether Frege had this in mind, but it is an interesting possibility, and, moreover, it is also one we can defeat handily. For, if the analogy between predicates and concepts is to hold rigorously, as one would expect in an artificially constructed language like *Begriffsschrift*, then just as a concept is a function, so too the predicate that stands for it would have to be a function. And, if so, we shall run into the same difficulty with ‘ $\eta$  is a *Funktionsname*’ as we had earlier with ‘ $\eta$  is a concept’. In particular, if we pursued the same analysis for ‘ $\eta$  is a *Funktionsname*’, we shall have a formal mode analogue to the fact we unearthed earlier, namely, that “No concept is an object,” in the fact that “No *Funktionsname* is an *Eigenname*” would also be nonsense. It is this fact we had in mind when we argued, *contra* Dummett, that the same ills affect the formal mode of speech.

Wittgenstein (1922: 55) clearly had Frege’s predicament about the concept *horse* in mind when he spoke (4.126) about ‘formal concepts’:

When something falls under a formal concept as one of its objects, this cannot be expressed by means of a proposition. Instead, it is shown in the very sign for this object. (A name shows that it signifies an object, a sign for a number that it signifies a number, etc.)

Formal concepts cannot, in fact, be represented by means of a function, as concepts proper can.

Frege’s *concept* and *object* are just such formal concepts. The *Tractatus* story is that we cannot *say* that a given expression stands for a concept (object);



that the expression stands for a concept (object) can only be *shown*. This doctrine has been regarded somewhat askance ever since Ramsey (1931) ridiculed it. Our own objection to Wittgenstein's doctrine is to its reliance on some kind of self-evident communication involved in this showing. But there is a piece of Wittgenstein's story that is quite accurate. When Frege connects the semantical and ontological parts of the concept/object distinction, he does so because he is describing a formal language, one in which syntactic distinctions mark semantical ones. The concept/object distinction is being marked in the language by different types of symbols, and it is this difference in the symbols that regulates permissible inferences. The philosophical strategies of moving to the formal mode or of semantic ascent are of no help with these issues because semantic and ontological features are embodied in syntactical differences. On our view, the fact that we can understand a claim like "No concept is an object" reflects the extent to which the syntax and semantics of our natural language are *not* captured by the formal structures mathematicians and logicians have proffered.

## Names and Descriptions

### 6.1 Introduction

We saw in Chapter 3 that Frege and Russell chose different strategies to deal with the Paradox of Identity. The problem for each was the informative character of definite descriptions. Frege (1892*c*) continued to regard both ordinary proper names as well as definite descriptions as belonging to the same syntactic category: both were *Eigenname*. He identified the informativeness of these expressions with the sense they expressed, but he does not appear ever to have attempted to link up this sense he attached to an *Eigenname* in any systematic way with the semantic role of predicate expressions. Is the sense attached to a proper name to be identified with a concept, or a combination of concepts, denoted by some corresponding predicate? This does not seem right, for concepts are extensionally equivalent while senses are not. Is the sense attached to a proper name to be identified with the sense of a predicate – and if so, how? These are issues Frege simply did not address. Russell (1905), however, met these issues head-on. He took the informativeness of definite descriptions as evidence that they were predicative in nature: he regarded their status as singular terms as a *surface* feature of what are at bottom, logically predicative constructions. Russell had a much more comprehensive theory of definite descriptions than Frege did, one that so struck the philosophical community that its very methodology served, in Ramsey’s words, as a “paradigm of philosophy.” Specifically, Russell (1905: 47) claimed his theory solved three puzzles which “a theory as to denoting ought to be able to solve.” One involves the notion of *identity*, the second involves the notion of *existence*, and the third involves the notion of *truth*. In this

chapter, we will compare and contrast how Frege and Russell treated proper names and definite descriptions and pay particularly close attention to Russell's solution to the Paradox of Identity, redeeming the promissory note we extended in Chapter 3. We will take up the notion of existence in Chapter 7 and the notion of truth in Chapter 8. We will find as we continue our investigation of Frege that a thorough appreciation of Russell's theory of descriptions will provide us with an important tool for understanding Frege's semantic theory.

## 6.2 Russell's Theory of Descriptions

Russell (1905) marked a sharp distinction between *genuine* or *logically proper* names on the one hand, and definite descriptions on the other. A genuine or logically proper name refers to an object and functions solely to introduce that object into the proposition expressed by the sentence containing the name: the meaning of the name is the object it stands for.<sup>1</sup> Definite descriptions, on the other hand, are essentially informative, which is to say that their meaning is secured not by the objects satisfying the predicate but by the properties purportedly ascribed. Most ordinary proper names, he maintained, are not genuine proper names but, on examination, turn out to be *disguised* or *truncated* definite descriptions. This is the basis for what has now come to be known as the *Frege/Russell Description Theory of Names*.<sup>2</sup>

Russell provided an analysis of sentences containing definite descriptions on which they were seen to “lack meaning in isolation.”

It is of the utmost importance to realize that ‘the so-and-so’ does not occur in the analysis of propositions in whose verbal expression it occurs, that when I say ‘The author of *Waverley* is human’, ‘the author of *Waverley*’ is not the subject of that proposition, in the sort of way that Scott would be if I said ‘Scott is human’, using ‘Scott’ as a name. I cannot emphasize sufficiently how important this point is, and how much error you get into metaphysics if you do not realize that when I say ‘The author of *Waverley* is human’ that is not a proposition of the same form as ‘Scott is human’. It does not contain a constituent ‘the author of *Waverley*’.  
(Russell 1918: 251–2)

A sentence of the form ‘The *F* is *G*’, where ‘The *F*’ is a definite description, is not assigned the same logical form as is a sentence of the form ‘*s* is *G*’, where *s* is a proper name. In this latter case, the proper name is treated logically as an individual constant, and the sentence is assigned the logical form *Gs*. By contrast, Russell (1905) analyzed a sentence of the form

‘The *F* is *G*’, for example,

The present King of France is bald, (6.1)

into a conjunction of three clauses:

There is at least one thing that kings France, (6.2)

There is at most one thing that kings France, (6.3)

That thing is bald. (6.4)

(6.2) is the *existence clause*, (6.3) is the *uniqueness clause*, and (6.4) is the *predication clause*. Every sentence of the form ‘The *F* is *G*’ is analyzed in this way. Sometimes it is expressed in a logically equivalent but shorter form, *One and only one thing Fs and that thing Gs*. In this shortened form, (6.1) becomes *One and only one thing kings France and is bald*.

For Russell (1905), a phrase like

*The King of France* (6.5)

belongs with the quantified phrases

*Every King of France*, (6.6)

*Some King of France*, (6.7)

*No King of France*. (6.8)

These are none of them *directly referring expressions* like genuine proper names, which simply stand for objects; and the usual quantified sentences involving these phrases are not to be understood as “subject/predicate” sentences.<sup>3</sup> Were we to complete any of these expressions with the predicate ‘is bald’, we would not thereby have created a sentence in which the numbered expression stands for something and ‘is bald’ attributes something to it. This is clearest for (6.8): we cannot suppose it refers to *no King of France* about whom we then say something, namely, that he is bald. Who could we possibly be ascribing this to?

*Everything*, *nothing*, and *something* are not assumed to have any meaning in isolation, but a meaning is assigned to *every* proposition in which they occur. This is the principle of the theory of denoting I wish to advocate: that denoting phrases never have any meaning in themselves, but that every proposition in whose verbal expression they occur has a meaning. (Russell 1905: 42–3)

A definite description like ‘the present King of France’ is not a proper name. Accordingly, a sentence like (6.1) is not an atomic sentence, or, in the terminology of the day, it is not a subject/predicate sentence. Not,

of course, because ‘ $x$  is bald’ fails to be a predicate. No, unlike ‘ $x$  exists’, whose status as a predicate has generated considerable controversy, ‘ $x$  is bald’ is unproblematically a predicate. The sentence fails to be a subject/predicate sentence because ‘the present King of France’ is not a subject expression. As a result, we cannot, as Frege did, insert ‘the present King of France’ into the function-expression ‘ $x$  is bald’ to form a complete sentence. On Russell’s view, ‘the present King of France’ – contrary to appearances – is not a coherent, meaningful unit of a sentence. It is no more a coherent, meaningful unit of the sentence than ‘of France is’. It only has meaning, as Russell puts it, as part of the larger container. So the smallest unit containing a definite description that is meaningful is a full sentence.

Descriptions are eliminated *contextually*. Russell (1905) provides no explicit definition of ‘the  $F$ ’, but there is a way of paraphrasing sentences involving ‘the  $F$ ’ in which the apparent referring expression has been eliminated.<sup>4</sup> So, for example, we speak not of ‘the King of France’ but of ‘a thing that kings France’. Second, descriptions are not names, and so the logical rules governing names need not, unless otherwise argued, apply to descriptions. In particular, substitution rules that apply to names need to be modified when applied to descriptions.

### 6.3 The Scope Distinction

Although definite descriptions are eliminated from any context in which they occur, their persistence in everyday speech compelled Russell eventually to include in his symbolism a notation that captured the illusion that they were singular terms.<sup>5</sup> In advance of this concession to everyday speech, Russell nonetheless found it necessary to highlight a scope distinction, perhaps the most revolutionary and important aspect of his theory. Russell (1905) spoke only of a *primary* and *secondary* occurrence of descriptions. The full scope distinction did not make its entrance until Whitehead and Russell (1910). Here is how it goes. The sentence

The present King of France is *not* bald, (6.9)

can be interpreted in two distinct ways. It might be the case that the *not* is only operating on the third conjunct:

There is at least one thing that Kings France, and  
there is at most one thing that Kings France, and  
that thing is *not* bald. (6.10)

In this case, the description is said to have *large* or *wide* or *broad* scope. The sentence says that one and only one thing kings France and that thing has the property of *being not-bald*. The other reading takes the *not* to be operating on the conjunction:

$$\begin{aligned} &\text{It is } \textit{not} \text{ the case that:} \\ &[\text{there is at least one thing that kings France, and} \\ &\text{there is at most one thing that kings France, and} \\ &\text{that thing is bald.}] \end{aligned} \tag{6.11}$$

In this case, the description is said to have *small* or *narrow* scope. Sentence (6.11) denies that one and only one thing kings France and is bald, and it is therefore the negation of (6.1). Sentence (6.10) affirms that one and only one thing presently kings France, but denies that that thing is bald. So it is inconsistent with (6.1), but it is not its negation because both are false.

Russell introduced a special notation to capture the superficial grammar of a description as a singular term: the ambiguity of everyday constructions is eliminated in the symbolism wherein scope issues are clearly marked. Here is how Whitehead and Russell (1910: 172) put it:

[W]riting ‘ $(\iota x) (\Psi x)$ ’ for ‘the term  $x$  which satisfies  $\Psi x$ ’,  $\Phi(\iota x)(\Psi x)$  is to mean  $(\exists b) : \Phi x. \leftrightarrow x.x = b : \Psi x$ . This, however, is not yet quite adequate as a definition, for when  $(\iota x)(\Phi x)$  occurs in a proposition which is part of a larger proposition, there is doubt whether the smaller or the larger proposition is to be taken as the “ $\Psi(\iota x)(\Phi x)$ .” . . . In order to avoid ambiguities as to scope, we shall indicate the scope by writing “[ $(\iota x)(\Phi x)$ ]” at the beginning of the scope, followed by enough dots to extend to the end of the scope. . . . Thus we arrive at the following definition:

$$14.01 \quad [(\iota x)(\Phi x)].\Psi(\iota x)(\Phi x). =: (Eb) : \Phi x.x.x = b : \Psi b \text{ Df}$$

Whenever  $(\exists y)((\forall x)(Fx \equiv (x = y)))$ , ‘ $(\iota x)(\Phi x)$ ’ behaves, formally, like an ordinary argument to any function in which it may occur,

[P]rovided  $(\iota x)(\Phi x)$  exists, it has (speaking formally) all the logical properties of symbols which directly represent objects. Hence when  $(\iota x)(\Phi x)$  exists, the fact that it is an incomplete symbol becomes irrelevant to the truth-values of logical propositions in which it occurs. (Whitehead and Russell 1910: 180)

So, when a definite description denotes an existent, the description acts just like a proper name in a truth-functional context, and one can be substituted for another codenotational name and preserve truth value.

This is just Frege's Substitution Principle for Reference – actually the *Corrected* Principle 2.5.4.

[W]hen  $(\exists!)(\iota x)(\Phi x)$ , the scope of  $(\iota x)(\Phi x)$  does not matter to the truth value of any proposition in which  $(\iota x)(\Phi x)$  occurs. (Whitehead and Russell 1910: 184)

That is, provided that the  $\Phi$  exists, and also that ' $\Psi$ ' is a context of the requisite kind,

$$[(\iota x)\Phi x]\Psi(\iota x)\Phi x \equiv \Psi[(\iota x)\Phi x](\iota x)\Phi x.$$

Note, however, that this is not to say that when the  $\Phi$  exists, 'the  $\Phi$  is  $\Psi$ ' will express the very same proposition as ' $\Psi a$ ' (where ' $a$ ' is to be a proper name for the  $\Phi$ ). For, on Russell's view, the object itself enters into the proposition expressed by the atomic sentence ' $\Psi a$ '; the proposition is *object dependent*.<sup>6</sup> The proposition expressed by 'the  $\Phi$  is  $\Psi$ ', on the other hand, is *object independent*: it does not contain the object as the appropriate constituent, but the denoting complex. Whatever the story about the constituents of propositions, Russell quite clearly made sure that there was a distinction in logic to avoid the improper implication that they express the same proposition when *the*  $\Phi$  exists. For, as Smullyan (1948) reminds us, it is only when we are working with *logical propositions* that the equivalence holds. When, however, our construction is embedded in a context that is not truth functional, it does not hold.

Whitehead and Russell (1910) introduced a cumbersome notation for capturing scope distinctions. We use the notation introduced by Fitting and Mendelsohn (1998), modeled on lambda-abstraction.  $\langle \lambda x.Fx \rangle$  is termed a "predicate abstract." A predicate abstract attaches to a singular term:  $\langle \lambda x.Fx \rangle(a)$  says that the object  $a$  has the property *being F*. We now distinguish the two formulas:

$$\neg \langle \lambda x.Fx \rangle(a), \tag{6.12}$$

$$\langle \lambda x.\neg Fx \rangle(a). \tag{6.13}$$

(6.12) denies the claim that  $a$  has the property *being F*; (6.13) says *of a* that it has the property *being not-F*. The former is the small-scope reading of the description; the latter is the large-scope reading of the description.

Russell's small-scope construction is very clearly similar to Frege's construction for 'that' clauses, but not exactly so. Each regarded the construction as affecting the reference of expressions inside the 'that' clause, but Frege went further than Russell in taking the expressions to shift

their reference to their senses. There is no clear analogue in Frege for Russell's large-scope construction.<sup>7</sup>

#### 6.4 Russell's Three Puzzles

##### *Truth*

Let us follow up our discussion of the scope distinction with the puzzle to which the example we have been using is immediately relevant, a puzzle involving the notion of truth. At risk is the Law of Excluded Middle, which requires of every proposition that it be either true or false. Russell sought to protect this law.

Here is the problem. Consider the pair of sentences (6.1) and (6.9). Now, says Russell, if we examine all the things that are bald, we will not find the present King of France among them. If we examine all the things that are not bald, we will not find the present King of France among them. So, apparently, the present King of France is neither bald nor not bald. This violates the Law of Excluded Middle.

Let  $Bx$  abbreviate the predicate  $x$  is bald. Let  $k$  abbreviate *the present King of France*. Using predicate abstract notation, (6.1) is entered this way:

$$\langle \lambda x. Bx \rangle (k). \quad (6.14)$$

But the sentence (6.9) is ambiguous: it can be understood in either of the following two ways:

$$\langle \lambda x. \neg Bx \rangle (k), \quad (6.15)$$

$$\neg \langle \lambda x. Bx \rangle (k). \quad (6.16)$$

It is the small-scope reading, (6.16), that is the contradictory of (6.14). The large-scope reading, (6.15), is the contrary of (6.14): both can be false, and both will be false if there is nobody presently kinging France. So (6.1) says that a unique object has a particular property and (6.16) denies that there is a unique object having that particular property. Exactly one of these is true; exactly one of these is false. Note that (6.16) does not ascribe a property to a nonexistent object. If it did, it would be false, because on Russell's (1905) view, nonexistent objects do not have any properties.<sup>8</sup> That is why (6.15) is false: it ascribes the property *being not bald* to a nonexistent object.<sup>9</sup>



*Identity*

Russell introduces Frege's Paradox of Identity in a slightly different form. We shall present it in Russell's (1905: 47–8) own words:

If *a* is identical with *b*, whatever is true of the one is true of the other, and either may be substituted for the other in any proposition without altering the truth or falsehood of that proposition. Now George IV wished to know whether Scott was the author of *Waverley*; and in fact Scott *was* the author of *Waverley*. Hence we may substitute *Scott* for *the author of 'Waverley'*, and thereby prove that George IV wished to know whether Scott was Scott. Yet an interest in the law of identity can hardly be attributed to the first gentleman of Europe.

George IV, Russell says, wanted to know whether Scott was the author of the *Waverley* novels. Presumably, this is true:

George IV believed that Scott is Scott. (6.17)

After all, that Scott is Scott is just an instance of the Law of Identity, something George IV certainly was aware of and certainly believed to be true. So what was the source of his inquiry? Presumably, despite the truth of (6.17), this was false:

George IV believed that Scott is the author of  
the *Waverley* novels. (6.18)

But, given the truth of

Scott = the author of the *Waverley* novels, (6.19)

and Frege's Substitution Principle for Reference 2.5.1, (6.17) and (6.18) should both have the same truth value.<sup>10</sup>

Russell's quick way out of the paradox is to deny that there is any substitution going on here:

The proposition 'Scott was the author of *Waverley*', ... in its unabbreviated form ... does not contain any constituent 'the author of *Waverley*' for which we could substitute 'Scott'. (Russell 1905: 51–2)

But this is not the whole of the story. For, he continues:

This does not interfere with the truth of inferences resulting from making what is *verbally* the substitution of 'Scott' for 'the author of *Waverley*', so long as 'the author of *Waverley*' has what I call a *primary* occurrence in the proposition considered. (Russell 1905: 52)

So Russell's handling of this apparent substitution failure essentially involves the scope distinction. If the description has large scope, then substitutivity is preserved – so long as the thing referred to exists. But if the description has small scope, substitutivity is not necessarily preserved even if the thing referred to exists, for example, as we pointed out earlier, when the description is embedded in a context that is not a truth-functional context.<sup>11</sup>

How does Russell handle Frege's Paradox of Identity? As a matter of fact, his solution is quite similar to Frege's.<sup>12</sup> The failure of the substitution of coreferential singular terms to preserve truth value in the case of the two sentences under the small-scope interpretation – (6.17) and (6.18) – is taken as evidence that neither is *about* Scott in any direct way. They would be about him on the large-scope interpretation, so Russell is actually partial to Frege's Corrected Substitution Principle for Reference 2.5.4. Russell, however, is marking a distinction that it is quite important for us to recognize: when a definite description has large scope, then it, like a logically proper name, is *about* that which the term *denotes*. But only in the case of a logically proper name does the term also *refer* to its object.

Russell's treatment of the problem only works when a sentence is part of a larger construction, however. What about the identity sentence taken by itself? This poses a bit of a problem. For, of course, Russell maintained the *Begriffsschrift* Substitution Principle 3.3.1 that generated the paradox discussed in Chapter 3. Here is another example of this sort with two true identities:

Bertrand Russell = the author of *The Principles of Philosophy*, (6.20)

Bertrand Russell = Bertrand Russell. (6.21)

Assuming *Begriffsschrift* Substitution Principle 3.3.1 and the view that identity relates objects, it would appear, once again, that there can be no informative identities. For the two terms – 'Bertrand Russell' and 'the author of *The Principles of Philosophy*' – appear to be coreferential, so the two sentences, (6.20) and (6.21), should have the same cognitive value. How can Russell avoid the unwanted conclusion?

The answer resides in the fact we have already noted, namely, that Russell takes the *Begriffsschrift* Substitution Principle 3.3.1 to hold only for directly referential expressions like genuine proper names. But it does not, on his view, hold for definite descriptions or ordinary proper names that are really definite descriptions in disguise. *Direct reference* is

essentially a story about meaning, not reference: in the case of a directly referential expression, the meaning *is* its reference. So, in Russell's hands, cognitive value is preserved when we substitute one directly referential term for another referring to the same thing because they have the same meaning.

The important point to take away from this discussion is that Frege and Russell clearly agree on the matter of the *Bedeutung* of singular terms.<sup>13</sup> And they disagree on the *Sinn* of a singular term only in that Russell clearly admits of cases in which the *Sinn* of a term is its *Bedeutung*, namely, cases of direct reference, whereas Frege very clearly does not. Given Russell's (1905) distinction between *reference* – which is *direct reference* – and *denotation*, Frege's notion of *Bedeutung* corresponds most closely to Russell's use of *denotation*, what a sentence is *about*, not *reference*. We read Russell's (1905: 493) *Principle of Acquaintance*,

[I]n every proposition that we can apprehend (i.e. not only in those whose truth or falsehood we can judge of, but in all that we can think about), all the constituents are really entities with which we have immediate acquaintance,

to say that all the categorematic elements of the sentence – not just the logically proper names – present directly the elements of a proposition. Understanding Russell's framework this way, the *Sinn* of the sentence is the proposition expressed, so, on his view, true identities differ so long as the items denoted are not directly referred to, that is, included in the proposition.

### *Existence*

On the one hand, Russell operates with Mill's (1843) assumption that a proper name – a genuine proper name – has denotation but no connotation. On the other hand, he believes most ordinary proper names are not genuine proper names at all, but rather disguised descriptions:

In a detective story propositions about “the man who did the deed” are accumulated, in the hope that ultimately they will suffice to demonstrate that it was A who did the deed. We may even go so far as to say that, in all such knowledge as can be expressed in words – with the exception of “this” and “that” and a few other words of which the meaning varies on different occasions – no names, in the strict sense, occur, but what seem like names are really descriptions. We may inquire significantly whether Homer existed, which we could not do if “Homer” were a name. The proposition “the so-and-so exists” is significant, whether true or false; but if *a* is the so-and-so (where “*a*” is a name), the words “*a* exists” are meaningless. It is only of descriptions – definite or indefinite – that existence can be significantly asserted; for, if “*a*” is a name, it *must* name something: what

does not name anything is not a name, and therefore, if intended to be a name, is a symbol devoid of meaning, whereas a description, like “the present King of France,” does not become incapable of occurring significantly merely on the ground that it describes nothing, the reason being that it is a *complex* symbol, of which the meaning is derived from that of its constituent symbols. And so, when we ask whether Homer existed, we are using the word “Homer” as an abbreviated description: we may replace it by (say) “the author of the *Iliad* and the *Odyssey*.” The same considerations apply to almost all uses of what look like proper names. (Russell 1919: 178–9)

Russell (1905) sought to explain how a nondesignating name could be meaningful, and therefore how any sentence containing such a name could also be meaningful. Because he believed that genuine proper names were directly referential, a nondesignating proper name would not simply lack a referent. It would lack meaning as well.

Russell (1903*b*) also held that proper names were directly referential. To say that a term is nondesignating is to say that the item purportedly referred to by the term does not exist. ‘Pegasus’ is a nondesignating term. So is the planet ‘Vulcan’, ‘the ether’, which scientists postulated toward the end of the nineteenth century as the medium through which light radiated, and, of course, ‘the present King of France’. Because these expressions were meaningful, they had to refer to something, and since there was nothing that existed answering to the term, they had to refer to something that did not exist. Russell’s (1903*b*) strategy was to enlarge the scope of reality and include not only things that exist but also things that had *being*. This strategy, which we associate most closely with the Austrian philosopher Alexius Meinong, has a very serious flaw that Russell (1905) himself pointed out. Contradictions in language are meaningful and accepted by all. Contradictions in reality are not. If we require a being in reality to correspond to every meaningful expression, we are thereby committed to the view that *there are* these contradictory objects corresponding to these contradictory terms.<sup>14</sup> Meinong (1904) explicitly embraced this position. Russell (1905), however, reflecting the majority view, found it unacceptable. He required an alternative account of the meaningfulness of these terms.

With this brief introduction, we turn to consider Russell’s treatment of the Paradox of Nonbeing, in many ways the most pressing of the puzzles and the one Russell is most famous for. Here is how it goes:

If a person denies the existence of something, he must refer to it. If he refers to something, it must exist. So, if a person denies the existence of something, that thing must exist.

Denials of existence, therefore, are at best false and at worst meaningless.<sup>15</sup> Consider:

The present King of France does not exist. (6.22)

According to Russell, (6.22) is not referring to something, the present King of France, and saying of him that he does not exist. On the contrary, it says that there is nothing that presently kings France. (6.22) is the denial of

The present King of France exists. (6.23)

The analysis of (6.23), unlike the analysis of (6.1), only has two clauses. There is no predication clause because ‘*x* exists’ is assumed not to be a predicate; only the existence and uniqueness clauses remain:

There is at least one thing that presently kings France, and (6.24)

there is at most one thing that presently kings France. (6.25)

Since there is no third clause, we do not have two places available in which to insert the negative particle when interpreting (6.22). Russell assumes not simply that ‘the present King of France’ is not a subject expression, but further that ‘*x* exists’ is not a (first-order) predicate. He need not have made this further assumption, but we defer discussion of this matter until Chapter 7.

We cannot resist noting here the importance of the scope distinction. The description in (6.22) can only be understood with small scope. The large-scope reading is syntactically impossible. So we find, very clearly, that it is the scope distinction that enables Russell to solve the problems he does, and it is the scope distinction that is the crucial element of his theory of descriptions.

## 6.5 Frege and Russell on Definite Descriptions

Early on, Frege (1879) recognized the importance of identities in mathematics. The lengths of the two legs of a right triangle *a* and *b* and the length of the hypotenuse *c* are related by a beautiful and important identity known as the *Pythagorean Theorem*:  $a^2 + b^2 = c^2$ . Frege (1879) justified including the identity sign among his logical connectives for just this purpose, that is, to express informative identities. Oddly, however, Frege (1879) proposed no machinery for expressing such interesting identities. Nor was any included in Frege’s (1892*c*) presentation and defense

of the sense/reference distinction.<sup>16</sup> We had to wait until Frege (1893; Section 11) for a definite description operator, and this he defined in terms of class abstraction:

DEFINITION 6.5.1 (FREGE'S DEFINITE DESCRIPTION OPERATOR)

- (I) 'the  $F$ ' denotes  $x$  if  $x$  is the sole element of the set  $\{x \mid Fx\}$ ;
- (II) 'the  $F$ ' denotes the set  $\{x \mid Fx\}$  otherwise.<sup>17</sup>

Frege (1892*b*) had acknowledged that a proper name might have a sense and no reference – 'Odysseus', for example – and in that case, he argued, a declarative sentence containing that name – 'Odysseus was set ashore at Ithaca while sound asleep' – would express a thought but lack a truth value. That text is famous for espousing the view that some propositions lack a truth value, and for the view embraced many years later by Strawson (1950) that the existence of a referent is *presupposed*, not *implied*. However, Frege (1893) banned declarative sentences that lacked truth values, and given the compositional connection between the reference of a sentence and the reference of its parts, he required that every *Eigenname* in his *Grundgesetze* system have a reference.

A definite description is, for Frege, an *Eigenname* – a proper name – and the truth value of a sentence containing a definite description is determined in much the same way as that of any sentence containing a name. A sentence of the form  $Fa$  will be true if, and only if, the object referred to by the *Eigenname*  $a$  falls under the concept referred to by  $F$ . It was important for Frege (1893) that a definite description should always stand for something so that a sentence containing it would have a truth value. This was, in effect, his consistency proof in *Grundgesetze*, namely, that there should correspond an object to every name constructible in his system.<sup>18</sup> In the case that one and only one thing is  $F$ , then, as in ordinary language, 'the  $F$ ' will stand for that thing. But what to do if nothing is  $F$  or if more than one thing is  $F$ ? In these cases, Frege arbitrarily assigned it a reference just so that it had one.

The reader must recognize that Frege's motivation in formulating a definite description operator is significantly different from Russell's. As a solution to the Paradox of Nonbeing, for example, Frege's (1893) is an abject failure. Because he was so intent in *Grundgesetze* on assuring that every *Eigenname* have a reference, then in accordance with his Definition 6.5.1 of the description operator, a sentence like

The ancient Greek who first typeset the *Iliad* does not exist (6.26)

turns out to be false. For, since there is no ancient Greek who first typeset the *Iliad*, the description stands for the set. And the set, clearly, exists. On Frege's account, every nondesignating singular term arbitrarily designates the empty set. A more benign, but no less troublesome, consequence is that these two identities both turn out to be true:<sup>19</sup>

George Washington's eldest son = Bill Clinton's eldest son, (6.27)

George Washington's eldest son = George Washington's  
eldest son. (6.28)

However, there is an aspect of Frege's (1893) treatment of definite descriptions that makes it more like Russell's. For he finally incorporated descriptions into the formal syntax of *Grundgesetze* so that theorems involving them could be formulated and proved. Although Russell regarded both his definite description and class abstraction operators as generating incomplete symbols, Whitehead and Russell (1910) did not define the former in terms of the latter. Frege (1893), by contrast, introduced his class abstraction operator as primitive, and used it to define his definite description operator. Where Russell connected up descriptions with predicates directly, Frege connected them up via classes. But at last Frege was able to construct complex *Eigennamen* out of predicative expressions, and so provide a technical device to construct definite descriptions. The two men are so close, it is a wonder that Frege did not identify the informativeness of the description with the concept used to pick out the object. He could not. The problem for Frege is that concepts are extensionally identified, so any two extensionally equivalent concepts have the same cognitive value. Russell did not have that problem because his "concepts," if we might so call them, were attributes: strip a name from a sentence and what is left designates a *propositional function*, not, as Frege would have it, a truth function.

Frege's (1893) view of descriptions can be made to look quite like Russell's (1905), differing only in the second clause. Russell's (1905) definition goes like this:

DEFINITION 6.5.2 (RUSSELL'S DEFINITE DESCRIPTION OPERATOR)

- (I) 'The *F*' denotes *x* if *x* is the sole element of the set  $\{x \mid Fx\}$ ;
- (II) 'the *F*' does not denote anything otherwise.

Presented this way, Russell's account differed from Frege's in that Russell treated designation for a definite description as a *partially* defined

function. Unlike Frege, he did not require that the description always denote. But, since he accepted the Law of Excluded Middle, he had to say something about the way in which the truth value of a sentence containing a definite description was to be determined. It is worth remarking, further, that presented this way, Russell (1905) is rejecting Frege's Compositionality Principle for Reference 2.3.1. For, if the sentence contains a nondesignating description, then the reference of the containing sentence is *not* determined by the reference of the contained description.

Let us recall examples (6.12) and (6.13). These two will not differ in a classical, that is, nonmodal, context unless *a* fails to denote. For if it denotes nothing, then it will not denote anything that has the *F* property, and so (6.12) is true. But if it denotes something, then it will not denote anything that has the property of not being *F* so (6.13) is false. Adapting language from modal logic, we might speak of (6.13) as a *de re* attribution of a property to *a*. We can characterize Russell's story about truth as follows: If a singular term *a* stands for something, then a sentence of the form *Fa* is true if, and only if, the object referred to by the *Eigenname* *a* falls under the concept *F*; if the term *a* fails to stand for anything, then every *de re* attribution of a property to *a* is false.

This particular characterization of the *iota*-operator shows Russell's to be an early version of *Free Logic*. Russell's scope distinction enables us to handle situations where we have putative singular terms, putatively designating objects, which do not actually succeed in doing so. We can, at least on the surface, treat them as singular terms, but the sentence is not ascribing a property to the object, there being none, and so its truth does not depend upon whether the object has or lacks the property. Russell saw the matter as one in which failure of designation corresponded to failure of existence. In fact, as we noted in Section 6.3, in truth-functional contexts, the two are essentially indistinguishable. But in non-truth-functional contexts, the two are distinguishable. In truth-functional contexts, an object exists if it has any property; but in non-truth-functional contexts, it has being if it has any property, and existence if the having of that property is unqualified.<sup>20</sup>

We find Russell's theory a good deal more intriguing than Frege's, and largely because of the introduction of the scope distinction, which has proved so fruitful in understanding other logical phenomena, especially, in recent times, the role of singular terms in modal contexts. It is important to recognize that the scope distinction is a vital and intrinsic component of the theory, not an afterthought. For it is the critical difference between Russell and Frege, both of whom sought to maintain



the classical logic principles (i) that every proposition is either true or false, and (ii) that a proposition is true if, and only if, its negation is false. Frege preserved these principles by arbitrarily assigning a denotation when none satisfied the description. By allowing singular terms only partial denotation, Russell preserved these principles by manipulating compositionality via the scope distinction.

## Existence

### 7.1 Introduction

The matter of *existence* is one of the most difficult in philosophy. The topic is infused with a particularly noxious mix of dogma and confusion. Needless to say, Frege's influence on modern thought is deep. There are three distinct aspects of the issue on which he made contributions:

- First, there is the *Context Principle* – “never to ask for the meaning of a word in isolation, but only in the context of a proposition” (Frege 1884*b*: x) – which he employed to promote his own view that numbers are objects and to undermine the then current psychologism in mathematics;
- Second, there is his treatment of nonreferring singular terms and the truth value of sentences containing them;
- Third, and perhaps most significantly, there is his doctrine that *existence* is a property of properties, not of things.

We have little to offer that will help clear the general fog about Frege's Context Principle and its application.<sup>1</sup> We spoke about Frege's treatment of nonreferring singular terms in Chapter 3, and again in Chapter 6. We will say more in Section 7.6. However, we will focus in this chapter primarily on the third issue, namely, whether existence is a first-order property.

Frege is widely credited with providing a precise interpretation in the language of modern logic of Kant's (1781: 504) well-known declaration: “‘*Being*’ is obviously not a real predicate. . . .”<sup>2</sup> Frege's discussion of the

issue is self-consciously derived from Kant's, and framed with the same explicit connection to the Ontological Argument for God's existence. Frege's view about existence was adopted by Russell, and we speak in Section 7.2 of a Frege/Russell view about existence. But the similarity between the Frege/Russell view, on the one hand, and Kant's view, on the other, has limits. To set the record straight, we return in Section 7.7 to look more closely at what Kant actually said.

Frege's idea, as we have mentioned, was to treat *existence* as a property of properties, not as a property of objects. He took the existential quantifier, which in his categorial grammar stood for a property of properties, to express what we ordinarily want to express when we say that a thing *exists*. He provided little in the way of argument for this position in his published writings, but important insights into his reasoning can be found in the posthumously published dialogue he engaged in with his colleague at Jena, the theologian Bernard Pünjer. In this little-studied piece, we find Frege (1884*a*) urging that the quantifier should be understood to be carrying existential import in order to explain how there can be informative existence claims. This dialogue has two important consequences for our understanding of Frege: first, it frames the problem about existence in a way analogous to the problem about identity, and second, it sets out in sharp relief how differently Frege treated these two notions.

The problem about *existence* is strikingly similar to the better known Paradox of Identity we considered in Chapter 3. Parallel to Frege's (1892*c*) question,

*How are informative identity claims possible?*

Frege (1884*a*) addresses the question,

*How are informative existence claims possible?*

The philosophical literature has treated Frege's Paradox of Identity differently from the Paradox of Nonbeing. This is incorrect. They are much more closely related than has otherwise been recognized. The Paradox of Nonbeing is formulated in such a way that the negative existential poses the problem. The Paradox of Identity, on the other hand, is formulated with the positive identity claim the problem. Let us formulate them both positively, just to maintain the analogy. (We could just as well have formulated them both negatively.)

The Paradox of Identity goes like this. How can there be informative identities? Assuming identity relates objects, if  $a = b$  is true, then it is

trivially true (that is, it must have the same cognitive value as  $a = a$ ), because self-identity is a condition for referring to an object. The self-identity condition renders  $a = a$  true, and the direct reference condition renders it indistinguishable from true  $a = b$ .

The Paradox of Nonbeing goes like this. How can there be informative existence claims? Assuming existence is a property of objects, if  $a$  exists is true, then it must be trivially true (that is, it must have the same cognitive value as  $b$  exists where  $a = b$ ), because existence is a condition for referring to an object. The existence condition renders  $a$  exists true, and the direct reference condition renders it indistinguishable from  $b$  exists, where  $a = b$ .<sup>3</sup>

We begin, in Section 7.2, to set out the Frege/Russell story about existence. In Section 7.3, we argue for the view that ‘ $x$  exists’ is a first-order predicate, and in Section 7.4 we use the machinery of Russell’s account of descriptions to show how it enables us to avoid paradoxes about existence even when it is taken as a property of objects. We then return, in Section 7.5, to a close analysis of the text of the dialogue with Pünjer, and reveal the sense/reference confusion in Frege’s argument. After a brief discussion of nonreferring singular terms in Section 7.6, we turn back in Section 7.7, to see how different the Frege/Russell view is from Kant’s.

## 7.2 The Frege/Russell View About ‘Existence’

Frege and Russell, inspired by Kant (1781), held that ‘ $x$  exists’ is not a real predicate. The work we want from it, they held, is adequately and correctly provided by the existential quantifier. Frege (1884*b*) was quite precise about the notion. His view was that it was not a *first-level* predicate, one that expressed a property of objects. He proposed in *Grundlagen* and subsequent that it was actually a *second-level* predicate, one that expressed a property of concepts. The view he held was that ‘ $F$ s exist’ is to be understood as expressing something about the concept  $F$ , namely, that it has at least one instance. Russell (1918: 232) says in similar language: “Existence is essentially a property of a propositional function. It means that that propositional function is true in at least one instance.” So to say that  $F$ s exist is to say that the function  $x$  is  $F$  comes out true for at least one value of  $x$ . The wisdom in these words is more commonly expressed today as

PRINCIPLE 7.2.1 (FREGE/RUSSELL ON ‘EXISTENCE’) *To assert that  $F$ s exist is to say that there are  $F$ s, and to deny that  $F$ s exist is to say that there aren’t any  $F$ s.*

It is important that we distinguish Principle 7.2.1 from the closely related

PRINCIPLE 7.2.2. (FREGE/RUSSELL ON EXISTENCE) (i) 'x exists' is not a first-order predicate; (ii) Existence is not a property of objects but of properties; and (iii) Existence is completely expressed by means of the quantifier 'There is'.

As we will see, Principle 7.2.1 is true (in nonmodal contexts) but Principle 7.2.2 is false.

The quantifier  $(\exists x)$  is read as 'There is'. Frequently it is also read as 'There exists', for it is with the quantifier that we usually express existence claims.<sup>4</sup> It is a widely respected philosophical view that we *always* express existential claims with the quantifier. The Frege/Russell scheme is implemented in first-order logic so:

$$\begin{aligned}Fs \text{ exist} &\leftrightarrow (\exists x)Fx, \\Fs \text{ do not exist} &\leftrightarrow \neg(\exists x)Fx.\end{aligned}$$

In formalized English, 'Unicorns exist' becomes ' $(\exists x)$  ( $x$  is a unicorn)' and 'Unicorns do not exist' becomes ' $\neg(\exists x)(x$  is a unicorn)'. There are complications with the Frege/Russell scheme when applied to singular assertions and denials of existence, for example, when we say 'Homer exists' or 'Homer does not exist'. Frege never adequately addressed these cases in his published work, but in the dialogue with Pünjer, Frege (1884*a*) proposed an analysis similar to his *Begriffsschrift* treatment of identity: 'Homer does not exist', he said, is about the name 'Homer', not the man, and it says of the name that it does not designate.<sup>5</sup> Russell's view was also complicated. As we saw in Section 6.4, Russell maintains that it makes no sense to speak of an object as existing or not, and that sentences that purport to do so are plain nonsense. *Logical* or *genuine* proper names serve simply to introduce their referent into the proposition, and so no meaningful assertions or denials of existence are generated using these types of expressions. But ordinary or garden-variety proper names have a different logical analysis. Russell (1905) offered a suggestion, which Quine (1948) later refined, of associating with a given singular term a predicate that purports to apply truthfully to at most one thing. Russell sought a predicate that would be widely acceptable, like ' $x$  authored the *Iliad*'; Quine, however, simply invents a predicate for the purpose, ' $x$  homerizes'. These garden-variety proper names are then logically regarded as, in effect, predicative constructions. They are *disguised* or *truncated* descriptions. Using Quine's predicate (only because it is briefer), 'Homer

exists' becomes ' $(\exists x)(x \text{ homerizes uniquely})$ ', and 'Homer does not exist' becomes ' $\neg(\exists x)(x \text{ homerizes uniquely})$ '.

### 7.3 Is 'Exists' a Predicate?

The sentence

Something exists (7.1)

is readily understood. As soon as we understand it, we are impressed with its truth. Descartes reminds us that

I exist (7.2)

is true whenever it is affirmed in speech or thought. (7.1) cannot be far behind. Quine (1948) is a bit more generous. He says:

Everything exists. (7.3)

Here is the famous opening passage of "On What There Is":

A curious thing about the ontological problem is its simplicity. It can be put in three Anglo-Saxon monosyllables: 'What is there?' It can be answered, moreover, in a word – 'Everything' – and everyone will accept this answer as true. However, this is merely to say that there is what there is. There remains room for disagreement over cases; and so the issue has stayed alive down the centuries. (Quine 1948: 1)

The indefinite pronoun 'something' in (7.1) is just the first-order existential quantifier  $(\exists x)$ . The indefinite pronoun 'everything' in (7.3) is the first-order universal quantifier  $(\forall x)$ . A well-formed first-order representation of either (7.1) or (7.3) requires that the quantifier be completed by a predicate, which we designate by  $\mathcal{E}x$ , to get:

$(\exists x)\mathcal{E}x$  (7.4)

or

$(\forall x)\mathcal{E}x$  (7.5)

respectively.  $\mathcal{E}x$  might be a primitive predicate, or it might be a defined predicate –  $x = x$  and  $(\exists y)(x = y)$  are favorite choices<sup>6</sup> – but it must be a predicate.

From this vantage point, to adopt the Frege/Russell view is tantamount to adopting the view that ' $x$  exists' is a *universal* predicate, true

of everything. For using our existence predicate  $\mathcal{E}x$ , the following equivalences hold:

PRINCIPLE 7.3.1 (REDUNDANCY THEORY OF EXISTENCE)

$$\begin{aligned}(\exists x)F(x) &\equiv (\exists x)(\mathcal{E}x \wedge Fx) \\ \neg(\exists x)F(x) &\equiv \neg(\exists x)(\mathcal{E}x \wedge Fx).\end{aligned}$$

These formulas constitute the basis of what we call the *Redundancy Theory of Existence*. The theory is not committed to the claim that 'x exists' is not a predicate. It is not even committed to the claim that 'x exists' is an *unnecessary* predicate. It simply embodies the view that everything exists. But to make this latter claim, where the things said to exist are specified no more precisely than as 'everything',<sup>7</sup> we simply have to have a predicate in our language like  $\mathcal{E}x$ . Philosophers have tended to focus on more precise specifications of the things said or denied to exist, and so they have failed to notice the need for such a predicate in these cases. Moore (1936) famously conceded that 'x exists' is a predicate *in grammar*, identifying as the important philosophical issue whether it is a predicate *in logic*. From what we have just seen, 'x exists' is most definitely a predicate *in logic*.

Incidentally, it is important that we admit the error of Principle 7.2.1. We have found that although it is true, it is not the whole of the story. There are existence claims that cannot be treated in this way, namely those in which the thing said to exist is specified indefinitely. We will find in Chapter 8 a Redundancy Theory of Truth. Here too we can eliminate truth so long as we specify the thing said to be true explicitly. If specified indefinitely, it cannot be eliminated. Neither existence nor truth appear to be eliminable in favor of predication in this way.<sup>8</sup>

The problem with this Redundancy Theory of Existence is that it renders unclear how an existential claim '*F*s exist' can be informative (if true). For everything exists. Put negatively, the problem is how one can informatively (if truthfully) claim '*F*s do not exist'. The latter is the form of the problem Quine (1948) focuses on: If everything exists, where is the "room for disagreement over cases?"

#### 7.4 Russell's Machinery

Suppose we regard 'the present King of France' as an individual constant  $k$ .<sup>9</sup> Suppose, further, that we treat the remainder of (6.23) as a predicate 'x exists'. This we take to be the universal predicate  $\mathcal{E}x$ , true of everything.

We have a logical truth

$$(\forall x)\mathcal{E}x. \quad (7.6)$$

By Universal Instantiation,

$$\mathcal{E}k \quad (7.7)$$

is also a logical truth. So, if we represent (6.23) as (7.7), (6.23) turns out to be logically true and its negation turns out to be logically false. This is a formal characterization of the Paradox of Nonbeing.

Russell (1905) rejected this logical representation of (6.23). (6.23) is not, as he called it, a “subject/predicate” proposition. If we had a genuine subject, he said, it would be a logical or genuine proper name. In the case of a genuine proper name, we must be *directly acquainted* with the object named, and so the issue of its existence cannot possibly arise.<sup>10</sup> Both the affirmation and denial of existence in that case are nonsense. Since (6.23) (and its negation) are neither of them nonsense, he concluded, it cannot be that in (6.23) we are speaking about an object with which we are directly acquainted, and it therefore cannot be that we are predicating anything of such an object.

We saw in Section 6.3 that Russell (1905) analyzed the existence claim (6.23) as a conjunction of the two clauses (6.24) and (6.25). (6.24) is the existence clause and (6.25) is the uniqueness clause. Unlike (6.1), there is no third clause, no predication clause, for the obvious reason that Russell did not believe there was any predication in this case. By the same token, Russell proposed the denial of this conjunction,

$$\begin{aligned} &\text{It is not the case that: [there is at least one thing that kings} \\ &\text{France, and there is at most one thing that kings France].} \quad (7.8) \end{aligned}$$

as the representation of (6.22). The conjunction of the two clauses – (6.24) and (6.25) – is not a logical truth. Because the conjunction is not a logical truth, the negation of the conjunction is not a logical falsehood. So Russell claimed to capture informative and nontrivial affirmations and denials of existence.

The analysis Russell provided of (6.23) (and also of (6.22)) assumes two things: first, that ‘the present King of France’ is not a subject, *and* second, that ‘*x* exists’ is not a predicate. Russell (1905) never argued that ‘*x* exists’ is not a predicate, and it is not an essential component of the theory of descriptions, but an additional assumption. For the *pure* theory,



as we shall call it, that is the theory without this assumption, is sufficient to handle the Paradox with the help of the scope distinction.<sup>11</sup>

Suppose that 'x exists' were a genuine predicate so that (6.23) receives the usual tripartite analysis:

There is at least one thing that presently kings France,  
and there is at most one thing that presently kings  
France, and that thing exists. (7.9)

There are now, as usual, two distinct places to insert the 'not', and so there are two genuine options for interpretation. Using predicate abstract notation, we distinguish the small-scope reading

$$\neg \langle \lambda x. \mathcal{E} x \rangle (k) \tag{7.10}$$

from the large-scope reading

$$\langle \lambda x. \neg \mathcal{E} x \rangle (k). \tag{7.11}$$

The large-scope reading (7.11) is the *de re* reading: *One and only one thing kings France and that thing does not exist*. On the classical first-order interpretation, this *de re* reading is false. For if the thing does not exist, it has no properties. So, in particular, it does not have the *nonexistence* property. This is the problematic reading, the one that self-destructs. But we are saved by the *de dicto* reading, for this makes denials of existence possible. The *de dicto* reading is the small-scope reading (7.10). On the classical first-order interpretation, this *de dicto* reading is true, for it is simply not the case that one and only one thing kings France and has the *existence* property.

Of course, on the *de dicto* reading, 'x exists' is a predicate. It is just not predicated of anything. Let us be clear. The predicate 'x exists' is no different in this regard from any other garden variety predicate, for example, 'x is bald'. For, on the theory of descriptions, when we say that the present King of France is not bald, and we speak truthfully, the claim must be understood *de dicto*. 'x is bald' is a predicate though it is certainly not predicated of anything in this case for, as we all know, the present King of France does not exist.

We have seen that Russell's reasons for denying that *existence* is a property of objects are faulty, and so his solution to the Paradox of Nonbeing ultimately depends on the fact that denials of existence are only permitted the small-scope reading. The scope distinction permits us to identify something as a predicate without its being predicated in that context.

Russell's Theory of Descriptions, far from showing or even depending upon the view that existence is not a property, actually shows us how logically to understand it so that it is. The Theory of Descriptions provides just the machinery to explain informativeness and complete the account of the first-order existence predicate  $\mathcal{E}x$ . Russell had supposed that it made no sense to affirm or deny existence when he had direct acquaintance with an object. But his problem was a result of the peculiarities of *direct acquaintance*, not with its being an *object*. For once the direct acquaintance has been stripped away from the object, the problem of informativeness is eliminated.

Incidentally, Russell (1905) did not supply an argument to think that *existence* is not a first-order property; as we have seen, he assumed it.<sup>12</sup> Gilbert Ryle (1932: 42), in the very influential "Systematically Misleading Expressions," reveals how deeply the idea that *existence* is not a property is woven into Russell's theory. He describes Russell's discovery that descriptions are not subjects as *dependent* upon the observation that existence is not a property.<sup>13</sup>

Since Kant, we have, most of us, paid lip service to the doctrine that 'existence is not a quality' and so we have rejected the pseudo-implication of the ontological argument: 'God is perfect, being perfect entails being existent,  $\therefore$  God exists.' For if existence is not a quality, it is not the sort of thing that can be entailed by a quality.

But until fairly recently it was not noticed that if in 'God exists' 'exists' is not a predicate (save in grammar), then in the same statement 'God' cannot be (save in grammar) the subject of predication.

We have serious reservations about Ryle's characterization of Kant's view, but it is Russell's view we are interested in pursuing further. Here is an extended passage in which Russell (1919: 164–5) informally explains his position about *existence*:

We say that an argument  $a$  "satisfies" a function  $\phi x$  if  $\phi a$  is true; this is the same sense in which the roots of an equation are said to satisfy the equation. Now if  $\phi x$  is sometimes true, we may say there are  $x$ 's for which it is true, or we may say "arguments satisfying  $\phi x$  exist." This is the fundamental meaning of the word "existence." Other meanings are either derived from this, or embody mere confusion of thought. We may correctly say "men exist," meaning that " $x$  is a man" is sometimes true. But if we make a pseudo-syllogism: "Men exist, Socrates is a man, therefore Socrates exists," we are talking nonsense, since "Socrates" is not, like "men," merely an undetermined argument to a given propositional function. The fallacy is closely analogous to that of the argument: "Men are numerous, Socrates is a man, therefore Socrates is numerous." In this case it is obvious that the conclusion is nonsensical, but in the case of existence it is not obvious. . . . For

the present let us merely note the fact that, though it is correct to say “men exist,” it is incorrect, or rather meaningless, to ascribe existence to a given particular  $x$  who happens to be a man. Generally, “terms satisfying  $\phi x$  exist” means “ $\phi x$  is sometimes true”; but “ $a$  exists” (where  $a$  is a term satisfying  $\phi x$ ) is a mere noise or shape, devoid of significance. It will be found that by bearing in mind this simple fallacy we can solve many ancient philosophical puzzles concerning the meaning of existence.

The passage very clearly embodies a confusion between the way in which we specify an object and the object itself. For a *term satisfying  $\phi x$*  has got to be an object, and if it makes sense to say of a term satisfying  $\phi x$  that it exists, then it makes sense to say of an object that it exists. The contrast is between this description and the name  $a$ , where this name  $a$  is a logical constant. Since it is a logical constant, ‘ $a$  exists’ is treated as a subject/predicate proposition, with all the problems already noted. The argument Russell points to in the above passage is one that trades on this confusion.

Here is another argument Russell (1918: 234) provides for this view. It too is flawed.

[Y]ou sometimes know the truth of an existence-proposition without knowing any instance of it. You know that there are people in Timbuctoo, but I doubt if any of you could give me an instance of one. Therefore you clearly can know existence-propositions without knowing any individual that makes them true. Existence-propositions do not say anything about the actual individual but only about the class or function.

To be sure, we might know that people in Timbuctoo exist without knowing any particular such exister, but by the same token we might know that people in Timbuctoo walk, without knowing any particular such walker. This provides scant evidence that walking-propositions say nothing about actual individuals.

But what of Russell’s view, which, like Frege’s, takes existence to be a property of properties, not a property of objects? On this view, ‘ $x$  does not exist’ is like ‘ $x$  is extinct’, which applies to a species but does not distribute over the individual members of the species. We say, for example, ‘The dodo is extinct’, but we by no means imply ‘This dodo is extinct’, identifying a particular dodo. Russell’s analogy is between ‘Men exist’ and ‘Men are numerous’: in neither case, he says, does the property apply to individual men. If we set aside arguments he provides for this view – and these arguments, as we have noted are terribly flawed – then we only have the linguistic evidence to go on. And Russell’s story is just not compelling.

There is no fallacy in the argument

Men exist  
Socrates is a man  
Therefore, Socrates exists.

Each of these statements seems to make perfectly good sense. And assuming the first premise to mean ‘*all* men exist’, then we cannot possibly have true premises and a false conclusion. The contrast with the argument

Men are numerous  
Socrates is a man  
Therefore, Socrates is numerous

is striking. The two premises might very well be true, but the conclusion – insofar as it makes sense – certainly is not. We cannot distribute *numerosity* over individual men, but *existence* is quite different. For how can men exist if no individual man does?

### 7.5 Frege’s Mistake

Frege, as we have said before, treated *existence* as a property. Not a property of objects – a first-level property, as he called it – but rather a property of properties – that is, a second-level property. He identified it with the property designated by the existential quantifier ‘There is’. To say ‘*F*s exist’ is to say ‘There are *F*s’, and to say that there are *F*s is to speak about the concept *F*. It is to say about this concept that it has at least one instance. Here is the famous passage from *Grundlagen*:

By properties which are asserted of a concept I naturally do not mean the characteristics which make up the concept. These latter are properties of the things which fall under the concept, not of the concept. Thus “rectangular” is not a property of the concept “rectangular triangle”; but the proposition that there exists no rectangular equilateral rectilinear triangle does state a property of the concept “rectangular equilateral rectilinear triangle”; it assigns to it the number nought.

In this respect existence is analogous to number. Affirmation of existence is in fact nothing but denial of the number nought. Because existence is a property of concepts the ontological argument for the existence of God breaks down. (Frege 1884*b*: 64–5)

How did the Ontological Argument come in here? That is just Frege’s nod to Kant, whose comments about existence are framed entirely within the

context of this argument. We should not underestimate, however, the importance of this connection for Frege. Some two years earlier, in a letter dated August 29, 1882, to Anton Marty – like Meinong, a student of Brentano's – he explains some of the notation of his *Begriffsschrift*. He shows how, as he puts it, “every particular judgment is an existential judgment.” His example is

$$\vdash \exists x \top a^2 = 4$$

that is, “There is at least one square root of 4.” He continues:

Existential judgments thus take their place among other judgments. I should still like to show you how Kant's refutation of the ontological argument becomes intuitively very obvious when presented in my way and what the value of the concavity is, which is my sign of generality, but I fear have already overburdened you with my long letter. (Frege 1979: 102)

The very same connection was reaffirmed years later in a passage we quoted earlier: “The ontological proof of God's existence suffers from the fallacy of treating existence as a first-level concept” (Frege 1891: 146). It is of no small importance, then, to Frege, that his treatment of existence should coincide with Kant's and serve to thwart the Ontological Argument. But the promise to Marty of the spelling out of this connection is one of the great teasers of modern philosophy.

In *Grundlagen*, however, the Ontological Argument is a side issue. Frege's primary concern is with the definition of *cardinal number*, and so the crucial connection for him is between the notion of *existence* and the notion of *cardinal number*. A statement like ‘There is at least one *F*’ is on a par with a statement like ‘There are four *F*s’. These are “statements of number,” statements that answer the question ‘How many?’ Once he had worked out the quantifiers – all *F*s, some *F*s, no *F*s – there appeared to be no difference in kind to precise the quantity – one *F*, two *F*s, three *F*s, and so on – and so to regard arithmetic as on a par with logic. No *F*s in fact provided the precision – 0 *F*s – on which his ingenious construction was based.<sup>14</sup> “The content of a statement of number,” Frege said, “is an assertion about a concept.” This is because, if we might so phrase it, the content of an existential statement is an assertion about a concept.

[T]he content of a statement of number is an assertion about a concept. This is perhaps clearest with the number 0. If I say “Venus has 0 moons”, there simply does not exist any moon or agglomeration of moons for anything to be asserted of; but what happens is that a property is assigned to the *concept* “moon of Venus”, namely, that of including nothing under it. If I say “the King's carriage is drawn

by four horses”, then I assign the number four to the concept “horse that draws the King’s carriage”. (Frege 1884*b*: 59)

The interpretation of *existence* and the connection between *existence* and *number* is repeated eight years later, when the sense/reference distinction had been clearly drawn:

I have said that to assign a number involves saying something about a concept; I speak of properties ascribed to a concept, and I allow that a concept may fall under a higher one. I have called existence a property of a concept. How I mean this to be taken is best made clear by an example. In the sentence ‘There is at least one square root of 4’, we are saying something, not about (say) the definite number 2, nor about  $-2$ , but about a concept, *square root of 4*; viz. that it is not empty. (Frege 1892*b*: 187–8),

Now there is an important and contentious issue in here that has occupied a great deal of philosophical space. Since a statement of number is a statement about a concept, Frege held that every count of objects had to be carried out within the framework of a concept that divided its subject matter. How many things one had depended upon how one sorted them. In the case of a deck of cards, we have *one* deck, *four* suits, *thirteen* playing cards.

If I give someone a stone with the words: Find the weight of this, I have given him precisely the object he is to investigate. But if I place a pile of playing cards in his hands with the words: Find the Number of these, this does not tell him whether I wish to know the number of cards, or of complete packs of cards, or even say of honour cards at skat. To have given him the pile in his hands is not yet to have given him completely the object he is to investigate; I must add some further word – cards, or packs, or honours. (Frege 1884*b*: 28–9)

Peter Geach (1967) claimed that Frege should have treated identity in the same way he handled cardinality. To say that *a* is *identical with b*, he argued, is just to say that *a* is *one and the same as b*, so if ‘one’ had to be completed by a concept word, ‘same’ would have to as well. Frege’s view, as we see, is that we might have one *F* but two *G*s. Parallel to this, Geach’s doctrine of *relative identity* is that *a* and *b* might be the same *F* but different *G*s. Many in the ensuing discussion thought Frege was right to treat these differently, and that although counting was relative, identity was absolute.<sup>15</sup> However, we tend to agree with Bennett and Alston (1984), who argued that Geach was correct in holding Frege to the same position about identity that he maintained about number, but that Geach was wrong about which position to hold: both should be absolute, that is, about objects.

Our contribution is to add existence into the mix. If a statement of identity is about objects, then a statement of existence should also be about objects.<sup>16</sup> We are at a loss to understand why Frege handled the notions of identity and existence so differently. But we are certain, and will argue it now, that he was wrong to do so.

We are hampered in our investigation because, as we mentioned at the beginning of this chapter, Frege provided no argument in his published writing for the view about existence that he is famous for. In fact, he has nothing much more to say than the sort of pronouncements of which the above quotations are a good example. We see that he held the view at least as early as 1882 and never changed his mind about it, despite other radical changes in his philosophical vision. But why did he hold it? If Kant (1781) had held the same view, we might take Frege's approving nod in his direction as license to accept wholesale Kant's arguments for the position he and Frege shared. But, as we will see in Section 7.7, Kant's position was importantly different from Frege's. The only relevant text we can locate is the transcript in his *Nachlass* of a dialogue with his colleague, the theologian Bernard Pünjer, on the topic of existence, together with Frege's commentary on the issues. The editors date the transcript "before 1884," Pünjer's death in 1885 marking, as they put it, "an upper limit" on its occurrence. This clearly places it sometime between the publication of *Begriffsschrift* in 1879 and the publication of *Grundlagen* in 1884. This dating is extremely important. It is a generally accepted fact among Frege's commentators that *Grundlagen* antedates Frege's sense/reference semantics. Where we differ from the generally accepted view is in our identification of a *Begriffsschrift* semantic story, one that has not been widely recognized, and which we have every reason to believe was the operative story in *Grundlagen*. Our disagreement with Dummett's (1991: 66–7) assessment of the situation should be evident to the reader:

The brilliance and clarity of *Grundlagen*, and the cogency of many of its arguments, make it difficult for us to take in the fact of Frege's blindness, during the whole of his early period, to what seems to us an obvious need for a distinction. He simply had no consciousness, until he formulated the principles of his middle-period theory, of the necessity for distinguishing between the significance of an expression and that which it signifies. . . . The content or meaning (*Bedeutung*) of an expression was for Frege at that time simultaneously its significance and *what* it signified: the distinction became apparent to him only when he drew his distinction between *Sinn* and *Bedeutung*, and he was strictly accurate in saying that he had split the former notion of content into those two components. This explains the oddity of his later terminology: he chose to retain the term '*Bedeutung*' for

that which the expression signifies. It explains also why the term ‘concept’ plays so striking a double role in *Grundlagen*, being used sometimes for the sense of a predicative expression and sometimes for its reference. Naturally, no coherent exposition can be given of the doctrines of *Grundlagen* without acknowledging a distinction between significance and what is signified; but, in reading the book, we must bear in mind the fact that Frege was not himself making such a distinction. His failure to do so means that there was at that time a radical incoherence at the very heart of his thinking, though one that obtrudes very little in the argumentation of the book.

Dummett’s claim that Frege had “no consciousness” of a distinction between significance and signification – Dummett’s terms in this passage for sense and reference – in his early work, is unsupportable. We have documented a very different view in Chapter 4. This is not to say that Frege had worked out sense/reference semantics in *Grundlagen*. He did not. Nowhere is this more obvious than in his treatment of predicate expressions; for Frege had not clarified for himself how the sense/reference distinction was to apply in the case of predicate-expressions, any more than he had clarified how the sense/reference distinction was to apply to sentences. We disagree, then, with Dummett’s more sweeping claim about sense and reference; but the “radical incoherence” in the semantic/ontological underpinnings of *Grundlagen* is precisely what we have been pointing out here, and we probably find the incoherence a bit more obtrusive than Dummett did.

Without the worked-out sense/reference semantics, it is most reasonable to understand the arguments in the dialogue with Pünjer within his earlier semantic framework from *Begriffsschrift*. This is what we will now argue. The arguments are persuasive only within the context of the earlier semantics, and they have no plausibility within the later sense/reference semantics. For the sense/reference semantics provides a framework, much like Russell’s, in which the view that existence is a property of objects has a comfortable home. Just as he had abandoned his early view about identity, Frege should have abandoned the correlative view about existence.

We spoke about the *Begriffsschrift* semantic theory back in Chapter 4. The notion of content [*Inhalt*] was, as Frege later said, a hybrid: part reference and part sense. The *Inhalt* of an atomic statement was that of a Russellian singular proposition. In an atomic sentence  $S\alpha$ ,  $\alpha$  stands for an object and  $S$  stands for a concept – but intensionally understood, that is, as a property or attribute. The sentence as a whole stands for its *Inhalt*, which consists of an object and a property. The *Begriffsschrift*



Substitution Principle 3.3.1 that generated the Evening Star/Morning Star paradox was a clear reflection of this direct reference semantics: substitution of coreferential singular terms preserved cognitive value. Frege's two-layer semantics of sense and reference resulted from splitting up his *Begriffsschrift* notion of *Inhalt*. The reference of  $S$  was an *extensionally* understood concept, which when applied to the reference of  $\alpha$  – an object – yielded a truth value (the reference of the sentence). The sense of  $S$  was the *intensionally* understood concept of *Begriffsschrift*, which when applied to the sense of the name  $\alpha$  yielded a thought or proposition.

Given that  $a = a$  and true  $a = b$  differ in cognitive value, the author of *Begriffsschrift* held that these identity statements were about the names of the objects, not the objects themselves. The only way he could handle informativeness in *Begriffsschrift* was by shifting the subject of discourse. Instead of reading  $a = b$  to say that the objects referred to were identical, he took it to say that the two names were coreferential.  $a = a$  and  $a = b$  now differed in *Inhalt*, because the *Inhalt* of the former contained the name  $a$  where the *Inhalt* of the latter contained the name  $b$ .

Our suggestion is that the young Frege adopted the very same strategy to deal with the informativeness of singular existence statements that he had adopted to deal with the informativeness of identity statements. We turn now to the dialogue.

The focus of the dialogue is Pünjer's opening question: Does the sentence

Something does not have the characteristic of flying, but does fall under the concept "bird"

have the same meaning as the sentence

Among what *is*, is something that does not have the characteristic of flying, but does fall under the concept "bird"?

Frege's (1884*a*) answer is "Yes," but that nothing could have been added by the clause "among what *is*." In his commentary on the dialogue, the key point Frege insists on is that to say that something is  $F$  and exists is to say no more than that it *is*  $F$ .

Now I can readily grant that the expression 'there are men' means the same as 'Something existing is a man' only, however, on condition that 'exists' predicates something self-evident, so that it really has no content. The same goes for other expressions which you use in place of 'exist'. (Frege 1884*a*: 62)

What he has in mind as a replacement for “exists” is “self-identical”:

I shall use the fact: that instead of ‘exists’ one can also say ‘is identical with itself’ to show that the content of what is predicated does not lie in the word ‘exists’. ‘There are men’ means the same as ‘Some men are identical with themselves’ or ‘Something identical with itself is a man’. (Frege 1884*a*: 62)

The ascription of existence “has no content,” not in the sense of being meaningless, but rather in the sense of not adding anything new, of not being informative.

Neither in ‘A is identical with itself’ nor in ‘A exists’ does one learn anything new about A. Neither statement can be denied. (Frege 1884*a*: 62)

Again, he says,

[T]he judgments ‘This table exists’ and ‘This table is identical with itself’ are completely self-evident, and that consequently in these judgments no real content is being predicated of this table. (Frege 1884*a*: 62–3)

These comments are perfectly in line with our remarks in Section 7.3 about the redundancy of a first-order existence predicate. Singular existence claims are uninformative and so lack “real” content, so the sense of ‘exists’ that is consistent with the use of ‘There is’ to express existence must be this redundant, uninformative, no-content predication.

But existence claims are, as we know, informative:

But if the proposition ‘Leo Sachse is’ is self-evident then the ‘is’ cannot have the same content as the ‘there are’ of ‘There are men’, for the latter does not say something self-evident. Now if you express what is said by ‘There are men’ by ‘Men exist’ or ‘Among that which has being is some man’, then the content of the statement cannot lie in the ‘exist’ or ‘has being’ etc. (Frege 1884*a*: 62)

This raises a question:

But if the content of what is predicated in the judgment ‘Men exist’ does not lie in the ‘exist’, where then does it lie? I answer: in the form of the particular judgment. Every particular judgment is an existential judgment that can be converted into the ‘there is’ form. E.g. ‘Some bodies are light’ is the same as ‘There are light bodies’. ‘Some birds cannot fly’ is the same as ‘There are birds that cannot fly’, and so on. (Frege 1884*a*: 63)

So, if we take existence to be a property of objects, the claim that this particular object exists is uninformative. For this reason, although ‘There are men’ is equivalent to ‘Men exist’, the *exists* in this later sentence cannot be the uninformative *exists* noted earlier. So what is being predicated is not *of* an object. The particular judgment, for Frege, is a judgment that

relates concepts. An existence claim is a judgment of this kind: 'There is ( )' and '( ) exists' are each completed by a concept word, and are each about a concept.<sup>17</sup>

As we are by now well aware, Frege distinguished sharply between singular terms and general terms, between names of objects and names of concepts. Frege's story about existence, which remained unchanged through very radical changes in his semantic theory, is fundamentally unstable because of the difference in the way in which it treats singular existence claims and general existence claims. The explanations given in the previous paragraphs all deal with concepts. But what about singular existence claims? 'Frege' is an *Eigenname*. It names an object, a particular man. 'The author of *Grundlagen*' is also an *Eigenname*: it denotes the same object as 'Frege'. On Frege's view, the two sentences

Frege exists (7.12)

The author of *Grundlagen* exists (7.13)

cannot therefore assert anything about a concept. Frege never changed the logical role of descriptions in the way Russell did. Russell treated them as predicative constructions, so that (7.12) and (7.13) could be informative. To be sure, Frege later ascribed a predicative characteristic to these names by ascribing to them a *Sinn*, but this semantic characterization had no direct embodiment – as it did on Russell's theory – in syntax.<sup>18</sup>

These singular assertions and denials of existence are informative even though no concepts appear to be involved. Frege's strategy – just like the strategy of *Begriffsschrift* in dealing with identity – was to take the claims to be about the names, not the things named. Look at how he dealt with proper names in the dialogue:

If 'Sachse exists' is supposed to mean 'The word 'Sachse' is not an empty sound, but designates something', then it is true that the condition 'Sachse exists' must be satisfied. But this is not a new premise, but the presupposition of all our words – a presupposition which goes without saying. The rules of logic always presuppose that the words we use are not empty, that our sentences express judgments, that one is not playing a mere game with words. Once 'Sachse is a man' expresses an actual judgment, the word 'Sachse' must designate something, and in that case, I do not need a further premise in order to infer 'There are men' from it. The premise 'Sachse exists' is redundant, if it is to mean something different from the above-mentioned presupposition of all our thinking. (Frege 1884a: 60)

Here the direct reference assumption is explicit: if the name fails to refer to anything, it is empty, contentless. To preserve content, Frege changed

the subject matter to speak about the name. ‘Sachse exists’ is taken to mean “‘Sachse’ designates’.

Frege offers no unified treatment of singular existence claims and general existence claims. It is quite clear that this is an immediate consequence of the fact that he had no unified treatment of singular terms and predicate-expressions in his *Begriffsschrift* semantics: predicate-expressions stood for senses but singular terms stood for references. The informative general existence claim is about the intensional “concept” the predicate stands for. The informative singular existence claim is about the name of the object the singular term stands for – no doubt to capture the associated sense just as with the *Begriffsschrift* solution for identity. If we are correct about the *Begriffsschrift* semantic theory, his concepts were sense items,<sup>19</sup> and so his idea that the informative general existence claim was about a concept really was the story he should have incorporated as part of his sense/reference theory, amended, of course, because the sense is expressed, not referred to, and so it is not what the statement is about (any more than an identity statement is about the senses of the expressions flanking the identity sign).

In summary, then, Frege’s explanation of the informativeness of existence claims only works within this early direct reference semantics; within the sense/reference semantics, it is disjointed and implausible. Once he had adopted the sense/reference theory, he should have repudiated the view about existence just as he repudiated the view about identity. Within the sense/reference framework, the appropriate view about *existence* takes it to be a property of objects. It is a universal property. So, to say *of* something that it exists is not to say very much; to deny *of* something that it exists is – if reference presupposes existence – inconsistent. But to say *that* a thing exists can, indeed, be interesting; so too to deny *that* a thing exists. What is informative in each case is the thought expressed. The story is entirely analogous to the one we argued for in the previous section, but using the machinery Russell provided.

It remains something of a mystery why Frege never changed his view of existence, although we suspect that it was too deeply developed in conjunction with the definition of number for him to give it up.

## 7.6 Nonreferring Singular Terms

Frege’s (1892c: 157) correction of this early view of nonreferring singular terms is striking, but incomplete:

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 a *Bedeutung*, it is also doubtful whether the whole sentence does. Yet it is certain, nevertheless, that anyone who seriously took the sentence to be true or false would ascribe to the name ‘Odysseus’ a *Bedeutung*, not merely a sense; for it is of the *Bedeutung* of the name that the predicate is affirmed or denied. Whoever does not admit the name has a *Bedeutung* can neither apply nor withhold the predicate. But in that case it would be superfluous to advance to the *Bedeutung* of the name; one could be satisfied with the sense, if one wanted to go no further than the thought. If it were a question only of the sense of the sentence, the thought, it would be needless to bother with the *Bedeutung* of a part of the sentence; only the sense, not the *Bedeutung*, of the part is relevant to the sense of the whole sentence. The thought remains the same whether ‘Odysseus’ has a *Bedeutung* or not.

No longer do we hear of a name being an “empty noise” if it fails to denote. ‘Odysseus’ is meaningful even if it lacks a reference. This entails that the sentence

Odysseus exists (7.14)

also has a sense. (7.14) expresses a thought. In fact, it appears to express a quite different thought from

The cleverest of the Achaeans exists. (7.15)

For an individual might believe (7.14) to be true and yet also believe (7.15) to be false. The existence claim does have content. But, although Frege thought that singular existence claims made sense and were informative as well, he chose to regard them still as about the names, not what the names customarily stood for. He took (7.14) to express

‘Odysseus’ has a reference. (7.16)

But then the cognitive value of (7.16) would not be significantly different from

‘The cleverest of the Achaeans’ has a reference. (7.17)

Each would be about the words. The view is subject to the same sort of criticism Frege (1892c) had leveled against his *Begriffsschrift*: we do not appear to be expressing “proper knowledge” speaking thus about the words rather than what they designate. It is rather remarkable that he did not see this.

From the perspective of the sense/reference theory, the informativeness of existence statements is an issue to be dealt with at the level of sense, not at the level of reference. With the sense/reference distinction at hand, Frege should have located the informativeness of (7.14) in the

sense of 'Odysseus'. This would have enabled him to draw on the predicative character of *Sinne* and unify his treatment much as Russell was able to. At the level of reference, the sharp distinction between *Eigennamen* and *Begriffswörter* would be honored, but at the level of sense, the difference would be bridged by these predicative *Sinne*. For the singular term, the thought that *Odysseus exists* would be that a certain way of determining things has *one instance*, and for the general term, the thought that *Men exist* would be that a certain way of determining things has *at least one instance*. Using the term 'instances' would not bring objects into the picture because, as Frege said, the thought remains the same whether or not it in fact has instances. For he really believed that it was irrelevant to the thought itself whether or not a singular term denoted an object, and whether a general term applied truthfully to anything. The thought that a concept had instances remained whether it had instances or not, and so it was not about the instances. In a late, unpublished passage, Frege (1906: 191) explicitly embraces this attitude:

People certainly say that Odysseus is not an historical person, and mean by this contradictory expression that the name 'Odysseus' designates nothing, has no meaning. But if we accept this, we do not on that account deny a thought-content to all 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 to be 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. But we can immediately infer from what we have just said that something further must be associated with the proper name, something which is different from the object designated and which is essential to the thought of the sentence in which the proper name occurs. I call it the sense of the proper name. As the proper name is part of the sentence, so its sense is part of the thought.

But this is not the way in which he used the term 'concept' in his sense/reference years. A concept is the reference of a *Begriffswort*, and it is understood extensionally. The reasoning about intensional concepts that we discussed in the last paragraphs just is not compelling when we move to the level of reference. If 'Winged horses exist' means that the concept *Winged horse* has instances, then it must mean that the concept has at least one instance, and so there is a thing (an object) that is an

instance of the concept. How can the claim be about the concept and not the things that fall under it? And the denial ‘Winged horses do not exist’, which means that the concept has no instances, must mean that there is no thing that is an instance of the concept, which as much as says that each thing is not an instance of the concept. Once again, it is difficult to see how the claim is about the concept and not the object.

Lastly, we note that Frege appears to have breached his Principle 3.6.1, that sense determines reference. If an expression can have a sense and no reference, or even have a reference but not be committed to it in the way envisaged, then Frege has modified his basic theory in some very crucial way – indeed, a way that we are unable to assess. It is very clear that Frege’s commitment to Principle 3.6.1 is shaky. This is yet additional reason to reject Evans’s (1982) controversial claim that Frege was committed to “object dependent thoughts” much like Russell’s singular propositions: not only does it fly in the face of explicit statements of Frege’s, not only does it ignore the shift from the *Begriffsschrift* to the sense/reference semantics, but, finally, it relies on an unshakeable commitment to Principle 3.6.1, a stance, we see, that Frege did not adopt. Certainly Frege’s story about truth-value gaps itself has gaps. There is a genuine puzzle about what truth value – if any – Frege should ascribe to (7.14). But that is because he was unclear about the issue, not, as Evans (1982) said, because Frege held that a sentence like ‘Odysseus was set ashore at Ithaca while sound asleep’ could not *really* be used to express a thought.

## 7.7 Kant on Being

Here is the widely quoted paragraph from Kant’s (1781: 504–5) famous discussion of existence:

‘*Being*’ is obviously not a real predicate; that is, it is not a concept of something which could be added to the concept of a thing. It is merely the positing of a thing, or of certain determinations, as existing in themselves. Logically, it is merely the copula of a judgment. The proposition, ‘God is omnipotent’, contains two concepts each of which has its object – God and omnipotence. The small word ‘is’ adds no new predicate, but only serves to posit the predicate *in its relation* to the subject. If, now, we take the subject (God) with all its predicates (among which is omnipotence), and say ‘God is’, or ‘There is a God’, we attach no new predicate to the concept of God, but only posit the subject in itself with all its predicates, and indeed posit it as being an object that stands in relation to my concept. The content of both must be one and the same; nothing can have been added to the concept, which expresses merely what is possible, by my thinking its object (through the expression ‘it is’) as given absolutely. Otherwise stated,

the real contains no more than the merely possible. A hundred real thalers do not contain the least coin more than a hundred possible thalers. For as the latter signify the concept, and the former the object and the positing of the object, should the former contain more than the latter, my concept would not, in that case, express the whole object, and would not therefore be an adequate concept of it. My financial position is, however, affected very differently by a hundred real thalers than it is by the mere concept of them (that is, of their possibility). For the object, as it actually exists, is not analytically contained in my concept, but is added to my concept (which is a determination of my state) synthetically; and yet the conceived hundred thalers are not themselves in the least increased through thus acquiring existence outside my concept.

By whatever and by however many predicates we may think a thing – even if we completely determine it – we do not make the least addition to the thing when we further declare that this thing *is*. Otherwise, it would be exactly the same thing that exists, but something more than we had thought in the concept, and we could not, therefore, say that the exact object of my concept exists.

Engel (1963) says that the first paragraph “contains nearly everything Kant has to say” on existence as it connects with the Ontological Argument. We have included the beginning of the subsequent paragraph because it exhibits simply the controversial claim that is otherwise lost in the verbiage of the first paragraph.

It is a great misrepresentation, as Engel (1963) also notes, to attribute to Kant the well-known view that *existence is not a predicate*. We have already seen Ryle’s claim that Kant denied existence was a quality. This is just not so. Kant claimed something very different, namely that existence was not a *real* property. He gave a very precise sense to this locution. It is not, he said, a *predicate-that-enlarges-the-concept-of-a-thing*. Here is one way to understand this. Where  $F$  is the *nonmodal* property a thing must possess to belong to a certain set, then if we append  $\mathcal{E}$  – *existence* – we have not thereby altered the constituency of the set. That is,  $F \cap \mathcal{E} = F$ , for any  $F$ ; which is just to say  $F \subseteq \mathcal{E}$ , for any  $F$ , making  $\mathcal{E}$ , in effect, the universal set. Whatever one imagines, one imagines to exist: when one imagines one hundred thalers, one imagines they have the *existence* property (even though they might not). This is pretty much a restatement of the Redundancy Theory of Existence 7.3.1.

To the extent that Frege and Russell espouse this Redundancy Theory of Existence, it is fair to say they are following Kant rather closely. But in denying that it is a first-order property, they are going far beyond anything Kant says and perhaps actually conflicting with his view. Frege (1884 *a*), we recall, drew the analogy between *existence* and *being self-identical*. Everything is self-identical; and so, by the same token, everything exists.



This is just Principle 7.3.1, and the similarity with Kant is clear. But then Frege veered off in a different direction, saying that this does not really capture what we mean when we say that something exists because it does not capture the informative quality of existence claims. He therefore denied that it is a first-order predicate, apparently holding a rather different view from the one Kant espoused! Kant was not a direct reference theorist and he was not worried about the informativeness of existence claims. He never denied, as Frege did, that *existence* was really a first-order property.

Frege clearly *thought* he had captured Kant's view, and so did many who followed. Jonathan Bennett (1974: 231), for example, in his otherwise fine study *Kant's Dialectic*, says: "According to Kant, every existence-statement says about a concept that it is instantiated, rather than saying about an object that it exists."<sup>20</sup> But you cannot say what you say cannot be said. If the first part of the sentence is true, then the second part of the sentence is either meaningless (because it purports to be about an object) or, alternatively, is to be reinterpreted as being about a concept, undercutting the juxtaposition of truth with falsity implied by the transition *rather than*. Presumably this is an infelicitous wording: Bennett is very clearly trying succinctly to characterize Frege's view about existence. But Frege's view cannot be attributed to Kant. There are many reasons. First, it is an anachronism. Kant did not have the Fregean tools of art, the concept/object distinction, to make the claim. His use of the term 'concept' is quite different from Frege's. Second, Kant says nothing remotely like it. In fact, the words in the text actually favor the opposite position. We quote again:

The small word 'is' adds no new predicate, but only serves to posit the predicate in its relation to the subject. If, now, we take the subject (God) with all its predicates (among which is omnipotence), and say 'God is', or 'There is a God', we attach no new predicate to the concept of God, but only posit the subject in itself with all its predicates, and indeed posit it as being an object that stands in relation to my concept.

This certainly makes it look as though a statement like *God exists* is about the object itself, that Divine Being, *Him*, in all his glory. Third, Kant was dealing with the notion of existence within a modal framework; this is an entirely different playing field from the one Frege was on. This makes even the redundancy that Kant speaks about suspect when interpreted Frege's way. Fourth, Kant did not share Frege's direct-reference semantics that drove him to say that existence was a second-order, not a first-order, property. Finally, it is worth remarking that if Kant had held the Fregean

view Bennett (1974) attributes to him, then it is incredible that he said nothing about the Cartesian clear and distinct perception *I exist*, which most emphatically is about an object.

Is there anything in our discussion that leads one to think Frege's attribution of existential import to the quantifier, as opposed to any first-order predicate, correctly captures Kant's view? The only piece of the passage unaccounted for is Kant's saying that we *posit* the object. But positing is not connected in any clear way with a particular linguistic or mathematical form. Frege's suggestion should be abandoned completely.

## Thought, Truth Value, and Assertion

### 8.1 Introduction

We saw in Chapter 4 that Frege had applied the *Begriffsschrift* surrogate for identity, *identity of content*, to sentences as well as to names. Frege (1879) believed that names and sentences both stood for their contents. Frege (1892*c*) saw no need to change his treatment of sentences as names. The issue he addressed was not *whether* sentences refer, but *what* they refer to. He sought to correct his early account – as well as related views which take propositions, thoughts, states of affairs, or facts as the items designated by sentences. These items, he now thought, belonged at the level of sense.<sup>1</sup> In a very influential argument, Frege defended the view that the two truth values – *true* and *false*, or as Frege preferred, *the True* and *the False* – are the *only* candidates that are functionally related via the Compositionality Principle 2.3.1 to the reference of the parts of the sentence, and which, in turn, are functionally related to the reference of larger constructions in which the sentences are embedded. “If we are dealing with sentences for which the *Bedeutung* of their component parts is relevant,” Frege (1892*c*: 158–9) asked, “then what feature except the truth value can be found that belongs to such sentences quite generally and remains unchanged by substitutions of the kind just mentioned?”<sup>2</sup> In this chapter, we will examine carefully his views on truth, and, in particular, raise doubts about the inevitability of this result.

### 8.2 The Frege Argument

Alonzo Church (1956: 25) presents an elegant version of the argument Frege (1892*c*) advances for his view. He invites us to consider the true

sentence

Sir Walter Scott is the author of *Waverley*. (8.1)

‘The author of *Waverley*’ refers to the same person as ‘the man who wrote twenty-nine *Waverley* novels altogether’. Substituting one term for the other in (8.1), we obtain the true sentence

Sir Walter Scott is the man who wrote twenty-nine  
*Waverley* novels altogether, (8.2)

which, according to the Extensionality Principle for Reference 2.3.3, must have the same reference as (8.1). Church then paraphrases (8.2) as

The number, such that Sir Walter Scott is the man who  
wrote that many *Waverley* novels altogether, is twenty-nine; (8.3)

and this, as he remarks, if not synonymous with (8.2), is “at least so nearly so as to ensure its having the same denotation.” But ‘the number, such that Sir Walter Scott is the man who wrote that many *Waverley* novels altogether’ stands for the same number as ‘the number of counties in Utah’, namely, twenty-nine. Substituting one term for the other in (8.3), we get

The number of counties in Utah is twenty-nine, (8.4)

(again, true) which, according to the Extensionality Principle for Reference 2.3.3, has the same reference as (8.3). Hence each of (8.1) through (8.4) has the same reference. As we have transformed (8.1) in this series of steps to reach (8.4), the proposition or thought or state of affairs expressed has changed completely. So what is it that remains invariant in the transformations by which we reached (8.4) from (8.1)? Truth value. All of the sentences are true. So, Church (1956: 25) concludes,

Elaboration of examples of this kind leads quickly to the conclusion, as at least plausible, that all true sentences have the same denotation, and parallel examples may be used in the same way to suggest that all false sentences have the same denotation.

### 8.3 A Sharpening of Frege’s Argument

There have been several formal sharpenings of Frege’s argument in the literature.<sup>3</sup> Here is a version suggested by Davidson (1969). Assume that any two logically equivalent sentences have the same reference, and also that we have the device of class abstraction. Where ‘*p*’ and ‘*q*’ are any two

sentences that agree in truth value, consider the following sequence of formulas:

$$p, \tag{8.5}$$

$$\{x|(x = x \wedge p)\} = \{x|(x = x)\}, \tag{8.6}$$

$$\{x|(x = x \wedge q)\} = \{x|(x = x)\}, \tag{8.7}$$

$$q. \tag{8.8}$$

(8.5) and (8.6) are logically equivalent, and so, by assumption, they have the same reference. Similarly for (8.7) and (8.8). It remains to show that (8.6) and (8.7) have the same reference. We get (8.8) from (8.7) by replacing the singular term ' $\{x|(x = x \wedge p)\}$ ' by the singular term ' $\{x|(x = x \wedge q)\}$ '; and since, by assumption, ' $p$ ' and ' $q$ ' have the same truth value, these two singular terms have the same reference. So, by the Extensionality Principle 2.3.3, the two sentences (8.6) and (8.7) must have the same reference too. Assuming that sentences refer, then, any two sentences agreeing in truth value must have the same reference.

Formalization has the virtue of brevity, but it also renders the strategy of the argument clear. Take two sentences that have the same truth value but express different propositions. Then transform these different propositions into identities involving definite descriptions for the same object which, although substitutable one for the other *salva veritate*, pick out the object in the different ways embodied in the different propositions. Reference is supposed to be preserved throughout the argument. The substitution of coreferential terms presents no problem;<sup>4</sup> it is only the transformation that must be chosen with care so that it is perceived to preserve reference. In the Davidson version, (8.5) is transformed into the *logically equivalent* (8.6); in the Church version (8.1) is transformed into the *nearly synonymous* (8.2). We need not concern ourselves with these differences here.

#### 8.4 A Problematic Use of Frege's Argument

Frege (1892*c*) was primarily concerned with this extension of the sense/reference distinction to sentences, and with identifying and justifying his choice of the truth values as referents. He devoted fully half his essay to examining purported counterexamples, showing in each case that reference-shifting had occurred inside 'that' clauses.<sup>5</sup> Frege did not use the argument we have examined to show that reference shifted. He believed he was just highlighting and capturing the ordinary use of 'that' clauses. "If words are used in the ordinary way," Frege (1892*c*: 153) said,

“what one intends to speak of is their *Bedeutung*. It can also happen, however, that one wishes to talk about the words themselves or their sense.” When one encloses the words inside quotation marks, one speaks about the words. “In indirect speech,” Frege (1892c: 154) continued, “one talks about the sense, e.g., of another person’s remarks. It is quite clear that in this way of speaking words do not have their customary *Bedeutung* but designate [*bedeuten*] what is usually their sense.”

Frege’s description of our use of *oratio obliqua* constructions is surely incorrect. When we say “Frege thought that Kant was wrong about the epistemological status of arithmetical statements,” it is Kant whom Frege thought erred. *He* is the philosopher Frege spoke about, not (in any straightforward way) the sense of the name. And we are reporting what Frege thought about *him*, not the sense of the name. On the *de re* reading of the modal claim, *Necessarily the number of planets is greater than 7*, the sentence is about the number (that numbers the planets), not the sense of the description; it is that number (there are nine planets) that is necessarily greater than 7. Frege’s description of our use of these *oratio obliqua* constructions (says that . . . ), and, equally, his description of our use of these *oratio recta* constructions (says “. . .”), is clearly inadequate. It is a shaky base on which to construct a theory.<sup>6</sup>

Note, this view he adopts is not forced on him by logical principles. It is a choice of his, an independent observation of usage. Compositionality requires the logical syntax to be such that the reference of the whole has the right relation to the reference of the parts. Frege’s decision is in line with compositionality: the reference of the sentence in a ‘that’ clause is a function of the references of its constituent terms, with the additional machinery involved in the fact that reference shifts from the customary reference to the customary sense.<sup>7</sup> Russell’s (1905) alternative strategy involves a reevaluation of grammatical form because the standard parsing no longer makes the reference of the complex a function of the references of the parts.

This brings us to Quine’s (1953a) problematic use of Frege’s argument to cast doubt on the coherence of quantified modal logic. Let us use the usual symbol  $\Box$  for *Necessarily*, and let us assume (a) that when  $p$  and  $q$  are logically equivalent,  $\Box p$  and  $\Box q$  have the same truth value, and (b) that substitution of coreferential terms within the scope of  $\Box$  preserves truth value. There is little reason to doubt the first assumption. The second assumption is questionable – in fact, it is just the assumption Quine eventually dismisses. But on these two assumptions, Quine argues, it would turn out that  $\Box$  is a truth-functional operator.

For Quine's argument, we need only preface each step of the Davidson version with a  $\Box$ :

$$\Box p, \tag{8.9}$$

$$\Box \{x|(x = x \wedge p)\} = \{x|(x = x)\}, \tag{8.10}$$

$$\Box \{x|(x = x \wedge q)\} = \{x|(x = x)\}, \tag{8.11}$$

$$\Box q. \tag{8.12}$$

Since (8.8) and (8.6) are logically equivalent, as we saw earlier, (8.9) and (8.10) must have the same truth value, by assumption (a). Similarly for (8.11) and (8.12). We get (8.11) from (8.10) by substituting the singular term ' $\{x|(x = x) \wedge q\}$ ' for the coreferential ' $\{x|(x = x) \wedge p\}$ ', so, by assumption (b), (8.10) and (8.11) must have the same truth value. We have our conclusion: when  $p$  and  $q$  have the same truth value,  $\Box p$  and  $\Box q$  have the same truth value, that is,  $\Box$  is a truth-functional operator.

But  $\Box$  is notoriously *not* a truth-functional operator. For, although these two have the same truth value:

$$9 > 7, \tag{8.13}$$

$$\text{The number of the planets} > 7, \tag{8.14}$$

these two do not:

$$\Box 9 > 7. \tag{8.15}$$

$$\Box \text{The number of the planets} > 7. \tag{8.16}$$

So, by Modus Tollens, the problematic assumption in the argument, (b), is to be rejected: the positions occupied by the singular terms occurring inside the scope of  $\Box$  are, in Quine's well-known terminology, *referentially opaque*. Using Frege's substitutability criterion of *aboutness*, Quine (1953a) takes the failure of substitutability just noted as casting doubt on the possibility of understanding  $\Box \varphi x$  as a genuine open sentence purely *about* an object  $x$ .

There is an ambiguity in (8.15), however. On the one hand, it can be read *de dicto*, affirming the necessity of a proposition. This is the reading Quine assigns to (8.15). On the other hand, it can be read *de re*, affirming *of* an object that *it* has a property necessarily. This is the reading Quine (as well as Frege) overlooks. On this reading, the manner in which the number is picked out is irrelevant to the truth of the claim. In the case of (8.15), the *de dicto* and the *de re* readings agree in truth value, but not so in the case of (8.16).<sup>8</sup> This ambiguity plays into, and underscores, a

fundamental strategic error in Quine's use of the Frege argument: someone who finds the *de re* reading reasonably clear will conclude from the argument that the logical syntax Quine employs must be faulty – and that is the moral defenders of modal logic have drawn. Because there are these two intelligible readings that differ in truth value, (8.16) is outright ambiguous: traditional first-order logical notation is therefore inadequate when modal operators are added. A new syntax is required.<sup>9</sup> Hence the syntactical assumptions built into Quine's argument – in particular, whether one expression is a logical constituent of another – are unreliable.

How do these observations impact on Quine's argument concerning (8.9)–(8.12)? There are two types of substitutions in the argument. On the one hand, we substitute one *sentence* for another that is logically equivalent, to get (8.10) from (8.9) and, again, to get (8.12) from (8.11). On the other hand, we substitute one *singular term* for another having the same reference, to get (8.11) from (8.10). Whatever syntax is ultimately accepted, the *de dicto* reading should permit the substitution of the logically equivalent sentences but not necessarily of the coreferential singular terms. The *de re* reading should permit the substitution of the coreferential singular terms but not necessarily of the logically equivalent sentences.<sup>10</sup>

Quine is quite right about the *de dicto* reading, which is not, in his terminology, *referentially transparent*. The coherence of attaching  $\Box$  to open sentences is, however, untarnished. Of course, there still remains the technical problem of linking up the two readings in a formal setting.<sup>11</sup> But it is important to recognize that Quine did not eliminate a *de re* reading, which *is* referentially transparent. He simply overlooked it. Frege, of course, acknowledged only a *de dicto* reading of the modality, not the *de re* reading.<sup>12</sup> For Quine to rely on Frege's argument is just to stay within the framework of modality Frege set up. It remains a mystery that Quine relies so heavily on Gödel's (1944) sharpening of Frege's argument, for Gödel put it forward within a context in which he expressed the belief that Russell's theory of descriptions provided a path out of the Fregean conclusion that sentences stand for their truth values. In any event, we now show that Gödel's conjecture is correct.

### 8.5 A Way out of Frege's Argument

We have seen how we can use Russell's account of descriptions to maintain that  $\Box$  generates a context that is not truth functional. The next step is to



show how Russell's account of descriptions can be used to maintain that

'S' refers to \_\_\_\_ (8.17)

is not a truth-functional context. Frege relies on his syntactical treatment of definite descriptions as *Eigennamen*, that is, as individual constants. The alternative Russellian treatment blocks the Frege argument in exactly the same way as we showed it to block Quine's argument.

We have seen the Frege argument now in a number of different guises. Frege asks us to suppose that a declarative sentence has a reference. Let us do so. Surely, then, this should be true:<sup>13</sup>

'S' refers to S. (8.18)

Suppose this to be true:

'S' and 'T' have the same truth value. (8.19)

From (8.18) and (8.19), together with some steps involving substitution of coreferential definite descriptions, Frege concludes

'S' refers to T. (8.20)

Now, to infer (8.20) from (8.18) and (8.19) is to argue that the context (8.17) is truth functional. But Frege is actually assuming that it is truth functional in order to perform the appropriate substitution on the definite descriptions.

If (8.17) is a truth-functional context, the Frege result follows trivially: every true sentence stands for the same thing, and likewise every false sentence. If, on the other hand, (8.17) is not a truth-functional context, the Frege result is trivially false. In setting up Frege's argument, our focus on the Extensionality Principle for Reference 2.3.3 has perhaps been a bit infelicitous. For, as we noted in Section 2.5, the special Substitution Principle for Reference 2.5.1, which is based on it, is false. Substitution of coreferential terms preserves truth value only if the sentence is *about* the object referred to. That was the correction we ultimately came to with the Corrected Substitution Principle 2.5.4. What we are learning now is that *aboutness* is intimately connected with truth-functionality. Frege's argument was actually question begging. Frege could not accept that the descriptions would be about what they ordinarily refer if (8.17) were not a truth-functional context; for in that case reference would shift. But that is because he took descriptions to be *Eigennamen*.

Russell, however, provides an alternative explanation for the case.<sup>14</sup> Russell, as we have already remarked, thought that it was wrong to think

of a sentence as a name: a sentence does not, he said, stand for an object in the same way as a name stands for its bearer or a definite description stands for the object it uniquely describes. A sentence containing a name or definite description will, for the most part, be about that object; but one cannot talk of a sentence itself as being about anything. Nonetheless, let us set aside Russell's qualms for the purposes of the argument and grant Frege his supposition. What is being granted, to be clear, is that a sentence stands for something. But Russell need not grant that (8.17) is truth functional. Frege must prove it. That is the task of the Frege argument.

Russell has a story about descriptions that blocks the Frege argument. Let 'S' be the sentence '*k* is *G*', where '*k*' is some description. There are, on Russell's account of descriptions, two ways of understanding

$$\text{'S' refers to } k \text{ is } G. \quad (8.21)$$

There is the small-scope reading,

$$\text{'S' refers to } \langle \lambda x. Gx \rangle (k), \quad (8.22)$$

and the large-scope reading,

$$\langle \lambda x. \text{'S' refers to } Gx \rangle (k). \quad (8.23)$$

On the large-scope reading, the manner in which we pick out the object makes no difference; but on the small-scope reading it does make a difference, even if that object exists. Substitution of coreferential terms preserves truth in (8.23), but not in (8.22), just as in the modal case. The belief that the descriptions are about what they ordinarily refer to is, in effect, to impose the large-scope reading on the description: that is why substitution of coreferential descriptions preserves truth. But, of course, there is always this other reading of the descriptions, the small-scope reading, which, *because* the context is not truth functional, will block the substitutions. So Russell is able to explain why the substitutions work without admitting the context is truth functional.

The Frege argument, as we have reconstructed it, turns critically on Frege's understanding of definite descriptions as *Eigennamen*. The different predicates are absorbed into the description, and then the differences are ignored, for at the level of *Bedeutung*, two descriptions for the same object are interchangeable, preserving truth. Frege's substitution rule governing descriptions effectively treats descriptions as individual constants. Russell (1905) had a different substitution rule: substitutions

for definite descriptions depended significantly on whether the context in which the substitution was carried out was truth functional. For in a context that is not truth functional, the significance of the difference in the predicates used to pick out the individual is registered. There are many questions raised by these results about the connection between the Frege argument and the two theories of descriptions – Frege’s and Russell’s – but pursuing them will take us too far afield. The important result here is that the Frege argument is not as compelling as it has seemed.

### 8.6 Truth and Assertion

Russell objected to taking sentences as names of truth values. In a letter to Frege dated 2 February 1903, Russell (1903*a*: 155–6) says:

I have read your essay on sense and meaning, but I am still in doubt about your theory of truth-values, if only because it appears paradoxical to me. I believe that a judgment, or even a thought, is something so entirely peculiar that the theory of proper names has no application to it.

In another letter, dated 12 December 1904, he returns to the issue. Speaking of the proposition that Mont Blanc is more than 4000 meters high, Russell (1903*a*: 169) says “for me the *meaning* of a proposition is not the true, but a certain complex which (in the given case) is true.”

Russell’s unhappiness has been voiced as well by Black (1954: 229–30):

We may assume that if *A* and *B* are designations of the same thing the substitution of one for the other in any declarative sentence will never result in nonsense. This assumption would not have been questioned by Frege. Let *A* be the sentence “Three is a prime” and *B* the expression “the True.” Now “If three is a prime then three has no factors” is a sensible declarative sentence; substitute *B* for *A* and we get the nonsense “If the True then three has no factors.” The last form of words has no more use than “If seven then three has no factors” or indeed any form of words containing an expression of the form “If *X* then . . .” where “*X*” is replaced by a designation. Hence, according to our assumption, *A* and *B* are not designations of the same thing – which is what we set out to prove.

Black (1954: 229) considers this argument “a sufficient refutation of Frege’s view that sentences are designations of truth values.” However, it has become routine in logic to treat a sentence as a name of a truth value, and even to introduce the name of a truth value as a constant that occupies sentence position.<sup>15</sup> Given the simplicity, elegance, and

fruitfulness of this treatment, it is attractive to say that Black has not produced meaningfulness, but merely oddness. His argument does not constitute a refutation of Frege's view, nor does it indicate that further clarification will lead to anything that causes us to reject Frege's view.

Part of the oddness stems from the fact that, in natural language, singular terms and sentences belong to different syntactic categories. Frege assigns them to the same syntactic category: *Eigennamen*, he calls them. We have seen no logical difficulty in this treatment. But part of the oddness stems from the fact that sentences are asserted, while singular terms are not. Frege had much to say about this.

Words like 'judgment' and 'assertion' exhibit a process/product ambiguity. By an assertion we may mean either that which is asserted or the asserting of it.<sup>16</sup> This distinction is rather easy to miss. Frege (1915: 251) puts it this way: "When something is judged to be the case, we can always cull out the thought that is recognized as true; the act of judgment forms no part of this." In a conditional, neither the antecedent nor the consequent is asserted or judged to be true; nonetheless each part of the conditional is a complete thought, a proposition.

Frege (1893: 35) assimilated these nonassertive uses of sentences to designating: "I do not mean to assert anything if I merely write down an equation, but . . . I merely *designate* a truth value, just as I do not assert anything if I merely write down ' $2^2$ ,' but merely *designate* a number." He introduced the special sign ' $\vdash$ ' to indicate that what follows it is being asserted. ' $\vdash 2^2 = 4$ ' is not a name at all; it does not denote, or even purport to denote, anything. ' $\vdash$ ' is rather an illocutionary operator – 'It is hereby asserted that' – which attaches to a name of a truth value to make an assertion of the thought expressed by the sentence to which the operator is attached.

The treatment of the sign ' $\vdash$ ' in Frege (1893) is in marked contrast to its treatment in Frege (1879: 54), where, not having clearly distinguished between a content and the judging of the content to be true, he claimed that the assertion sign was a predicate:

Imagine a language in which the proposition 'Archimedes was killed at the capture of Syracuse' is expressed in the following way: 'The violent death of Archimedes at the capture of Syracuse is a fact'. Even here, if one wants, subject and predicate can be distinguished, but the subject contains the whole content, and the predicate serves only to present it as a judgment. *Such a language would have only a single predicate for all judgments, namely, 'is a fact'. . . . Our Begriffsschrift is such a language and the symbol  $\vdash$  is its common predicate for all judgments.*

Frege (1879) thought that attaching 'It is true that' turned an utterance into an assertion. This is the only place we find truth or falsity in *Begriffsschrift*. Frege (1879) accounted for the properties of the logical connectives rather in terms of their being *affirmed* or *denied*. Frege (1893) has entirely wrung out the assertive aspect from sentences, and especially from the copula, with which it has commonly been associated. Now sentences designate truth values, and the connectives are explained in terms of truth value. Their assertive role has been transferred to his special sign '⊢'. Unasserted sentences are nevertheless true or false, and so the truth values belong not to '⊢', but to the referential apparatus of the notation.

The sign ⊢ is actually a combination of two signs, a vertical stroke '⊥' and a horizontal stroke '–'. The horizontal stroke turns what follows it into the name of a truth value, and the vertical stroke indicates that what follows is being asserted. The horizontal stroke can occur without the vertical stroke (but not conversely). '–' is a one-place function-expression that attaches to a name to form the name of a truth value:  $-\Delta$  "is the True if  $\Delta$  is the True; on the other hand it is the False if  $\Delta$  is not the True" (Frege 1893: 38). Since it is a function that maps objects to truth values, it stands for a concept, a concept under which a single object, the True, falls, namely, the concept *being true* or *being the True*. As usual, Frege defines the function for any argument whatsoever. When the horizontal sign is attached to the name of a truth value, the whole refers to the same truth value, but when the horizontal sign is attached to a name for something other than a truth value, then the whole refers to the False. Although sentences and ordinary singular terms belong to the same syntactic category, not all names are assertible, but only names of truth values. In this way, Frege honors the intuitive difference that Russell and Black thought Frege had ignored.

### 8.7 Is 'True' a Predicate?

If we read '– $\Delta$ ' as 'It is true that  $\Delta$ ', it would appear that 'true' is a concept word, not a name for an object. But this is not so. Taking our cue from Frege's (1892b) distinction between the planet Venus and the concept *being none other than Venus* – the concept denoted by ' $x = \text{Venus}$ ' – we might take the concept *being true* to be *being identical with the True*, the concept denoted by ' $x = \text{the True}$ '.<sup>17</sup> Then we understand

It is true that 5 is a prime number (8.24)

to be an identity,

$$5 \text{ is a prime number} = \text{the True.} \quad (8.25)$$

So, 'true' is no more a concept word than is 'Venus': (8.24) says that something falls under the concept *being identical with the True*, not that something falls under the concept *true*.

Indeed, it is futile to persist in supposing that 'true' is a concept word once we have accepted Frege's function/argument analysis of sentences. For if we try to regard 'true' as a concept word that attaches to a name (of whatever entity) to create a more complex name, then since each of these are names, and so refer to objects, the argument from Section 8.2 shows that they refer to truth values. That is, once we try to regard *true* as a function, we are forced to take the truth values as objects – so there is no point in even trying to regard it as a function.<sup>18</sup>

It is a consequence of Frege's view that (8.24) and

$$5 \text{ is a prime number} \quad (8.26)$$

must have the very same reference, that is, the same truth value. In modern terms, he holds

$$\text{The proposition that } p \text{ is true iff } p. \quad (8.27)$$

However, Frege (1892c: 158) goes a bit further in his account of truth:

One might be tempted to regard the relation of the thought to the True not as that of sense to *Bedeutung*, but rather as that of subject to predicate. 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'.<sup>19</sup>

The claim is that (8.24) and (8.26) express the same proposition.<sup>20</sup> This is a much stronger claim than (8.27), and puts pressure on his theory. Frege (1915: 251–2) says:

This may lead us to think that the word 'true' has no sense at all. But in that case a sentence in which 'true' occurred as a predicate would have no sense either. All one can say is: the word 'true' has a sense that contributes nothing to the sense of the whole sentence in which it occurs as a predicate.

But his theory leaves no room for an expression to have sense, and yet not to contribute anything toward the sense of an expression containing it. This remains an open problem for Frege's story about truth.

## 8.8 The Correspondence Theory of Truth

There remains a nagging suspicion that Frege has got the story about the reference of a sentence wrong. What objects are these, the True and the False?

Facts are especially strong candidates for being the referents of sentences because of their tie with truth. It is commonly held that the truth of a sentence resides in its correspondence with the facts. Tarski (1944: 15), for example, says: "If... we should decide to extend the popular usage of the term '*designate*' by applying it not only to names, but also to sentences," then the following formulates the philosophical view of truth he seeks to make precise: "*A sentence is true if it designates an existing state of affairs.*" Viewed from this perspective, it would seem that Frege has got the analysis of sentences all wrong. A sentence does not stand for a truth value, one is inclined to say, but for a (possible) fact, and truth comes in when the (possible) fact named by the sentence obtains.

The problem with supposing that sentences designate (possible) facts is, as we have already noted, that the fine distinctions wanted are not forthcoming. For, given the result of the Frege argument in Section 8.2, all true sentences must name the same fact – the *Great Fact*, as Davidson (1969) calls it. Gödel (1944: 214) has remarked on the metaphysical character of Frege's notion of the True, "reminding one somewhat of the Eleatic doctrine of the 'One'." The resemblance is certainly striking. For the True – or the Great Fact or Reality – appears to be an undifferentiated totality much like Parmenides' Being. But there is a very significant difference: whereas Parmenides admitted Being, and Being only, Frege appears to admit both Being and Nonbeing. False sentences, too, are names. They are names of the False. So for Frege there appears to be a Great NonFact alongside of the Great Fact.

However, if we have both a Great Fact and a Great NonFact, then the neat relation between facts and truth no longer holds. Recall that facts were introduced in order to explain truth: a sentence was said to be true if the (possible) fact it named obtains. But, then, a false sentence, being false, stands for the Great NonFact, and since the Great NonFact obtains, the sentence must be true. Hence it turns out that if a sentence has any truth value at all, it has the truth value *true*. The idea behind the appeal to facts, however, is that a sentence is, metaphorically speaking, aimed at reality, and the sentence is true if it reaches its intended mark, and false if it does not. A natural attempt at patching up this account of truth would

be to say that a sentence is true if and only if it names the Great Fact, and false if, and only if, it names the Great NonFact. But this too fails. We do not expect reality to include both Reality and UnReality.

We have been assuming that there is something outside the realm of sentences (or thoughts) in virtue of which sentences (or thoughts) are said to be true; and we had, quite literally, assumed the True to be that thing. We had attempted to fit the True into what was essentially a Correspondence Theory of Truth. But Frege explicitly rejects the Correspondence Theory of Truth, and the True is not supposed to play anything like the role in his metaphysical scheme that Davidson and Gödel would have us think. Frege (1918: 327) argues that “truth does not consist in correspondence of the sense [of a sentence] with something else, for otherwise the question of truth would get reiterated to infinity.”

Frege casts the Correspondence Theory as an account of our description of a picture, assumed to depict something, as being a *true* picture:

It might be supposed from this that truth consists in a correspondence of a picture to what it depicts. Now a correspondence is a relation. But this goes against the use of the word ‘true’, which is not a relative term and contains no indication of anything else to which something is to correspond. 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. (Frege 1918: 326–7)

According to the Correspondence Theory, we are to imagine that we match up a picture with the item the picture is intended to represent, and if the two correspond, then the picture is said to be a true picture. But, Frege points out, truth itself is not a correspondence relation; rather, we must assume some correspondence scheme linking pictures with the things they are intended to depict, and then define truth for pictures in terms of their fidelity *modulo* this correspondence scheme. That is, we determine whether a particular picture is true by determining whether that picture *in fact* corresponds to the item it was intended to depict. But in that case, Frege (1918: 327) argues, the attempt to define truth as correspondence leads to a vicious regress:

But could we not maintain that there is truth when there is correspondence in a certain respect? But which respect? For in that case what ought we to do so as to decide whether something is true? We should have to inquire whether it is *true* that an idea and a reality, say, correspond in the specified respect. And then we should be confronted by a question of the same kind, and the game could begin again. So the attempted explanation of truth as correspondence breaks down.



The problem is that we do not lay down a correspondence scheme and afterward raise the question of truth: truth is already being assumed in the setting up of the scheme itself.

So what are these objects, the True and the False? When we speak of names as expressions that refer to objects, and we suppose that we use these names to speak about these objects, then we believe we should be able to identify them in some way. If a true sentence stands for the True, we suppose the sentence is about it; then we look to find the object referred to and see whether it is of such-and-such a sort so as to render the sentence true or false. But things are altogether different for sentences. In the case of sentences, it appears that first we determine the truth value of the sentence and thereby determine the object stood for. Can this be right? Frege thinks that it is. It is a mistake, from Frege's point of view, to search for and examine these abstract objects themselves. This is just the path he admonishes us from following in the famous Context Principle: "never to ask for the meaning of a word in isolation, but only in the context of a proposition" (Frege 1884*b*: x). Frege has given us a very fruitful and precise account of reference through his development of the compositionality principles. But his unfolding of these principles has moved us far from the original intuitions described back in Chapter 2 connecting up *standing for* and aboutness, and the question remains whether the Context Principle provides us with enough of a story to satisfy us that sentences do stand for these objects.

Frege not only developed modern logic, but, with his compositionality principles, he developed a very powerful picture with strong metaphysical overtones. In this chapter, we have been drawing out his insights pertaining particularly to the notion of *truth*. What we have found, interestingly enough, using Russell as our foil, is that these metaphysical implications are not *consequences* forced upon us by his logical analysis. Rather, they are beginning to look more like *assumptions* within which Frege works his syntactic engine. Frege's analysis leads him to deny one of the best entrenched traditional accounts of truth, the Correspondence Theory, and to suggest accounts of truth that are more in line with the minimalist notions that prevail today. This is particularly noteworthy because we usually find the Correspondence Theory a handmaiden to Realism. Frege's break with this tradition presents a challenge: Are we to regard Frege as a genuine Realist even though he rejects the Correspondence Theory of Truth, or is his rejection of the Correspondence Theory one more piece of evidence that he is not really a Realist at all?<sup>21</sup>

## Indirect Reference

### 9.1 Introduction

Frege's story about indirect contexts – 'that' clauses, like 'Harry believes that' and 'Joan said that' – is widely known and enormously influential. And yet it is only the briefest of sketches. Here is what he says in "On Sense and Reference":

In indirect speech one talks about the sense, e.g., of another person's remarks. It is quite clear that in this way of speaking words do not have their customary *Bedeutung* but designate [*bedeuten*] what is usually their sense. In order to have a short expression, we will say: in indirect speech, words are used *indirectly* or have their *indirect Bedeutung*. We distinguish accordingly the *customary* from the *indirect Bedeutung* of a word; and its *customary* sense from its *indirect* sense. The indirect *Bedeutung* of a word is accordingly its customary sense. (Frege 1892c: 154)

It is quite elegant, the way he has knitted together the sense/reference distinction and the problem of substitutivity in oblique contexts. But it is not as tight as one might think. For we know what the customary reference of an expression is supposed to be: the customary reference is the thing the word stands for. We also know what the customary sense of an expression is supposed to be: the customary sense is, roughly, the meaning of the expression. So we know what the indirect reference is supposed to be: Frege explicitly identifies the indirect reference with the customary sense. But what is the indirect sense supposed to be?

We can, of course, *speak* of the indirect sense of the sentence 'the cat is on the mat'. We can even *say*, in the Fregean spirit, that the indirect sense of that sentence contains the mode of presentation of its ordinary sense, the thought it expresses, namely, *that the cat is on the mat*. We might

even identify the indirect sense of the sentence ‘the cat is on the mat’ with the customary sense of the name of the thought, *that the cat is on the mat*; further, we might contrast it with a different way of introducing that thought, for example, as the thought J. L. Austin etched on the modern philosophical consciousness in his discussion of truth. It is not entirely clear that we can cash in this way of speaking for hard semantics. But even if we can, the problem remains that a sentence can be embedded and reembedded in ‘that’ clauses to whatever depth, and, to accommodate these multiple embeddings, we need some *general* method for determining indirect sense.

A good dictionary will provide the customary sense of a word, and, if the entry includes examples of things to which the word applies, it will provide the customary reference as well. But we do not expect to find the indirect sense included in the dictionary entry. Indirect sense should be structurally determined: there should be some way of computing the indirect sense of a word using the customary sense and reference of the word in combination with the properties of the ‘that’ operator. But, as Russell (1905: 50) observed, “there is no backward road from denotations to meanings.”<sup>1</sup> Indeed, saying how indirect sense is to be determined has turned out to be a task of considerable difficulty.

Russell, as already mentioned, appears to have understood the problem; but it is Carnap who is usually credited with raising the issue of the *infinite hierarchy* for Frege’s semantics. Carnap (1947) charged that Frege required an infinite number of distinct names for indirect senses. Davidson made the criticism more pointed. Since the indirect sense of a word cannot possibly be a function of its customary sense, Davidson (1968–9: 99) argued, Frege is committed to an infinite number of semantic primitives, an absurd requirement for any natural language:

Neither the languages Frege suggests as models for natural languages nor the languages described by Church are amenable to theory in the sense of a truth definition meeting Tarski’s standards. What stands in the way in Frege’s case is that every referring expression has an infinite number of entities it may refer to, depending on the context, and there is no rule that gives the reference in more complex contexts on the basis of the reference in simpler ones. In Church’s languages, there is an infinite number of primitive expressions; this directly blocks the possibility of recursively characterizing a truth predicate satisfying Tarski’s requirements.

Most philosophers do regard this prospect of an infinite hierarchy of indirect senses as a *reductio* of the theory.

Explicitly citing Russell's (1905) admonition and recoiling from the hierarchy, Dummett (1981*a*: 268–9) has urged that Frege's semantic theory be reshaped so that there is no indirect sense distinct from customary sense:

According to Frege, a word does not have a reference *on its own*, 'considered in isolation': it has a reference only in the context of a sentence. It is fully harmonious with this view to hold that, while a word or expression *by itself* has a sense, it does not by itself have a reference at all: only a particular occurrence of a word or expression in a sentence has a reference, and this reference is determined jointly by the sense of the word and the kind of context in which it occurs. The sense of a word may thus be such as to determine it to stand for one thing in one kind of context, and for a different thing in some other kind of context. We may therefore regard an expression occurring in an opaque context as having the same sense as in a transparent context, though a different reference. . . .

With this emendation, there is no such thing as the indirect sense of a word: there is just its sense, which determines it to have in transparent contexts a reference distinct from this sense, and in opaque contexts a referent which coincides with its sense. There is therefore no reason to think that an expression occurring in double *oratio obliqua* has a sense or a reference different from that which it has in single *oratio obliqua*: its referent in double *oratio obliqua* will be the sense which it has in single *oratio obliqua*, which is the same as the sense it has in ordinary contexts, which is the same as its referent in single *oratio obliqua*. This is intuitively reasonable: the replacements of an expression in double *oratio obliqua* which will leave the truth-value of the whole sentence unaltered are – just as in single *oratio obliqua* – those which have the same sense.

Dummett's view, then, is that the indirect sense of a word just *is* its customary sense. In ordinary contexts, the word stands for its customary reference, but in indirect contexts (at whatever level of indirectness), the word stands for its customary sense.

Let us, following Parsons (1981),<sup>2</sup> call an infinite hierarchy *rigid* if expressions that agree in customary sense agree at every level of indirect sense; otherwise it will be *nonrigid*. Parsons has shown that Dummett's reconstruction of Frege is equivalent to a rigid hierarchy. In a rigid hierarchy, there is a functional relation between customary sense and indirect sense, for, if two expressions have the same customary sense, they have the same indirect sense. This does not mean that each level of indirect sense is identical with customary sense. But it does mean that words having the same customary sense are interchangeable one with the other at any level of indirectness. Within this framework, Dummett's can be viewed as the *smallest* rigid hierarchy, because his functional relation is just identity: customary sense is identical with indirect sense, so all levels collapse, essentially, into the first. But there might be other rigid hierarchies in

which, for each  $i$ , the sense at level  $i$  is distinct from the sense at level  $i + 1$ . Parsons's surprising result is that all rigid hierarchies present equivalent semantic analyses of sentences with multiple embeddings.

The rigid hierarchy, then, constitutes one response to Davidson's charge of absurdity. Some hierarchies are not absurd. In the case of the rigid hierarchy, it might well be that there are infinitely many different indirect senses attached to a given expression. But, because the indirect sense is a function of the customary sense, we do not have to learn them all in order to understand a sentence with multiply embedded 'that' clauses. We need only know the customary sense.

There is much sympathy for Dummett's position among those favorable to a Fregean semantics.<sup>3</sup> But it has not won universal agreement. Some Fregeans prefer to take the bull by the horns and deny that the nonrigid hierarchy leads to absurdity in the way Davidson charges. The nonrigid hierarchy has been advocated by Church (1973), by Anderson (1980), and also by Heidelberger (1975). The challenge before them is to explain the structural connection between indirect sense and customary sense, given that the former is clearly not, on this view, a function of the latter.

The essential idea behind a nonrigid hierarchy is that expressions having the same customary sense might yet differ in indirect sense. Although 'vixen' and 'female fox' have the same customary sense, a person might not know that they do; the expressions are therefore substitutable one for the other in any singly embedded 'that' clause, but not necessarily in a doubly embedded one. The nonrigid hierarchist believes he can distinguish the contribution customary sense makes to the semantic interpretation of a deeply embedded sentence from the contribution made by some (at least) of the indirect senses. This is what differentiates him from the rigid hierarchist. The rigid hierarchist believes that it is difficult to pinpoint at which level of indirectness a failure of substitutivity is to be attributed: whatever reason he has for supposing that two expressions differ in indirect sense (at whatever level) is a reason for supposing they differ in customary sense. His solution is to push all differences back to the first level of customary sense. If one can know the customary sense of the expressions 'vixen' and 'female fox', and yet not realize that they have the same sense, then they do not have the same sense.<sup>4</sup>

We are inclined to believe that Dummett has correctly identified the critical juncture as that point where customary sense is distinguished from indirect sense. For, once indirectness has been established, there is no further problem introduced by deeper levels of indirectness. There

are, however, two very distinct logical ways in which the deeper levels of indirectness are to be treated. This has not been widely recognized. Our contention is that Dummett confused these two analyses, and in finessing one horn of Russell's dilemma he finds himself impaled on the other horn, which is the collapse of the sense/reference distinction.<sup>5</sup> The source of the confusion, in large measure, is the analogy, originally noted by Frege, between *oratio obliqua* and *oratio recta* constructions. There is an analogy, but Frege's misreading of the one has infected the reading of the other. We will examine *oratio recta* constructions in Chapter 10, and in this chapter we will focus on the *oratio obliqua* construction. After rehearsing Frege's theory in Sections 9.2 and 9.3, we provide whatever textual evidence we have for the infinite hierarchy in Section 9.4. Dummett's suggestion is introduced in Section 9.5, and a reconstruction of Russell's argument is advanced in Section 9.6. Finally, in Section 9.7, we identify the two readings, and show that if we keep the two analyses firmly separate, Dummett's interpretation would work.

## 9.2 The Sense/Reference Story

Let us begin by consolidating the story about sense and reference.

A proper name, like 'Margaret Thatcher', has both a sense and a reference. The reference of the name is the woman herself: she is whom you talk about, refer to, mean, if you like, when you ordinarily use the name in conversation. The sense of the name, on the other hand, is, very roughly, whatever it is that enables you to place, pick out, identify, or locate the person you speak about. The sense you attach to the name could be 'the first female prime minister of Great Britain' or it could be 'Sir Denis Thatcher's widow', or it could be something else.

Frege's syntactical analysis of

Margaret Thatcher drives a Peugeot, (9.1)

is that the name 'Margaret Thatcher' combines with the predicate '( ) drives a Peugeot' to form a declarative sentence. Frege regards a declarative sentence as a complex name. His semantic analysis of (9.1) comes in two parts. On the *Bedeutung* side, 'Margaret Thatcher' refers to the woman and '( ) drives a Peugeot' refers to a concept, and the complex name refers to the value of that function for that argument. So (9.1) is a name of a truth value. On the *Sinn* side, 'Margaret Thatcher' expresses a sense of the woman, '( ) drives a Peugeot' expresses a sense-function, and the two combine to form the sense of the whole sentence, the thought or proposition that Margaret Thatcher drives a Peugeot.

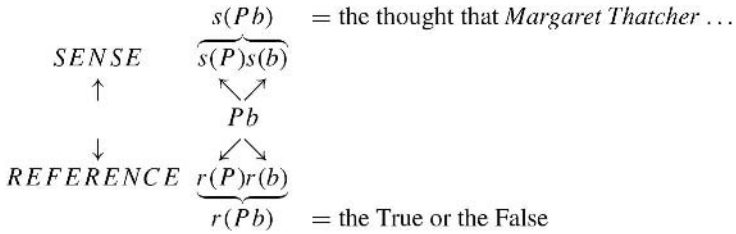


FIGURE 9.1

We continue to use  $r(\eta)$  for the *reference* of  $\eta$ , and  $s(\eta)$  for the *sense* of  $\eta$ . We use curly braces  $\underbrace{\quad}$  for the relevant combining of senses or references. Now, abbreviating ‘Margaret Thatcher’ to ‘b’, and ‘( )’ drives a Peugeot to ‘P’, the semantic analysis of (9.1) is given in Figure 9.1.<sup>6</sup>

Here is a summary of the principles governing Frege’s semantic theory that we have advanced to this point. Frege characterizes *Sinn* in a number of different ways: as conventional significance, as the common store of knowledge of the referent, as mode of presentation, as an individual’s way of picking out an object. Let us not focus on these differences and the unclarities they generate. The central facts are that the sense of a complex is composed out of the senses of its parts,

PRINCIPLE 3.6.3 (COMPOSITIONALITY FOR SENSE)  $s(\theta(\alpha)) = s(\theta)[s(\alpha)]$ ,

and the sense of a complex is uniquely determined by the sense of its parts,

PRINCIPLE 3.6.4 (EXTENSIONALITY FOR SENSE) *If  $s(\alpha) = s(\beta)$ , then  $s(\theta(\alpha)) = s(\theta(\beta))$ .*

These two principles capture the relation between the sense of a part and the sense of a complex.

Now let us look at the relation between the sense of an expression and its reference. A term refers to what the sense determines; so, although we speak of a term’s referring, it is the sense of the term that does the work. We can even say that it is the sense of the term that refers:

PRINCIPLE 3.6.1 (SENSE DETERMINES REFERENCE)  $r(\eta) = r(s(\eta))$ .

Principle 3.6.1 expresses one part of Frege’s view that sense determines reference; the other, the uniqueness of the referent, that is, the fact that  $r$  is a function, is given by

PRINCIPLE 3.6.2 (REFERENCE IS A FUNCTION) *If  $s(\eta) = s(\zeta)$ , then  $r(\eta) = r(\zeta)$ .*

Principles 3.6.1 and 3.6.2, which connect up sense and reference, together with Principles 3.6.3 and 3.6.4, which govern sense, enable us to derive Frege's two fundamental principles governing reference:

PRINCIPLE 2.3.1 (COMPOSITIONALITY FOR REFERENCE) *For any function-expression  $\theta(\Omega)$  and any name  $\alpha$ ,  $r(\theta(\alpha)) = r(\theta)[r(\alpha)]$ ,*

and

PRINCIPLE 2.3.3 (EXTENSIONALITY FOR REFERENCE) *For any function-expression  $\theta(\Omega)$  and any names  $\alpha, \beta$ , if  $r(\alpha) = r(\beta)$ , then  $r(\theta(\alpha)) = r(\theta(\beta))$ .*

Principle 2.3.1 says that all significant parts of the sentence refer. Principle 2.3.3 defines the functional relation between the reference of a complex name and the reference of its constituent singular terms. A name is complex for Frege if, and only if, Principle 2.3.3 holds for that name; so Principle 2.3.3 actually serves, as we noted in Section 2.3, as a parsing principle for identifying the significant parts of a sentence. Principles 3.6.3 and 2.3.1 are frequently identified as *Compositionality* Principles. Principles 3.6.4 and 2.3.3 are frequently identified as *Substitution* Principles. These six principles form the heart of the sense/reference story.

Let us now consider what happens when (9.1) is embedded in a 'that' clause, as for example, in

Ted Kennedy believes that Margaret Thatcher drives a Peugeot. (9.2)

We cannot replace 'Margaret Thatcher' by just any coreferential singular term and preserve the truth value of the sentence. Nor can we replace the embedded sentence by just any sentence having the same truth value and preserve the truth value of (9.2). Principles 2.3.1 and 2.3.3 fail when a declarative sentence is embedded in a propositional attitude context; and since these are derived from Principles 3.6.3 and 3.6.4, Principles 3.6.3 and 3.6.4 fail as well.

Frege could not be satisfied with leaving the matter like this. For one thing, he would be abandoning compositionality for a large class of sentences; and compositionality was a compelling idea for Frege (just as it is for many philosophers today). But there is another, more critical, reason. Frege had said that the sense of a declarative sentence is a thought or a proposition: this is what the sentence expresses. Sentence (9.1) therefore expresses the proposition that Margaret Thatcher drives a Peugeot. Frege had to be able to tell a convincing story that 'expresses that' and 'says that' related a sentence (or person) to a proposition, or else his



claim that a sentence expresses a proposition would be incomprehensible. And, of course, the same story would have to be told for the other propositional attitude verbs, because that is the role thoughts or propositions are supposed to play. So the story Frege told about oblique contexts is not an afterthought or an add-on to the basic account; it is a central component of the picture.

How did Frege handle these oblique contexts? We gave the outlines of his treatment in Section 3.6. Now we must look at it more closely. There are two parts to his solution. First, he says that ‘that’ shifts the reference of the words in its scope, and, second, he relativizes reference to the context in which the term occurs. The *customary* reference of the embedded sentence in (9.2) is its truth value; but it is not referring to its truth value in that context, so replacing it by an equipollent sentence need not preserve the truth value of (9.2). Just because two names have the same customary reference, that is, the same reference in one kind of context, they need not have the same reference in every context in which they occur. But if two names have the same reference appropriate for the context in which they occur, then they are substitutable in that context *salva veritate*. The appropriate reference of a term embedded in a ‘that’ clause is its *indirect* reference. So the truth value of (9.2) is a function of the indirect reference of its constituent ‘Margaret Thatcher’, and substituting another term for ‘Margaret Thatcher’ that has the same indirect reference should leave the truth value of (9.2) unchanged. Compositionality is preserved: (9.2) is regarded as having parts whose reference contributes to determining the reference of the whole.

Let us use  $s_0(t)$  and  $r_0(t)$  for the customary sense and reference, respectively, of  $t$ , and  $s_1(t)$  and  $r_1(t)$  for the indirect sense and reference, respectively, of  $t$ .<sup>7</sup> And let us abbreviate ‘Ted Kennedy believes that ( )’ to ‘K’. Then Frege’s semantic analysis of (9.2) is given in Figure 9.2.

Let us introduce some more formalization – just enough to clarify the analysis. First, we will use ‘ $\Theta$ ’ for ‘that’, with parentheses when needed to clarify scope. Frege’s view is that the indirect reference of an expression is what the expression refers to inside the scope of ‘that’. We will express this principle as follows:

PRINCIPLE 9.2.1 (INDIRECT REFERENCE)  $r_1(t) = r_0(\Theta(t))$ .

Principle 9.2.1, in effect, *defines*  $r_1(t)$ , the indirect reference of a term  $t$ : the indirect reference is the customary reference of an expression inside

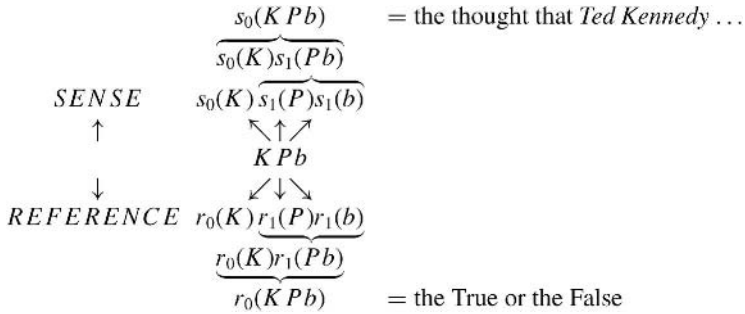


FIGURE 9.2

a ‘that’ clause. Next, we assume that indirect reference is compositional:

PRINCIPLE 9.2.2 (COMPOSITIONALITY FOR INDIRECT REFERENCE)

$$r_1(\theta(\alpha)) = r_1(\theta)[r_1(\alpha)].$$

This means that

$$\text{PRINCIPLE 9.2.3 (THAT)} \quad r_0(\Theta(\theta(\alpha))) = r_0(\Theta(\theta))r_0[(\Theta(\alpha))].$$

So the compositionality and extensionality principles necessary for handling ‘that’ clauses are

$$\text{PRINCIPLE 9.2.4 (COMPOSITIONALITY FOR THAT)} \quad r_0(\theta(\Theta(t))) = r_0(\theta)[r_1(t)],$$

and

$$\text{PRINCIPLE 9.2.5 (EXTENSIONALITY FOR THAT)} \quad \text{If } r_1(\alpha) = r_1(\beta), \text{ then } r_0(\theta(\Theta(\alpha))) = r_0(\theta(\Theta(\beta))).$$

Principle 9.2.4 tells us that the reference of a complex (which, in the case of a sentence, will be its truth value) is a function of the *appropriate* reference of the part. Principle 9.2.5 tells us that if we replace a term by another having the same reference *appropriate for the context in which the term occurs*, then the reference of the complex will remain unchanged. Since propositions, and senses in general, are detachable, independently existing entities, we have no difficulty speaking about *the customary sense* of a sentence or, which comes to the same thing, about *the indirect reference* of a sentence. It follows that we also have no difficulty identifying Principles 9.2.4 and 9.2.5 as precisely the principles we want

to link the *appropriate* reference of the part with the reference of the complex.

### 9.3 Some Loose Ends

There are some loose ends in the story that we want to identify, even though we cannot tie them up neatly.

The singular term ‘Margaret Thatcher’ and the predicate ‘( ) drives a Peugeot’ are parts of sentence (9.1): the sentence is constructed by concatenating these items. The semantic story, however, is a bit more complicated. On the *Bedeutung* side, *driving a Peugeot* is supposed to be a function that maps Margaret Thatcher into a truth value. There is no implication that the reference of the part (namely, Margaret Thatcher), is a part of the reference of the complex (namely, the True).<sup>8</sup> How do matters stand at the level of *Sinn*? We have seen that the sense of the predicate combines with the sense of ‘Margaret Thatcher’ to form a thought. Is this combining function/argument combining? If so, there would be no reason to suppose that the sense of ‘Margaret Thatcher’ is part of the thought expressed by (9.1) any more than there is a reason to suppose that Margaret Thatcher is part of the True. And there would be no reason to suppose that the proposition expressed by (9.1) had a structure that mimicked the structure of the sentence. Frege, however, adopted the part/whole reading, so that the thought expressed is a *structured proposition*: the sense of the singular term and the sense of the predicate are both parts of the thought.<sup>9</sup>

A proposition is the *Bedeutung* of a sentence in a ‘that’ clause. Which story holds here – function/argument or part/whole? If the former, there is no assurance that the sentence refers inside a ‘that’ clause to the very same thing it customarily expresses. But this is surely wrong: the proposition referred to in (9.2), that is, the one Ted Kennedy believes, must be the very same as the one expressed by (9.1). The same story must be told each time. We have already decided that the analysis is part/whole at the level of sense, so it must be part/whole at the level of reference. But then we do not have the function applied to the argument to yield a value, which is the way *referring* is supposed to work.

We have noted the problem on several occasions. The problem is that Frege takes denoting a function to be acting predicatively. As a result, he just does not have any way of specifying a predicate without its actually acting predicatively in that context. In Chapter 5, we spoke of the puzzle about the concept *horse* and traced it to the fact that Frege took referring

to a concept to be the performing of a certain function. In Chapters 6 and 7, we contrasted Frege's view with that of Russell, whose small-scope construction permitted the occurrence of a predicate that was not acting predicatively. Kaplan (1989) identifies this as an important aspect of the sense/reference distinction: he speaks metaphorically of the function being "multiplied through" at the level of *Bedeutung*, but not at *Sinn*.

There is another, closely related, problem in Frege's treatment of indirect reference. If we can refer in one context to the same thing expressed in another, what does the distinction between referring to something and expressing it come to?

In the simplest case, (9.1), there is no mistaking the reference of 'Margaret Thatcher', that is, the woman, with the sense of the name. However you choose to take the sense, it is clearly a very different thing from the woman herself. We tend to distinguish referring and expressing by these different kinds of things. We have been following Frege's lead, and he, as we discussed back in Chapter 5, distinguished names and predicates by the things they referred to. But in the cases we are considering, we have a proposition each time, so we cannot account for the difference between referring and expressing by appealing to a difference in the kind of thing referred to or expressed. Can we appeal to a difference in the way in which the proposition is engaged, if we might put it that way? In the straightforward case, again, when we talk of the reference of (9.1), we have a function *applied to* an argument, yielding a value; when we talk of sense, on the other hand, a function is not applied to an argument, but is bracketed or exhibited inert as part of the thought. Perhaps, then, this is how the distinction is to be understood when applied to a proposition: to refer to it is to assign a function/argument analysis and to express it is to assign a part/whole analysis. This would certainly get us a long way toward correctly understanding what Frege should be saying in these contexts.

However, in the passage quoted at the beginning of this chapter, Frege says that in indirect speech we mean to talk about the sense of the expressions inside the 'that' clause. Now, as we noted in Section 8.4, it certainly does not ring true that when we assert (9.2), we mean to talk about the sense or the meaning of the name 'Margaret Thatcher'. We would sooner say that we mean to be speaking about Margaret Thatcher, the woman herself: she is the one we claim Ted Kennedy believes to drive a Peugeot. *Referring* in this context has been lifted far off its intuitive moorings and is almost entirely an internally defined, theoretical notion for Frege. So it is far from obvious *how* to make the distinction between referring to and

expressing a proposition. Indeed it is far from obvious whether there is such a distinction.<sup>10</sup>

#### 9.4 The Infinite Hierarchy

Frege says nothing in his published writings about how the story of indirect sense and reference is to be extended to doubly embedded sentences, and to even more deeply embedded sentences. But in a letter to Russell dated December 28, 1902, Frege says that in

The thought that all thoughts belonging to class M are true  
does not belong to class M. (9.3)

the second occurrence of 'M' has its customary reference, but the first occurrence of 'M', which is in the italicized part, has its indirect reference. He contrasts (9.3) with

The thought that the thought that all thoughts belonging to  
class M are true does not belong to class M. (9.4)

And he says:

Since 'M' has different meanings in its two occurrences in [(9.3)], there must also be a difference in the meanings of 'M' in [(9.4)]. It can be said that in the twice-underlined part it has an indirect meaning of the second degree, whereas in the once-underlined part it has an indirect meaning of the first degree. (Frege 1980: 154)<sup>11</sup>

Clearly, in this passage, Frege believes that the reference of the term when doubly embedded must be different from its reference when singly embedded.

The picture that emerges from Frege's remarks, as shown in Figure 9.3, is the one many philosophers thought Frege held. In fact Parsons (1981) calls it the *Orthodox* view. Each term  $t$  has two sequences,  $S_t$  and  $R_t$ ,  $s_i(t)$  refers to  $r_i(t)$ , and for  $i > 0$ ,  $r_i(t) = s_{i-1}(t)$ . Each time a term is embedded in another 'that' clause, it shifts its reference one notch up to an item distinct from any it refers to in less deeply embedded contexts. So, whereas in (9.2), the singly embedded 'Margaret Thatcher' stands for its indirect

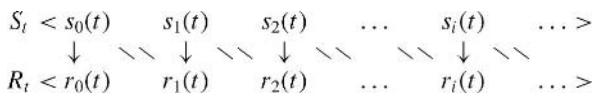


FIGURE 9.3

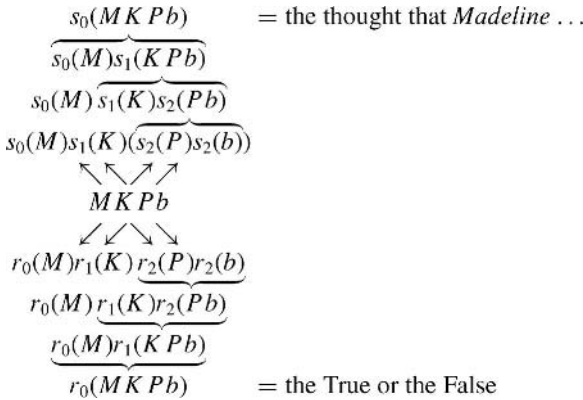


FIGURE 9.4

reference  $r_1$  ('Margaret Thatcher'), in

Madeline Albright said that Ted Kennedy believes that  
 Margaret Thatcher drives a Peugeot, (9.5)

the doubly embedded 'Margaret Thatcher' must stand for the distinct indirect indirect reference,  $r_2$  ('Margaret Thatcher'). Let us abbreviate 'Madeline Albright said that ( )' to 'M'. Then Frege's semantic analysis of (9.5) is given in Figure 9.4.

Carnap (1947: 130) objected to Frege's story. He complained that

Frege's method leads . . . to an infinite number of entities of new and unfamiliar kinds; and, if we wish to be able to speak about all of them, the language must contain an infinite number of names for these entities.

The proof he offered, as a number of commentators have pointed out, is faulty.<sup>12</sup> As a matter of fact, Frege's argument is also faulty, and for roughly the same reason. The problem with the argument is this. Frege fails to establish a distinct indirect reference of the second degree. Let us rewrite (9.3) and (9.4) as

The thought that all thoughts belonging to class  $M_1$  are true  
does not belong to class  $M_2$ , (9.6)

and

The thought that the thought that all thoughts belonging  
to class  $M_3$  are true does not belong to class  $M_4$ , (9.7)

respectively. Now Frege says that the two occurrences of 'M' in (9.6) cannot refer to the same thing, that is,

$$M_1 \neq M_2, \quad (9.8)$$

and also that the two occurrences of 'M' in (9.7) cannot refer to the same thing, that is,

$$M_3 \neq M_4. \quad (9.9)$$

Furthermore,

$$M_2 \neq M_4, \quad (9.10)$$

since one is extensional and the other not. But it does not immediately follow, as Frege seems to think, that

$$M_1 \neq M_3. \quad (9.11)$$

To make that inference, he appears to require that no sense can refer to itself, that is, for  $i = 0$  or  $i = 1$ ,

PRINCIPLE 9.4.1 (NO SELF-REFERENCE)  $s_i(t) \neq r_i(t)$ .

With Principle 9.4.1, the indirect reference (that is, the customary sense) of an expression would have to be distinct from its indirect sense, since, for Frege, the customary sense is what the indirect sense presents. To derive the hierarchy, Frege actually requires something stronger than Principle 9.4.1 to fill the gap in the argument, namely a hereditary No Self-Reference Principle, that is,  $s_i(t) \neq r_j(t)$  for every  $j \leq i$ . But the simpler statement of the assumption serves our expository purposes better.<sup>13</sup>

It is not clear whether Frege actually held anything like Principle 9.4.1. The strongest evidence that he did is in the argument we just examined from the letter to Russell. He appears to be assuming a principle of this sort to justify the distinction he wants. It is also worth noting that a No Self-Reference Principle would be in harmony with his rejection of the *direct reference* view we spoke about in Chapters 3 and 4. On the other hand, the very special relation one has to the sense of an expression, the epistemological grasp of the sense, which is quite similar to the immediate acquaintance that informed Russell's (1917) understanding of our grasp of a proposition, would appear to speak against this principle. Needless to say, those who favor the nonrigid hierarchy will be more inclined to attribute Principle 9.4.1 to Frege, while those who favor the rigid hierarchy will be less inclined to attribute Principle 9.4.1 to Frege.

Insofar as Dummett regards the indirect sense as standing for itself, his proposal squarely places him in the latter group.

Carnap's objection, unfortunately, sent philosophers down a strange, and irrelevant, path. The task was set: first, to find out which Fregean principles would generate such an infinite hierarchy – this became a problem because Carnap's argument for the hierarchy failed. See Linsky (1971), Forbes (1987). And second, to find out which principles could be safely jettisoned so as to avoid the infinite hierarchy. See Forbes (1987). The problem was seen as a straightforward cardinality problem: too many senses (in fact, infinitely many), and too many names of them (in fact, infinitely many). But why is this a problem? And why should this be a problem for Carnap? Carnap (1947), admitted infinitely many intensions into his system; and he certainly included enough names to talk about all of them. So why should *he* be disturbed about them?

The issue of the infinite hierarchy Carnap raised is not one of how many senses or how many names we need. The issue is whether there is some regular – perhaps algorithmic – way of determining the sense of an expression when embedded  $i$  times in 'that' clauses. Even if it turned out that there were infinitely many indirect senses required, we would not be particularly troubled if we had some rule for computing the  $i$ -th element of the sequence  $S_i$ . Given  $s_0(t)$ , the customary sense of  $t$ , we want to be able to determine  $s_1(t)$ , the indirect sense of  $t$ , and then to determine  $s_2(t)$ , the doubly indirect sense of  $t$ , and so on.

This is the problem Russell (1905) saw: given the customary sense of a term, compute its indirect sense. How can this be done? The customary sense is supposed to be what the indirect sense refers to. Therefore, we are being asked to compute the sense from the reference: we are being asked to forge a backward road from denotation to meaning. Not only must it be a unique sense (for it is a function), but somehow we should be able to figure out what this sense is simply by examining its referent. Now, if we accept the idea that there is no backward road, then we have no regular way of figuring out the items in the sequence  $S_i$ : we have, in effect, infinitely many semantic primitives, as Davidson would say. This is not a comfortable position for a semantic theory to be in. The obvious way of getting around this problem is to hold that the customary sense of a term, so to speak, self-presents itself. This is the route Dummett takes.

### 9.5 Collapsing the Hierarchy

In our characterization of the sequence  $S_i$  in Figure 9.3, we assumed that  $s_i$  was unique. We were simply following Frege's lead when he spoke of



*the* indirect sense of a term. Is this assumption justified? Could there be more than one sense for a given sense? There would seem to be strong reasons for thinking so. Surely a sense can be presented in different ways, just as any object, say, the Evening Star, can be.

Consider, for example, the following two identities:

The proposition that Giorgione was so-called because of  
his size = the proposition that Giorgione was so-called  
because of his size (9.12)

and

The proposition that Giorgione was so-called because of  
his size = the proposition that is expressed by sentence  
(2) in Quine's "Reference and Modality." (9.13)

These two clearly differ in sense, and the reason is that we have picked out one and the same proposition in two different ways: first, in a way that looks very much directly referential, that is, as *the proposition that Giorgione was so-called because of his size*, and second, somewhat indirectly, as *the proposition expressed by sentence (2) in Quine's "Reference and Modality."* We have two different ways of thinking about a proposition. But the clear Fregean intuition that senses, like other objects, can be presented in different ways, cannot require us to reject talk of *the* indirect sense of a term. If this notion of *the* indirect sense of a term is to work, then even though there might be more than one sense of a given sense, there cannot be more than one indirect sense: where we have the same customary sense, we must have the same indirect sense.

Dummett (1981*b*) presents an argument in which this understanding of indirect sense is defended. Let us suppose that 'is similar to' has the same customary sense as 'resembles', so that we can substitute one for the other *salva veritate* when singly embedded in a 'that' clause. Thus,

Barry thinks that Harvard is similar to Oxford, (9.14)

and

Barry thinks that Harvard resembles Oxford, (9.15)

must have the same truth value. If these two expressions were nonetheless to differ in indirect sense, then

Ayrton knows that Barry thinks that Harvard is  
similar to Oxford, (9.16)

and

Ayrton knows that Barry thinks that Harvard  
resembles Oxford, (9.17)

need not have the same truth value. “But this,” Dummett (1981*b*: 92) claims, “is contrary to intuition.”

The only case in which it might seem plausible to say that Ayrton knew that Barry thought that Harvard resembled Oxford, but did not know that he thought they were similar, is that in which Ayrton is ignorant of, or mistaken about, the sense of the word ‘similar’: but, if we admit this as a legitimate counter-example, then we likewise ought to deny that it follows from Barry’s thinking that Harvard resembles Oxford that he thinks they are similar; and, if we deny this, we reject Frege’s whole theory of senses as indirect referents.

That is, whatever reason we had for supposing that the two terms differed in indirect sense would equally be a reason for supposing that they differed in customary sense; so if two terms have the same customary sense, they must also the same indirect sense.

There is a very clear disanalogy being claimed between this case and the Evening Star/Morning Star case. Surely one can know what ‘the Evening Star’ refers to and know what ‘the Morning Star’ refers to without knowing that they are the very same thing. On the other hand, if one knows the sense of a word  $\alpha$  and one knows the sense of a word  $\beta$ , then if it is the same sense, one cannot help but know that it is the same. The slippage claimed between the reference and sense of ‘the Evening Star’ cannot be found between the sense and indirect sense of ‘the Evening Star’. And this means that unlike the Evening Star, we appear to have direct or immediate acquaintance with the sense of ‘the Evening Star’. Contrast

the proposition that Giorgione was so-called  
because of his size (9.18)

with

the proposition that is expressed by sentence (2) in  
Quine’s “Reference and Modality.” (9.19)

(9.18) is not simply a rigid designator,<sup>14</sup> picking out the same object, a particular proposition, in every possible world. For (9.18) does not embody a rigid *description* of the proposition. It seems, rather, to be *directly referential*. This is the direction Dummett’s suggestion is leading us: that in these indirect contexts, the senses of expressions are directly referred to, that the senses present themselves to us unmediated.

That is how the argument leads Dummett to the view that the customary sense of a term is the same as its indirect sense. For consider the argument to show that if (9.14) and (9.15) have the same truth value, then (9.16) and (9.17) must have the same truth value. We could only make sense of the claim that (9.16) is true while (9.17) is false by supposing that Ayrton was mistaken about the sense of the word ‘similar’. That is, we supposed that he did not realize that

the indirect sense of ‘similar’ = the indirect sense  
of ‘resembles’. (9.20)

Now, if (9.20) is true,

the indirect reference of ‘similar’ = the indirect  
reference of ‘resembles’ (9.21)

must be true; and if (9.21) is true,

the customary sense of ‘similar’ = the customary  
sense of ‘resembles’ (9.22)

must be true. So we could not see how, once we supposed he did not realize that (9.20) was true, he could fail to realize it without failing to realize that (9.22) was true. There is no way we could attribute the error to (9.20) without attributing it to (9.22) So there was no way we could distinguish the contribution of the higher level sense

the indirect sense of ‘similar’ (9.23)

from the lower level sense

the sense of ‘similar’. (9.24)

The very same reasons that led Dummett to require that (9.20) was true when (9.22) was true also lead him to require that (9.23) and (9.24) are the same thing.

Here are the details of this semantic analysis for our original sentence, (9.1). (9.1) expresses its customary sense and refers to its customary reference. When embedded in (9.2), however, it refers to its customary sense. ‘Ted Kennedy believes that ( )’ is not embedded, so it refers to its customary reference, a function that maps the customary sense of (9.1), that is, the thought it expresses, into a truth value. When (9.2) itself gets embedded, as in (9.5), the doubly embedded (9.1) does not change reference, but the singly embedded ‘Ted Kennedy believes that ( )’ is shifted to refer to its customary sense, and ‘Madeline Albright said that



point. Both sentences are well formed and meaningful – meaningful even for Frege, since on numerous occasions, as we documented in Section 3.6, he explicitly distinguished senses from other things, like ideas.

Let us now try to reconstruct Russell's argument. Each of the following is true and unproblematic.

'The center of mass of the solar system is a point' expresses the proposition that the center of mass of the solar system is a point. (9.27)

'The sense of "the center of mass of the solar system" is a point' expresses the proposition that the sense of "the center of mass of the solar system" is a point. (9.28)

Now since, in an indirect context, a term shifts its reference to its sense, in

The proposition that the center of mass of the solar system is a point, (9.29)

the expression 'the center of mass of the solar system' refers to its sense, namely,

the sense of 'the center of mass of the solar system', (9.30)

and in

the proposition that the sense of 'the center of mass of the solar system' is a point, (9.31)

the expression 'the sense of "the center of mass of the solar system"' refers to its sense, namely,

the sense of 'the sense of "the center of mass of the solar system"'. (9.32)

Dummett's suggestion is that iterated senses collapse. Recall the passage quoted in Section 9.5. Dummett says:

There is therefore no reason to think that an expression occurring in double *oratio obliqua* has a sense or a reference different from that which it has in single *oratio obliqua*: its referent in double *oratio obliqua* will be the sense which it has in single *oratio obliqua*, which is the same as the sense it has in ordinary contexts, which is the same as its referent in single *oratio obliqua*.

So it would seem (9.30) and (9.32) are the same, and therefore the propositions (9.29) and (9.32) are the same. Let us codify Dummett's principle as

PRINCIPLE 9.6.1 ( $\Theta$  COLLAPSE)  $\Theta\Theta(\theta(\alpha)) = \Theta(\theta(\alpha))$ .

Together with Principles 9.2.1 and 9.2.3 identified earlier, however, Principle 9.6.1 will enable Russell's argument to be sustained.<sup>15</sup>

We assume

'the center of mass of the solar system is a point' expresses  
 $\Theta(\text{the center of mass of the solar system is a point})$ . (9.33)

From which it follows, by Principle 9.2.3, that

'the center of mass of the solar system is a point' expresses  
 $\Theta(\text{the center of mass of the solar system})\Theta(\text{is a point})$ . (9.34)

Next, we assume the truism

' $\Theta(\text{the center of mass of the solar system})$  is a point' expresses  
 $\Theta(\Theta(\text{the center of mass of the solar system}) \text{ is a point})$ . (9.35)

And, once again, we use Principle 9.2.3 to get

' $\Theta(\text{the center of mass of the solar system})$  is a point' expresses  
 $\Theta\Theta(\text{the center of mass of the solar system}) \Theta(\text{is a point})$ . (9.36)

Now, from Principle 9.6.1, we have

$\Theta(\text{the center of mass of the solar system}) = \Theta\Theta(\text{the center of mass of the solar system})$ . (9.37)

So,

$\Theta(\text{the center of mass of the solar system is a point}) = \Theta(\Theta(\text{the center of mass of the solar system}) \text{ is a point})$ . (9.38)

It therefore follows, since these are the very same proposition, the two sentences

the center of mass of the solar system is a point (9.39)

and

$\Theta(\text{the center of mass of the solar system})$  is a point (9.40)

express the very same proposition.<sup>16</sup>

Russell is vindicated!

## 9.7 Reflections on the Argument

The Principle of  $\Theta$  Collapse 9.6.1 says that, no matter how deeply embedded in ‘that’ clauses, the sense of an expression – and therefore its reference – remains the same. Dummett suggested this principle to address Russell’s observation that there is no backward road from reference to sense. But apparently it collapses the sense/reference distinction, just as Russell said would happen if one tried to forge such a road.

Let us simplify Figure 9.3, where we introduced the relation between,  $S_t$ , the sequence of senses, and  $R_t$ , the sequence of references, for a term  $t$ . Let us suppose that we have only two items in each sequence, the customary and the indirect sense or reference. Now, as we look at Figure 9.6, we see that reference crops up twice, once in the sequence  $R_t$ , and once again in the relation between the items in the sequence  $S_t$  and the items in the sequence  $R_t$  that we represented by the downward arrow. Furthermore we see a relativized notion in the sequence  $R_t$ , and an unrelativized notion represented by the arrows. Figure 9.6 is obviously misleading. The two referring relations must be reduced to one. There are two ways of doing this.

One way is to take the indirect reference of  $t$ ,  $r_1(t)$ , to be the customary reference of the indirect sense of  $t$ ,  $r_0(s_1(t))$ . This reduces the two referring relations to the arrow, as shown in Figure 9.7. The other is to reduce the two referring relations to the items in the lower sequence, as shown in Figure 9.8.

On the first view, that given in Figure 9.7, we are supposing that the customary reference of the indirect sense of  $t$  is the indirect reference. The sense, speaking loosely, incorporates the context of occurrence. But then the customary sense and the indirect sense of  $t$  must be distinct so that the customary reference and the indirect reference of  $t$  are distinct. This seems to be the picture Frege had in mind in the passage we

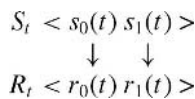


FIGURE 9.6

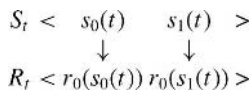


FIGURE 9.7

$$\begin{array}{ccc}
 S_t < s_0(t) & s_0(t) > \\
 & \downarrow & \downarrow \\
 R_t < r_0(s_0(t)) & r_1(s_0(t)) >
 \end{array}$$

FIGURE 9.8

examined in Section 9.4. This is the Orthodox picture, the one that leads to infinitely many distinct senses. It is this *context-free* reference pictured in Figure 9.7 that Dummett is rejecting with his suggestion.

In Figure 9.8, on the other hand, the indirect reference of  $t$  is not the customary reference of the indirect sense of  $t$ . It is the indirect reference of  $t$ , that is, what  $t$  refers to *in a certain context*.  $r_1(s_0(t))$  refers to the indirect reference *only in an indirect context*. Indirect reference simply sets up a context in which a word's function has been shifted. We cannot *say* what the sense determines in an indirect context. We can only use it in that context. So Figure 9.8 depicts a *context sensitive* reference, and this is the one Dummett favors.

If we could refer to the indirect reference in a nonembedded context, then

$$r_1(s_0(t)) = r_0(s_1(t))$$

would be true and the distinction between sense and reference collapses. That is exactly what Principle 9.2.1 permits us to do. This is how we collapsed the distinction in the previous section. Dummett was arguing for the second of the two readings, but, in doing so, he required speaking of these senses as if the first of the two readings were available. In opting for the picture in Figure 9.8, Dummett is opting for the view that the way in which the item is picked out is intimately connected with what it is that one is picking out: the sense of an expression is being picked out indirectly. But the argument that he used to identify customary and indirect sense required that he be able to identify the indirect sense from outside the indirect context. Each option is coherent on its own. The problem arises when both are chosen. A more drastic overhaul of the theory is needed, either to prevent *the sense of  $\eta$*  from being seen as a function, or, alternatively, from understanding the expressions inside the context from referring to something.



## Through the Quotation Marks

### 10.1 Introduction

We usually use language to speak about things other than itself. Of course, we can use language to speak about itself. The single-quote construction was devised for just this purpose, to render such speaking error free. The construction has the important characteristic that the very expression named is woven into the fabric of the sentence used to speak about it. To forestall errors of ambiguity, then, the construction takes upon itself the onus of ambiguity to assure clarity elsewhere.

The single-quote convention is to enclose a word or phrase in single quotation marks when we wish to speak about it. The convention was introduced into modern practice originally by Frege (1892*c*: 153–4), whose strategy was to treat direct quotation and indirect quotation in a parallel manner:

If words are used in the ordinary way, what one intends to speak of is their *Bedeutung*. It can also happen, however, that one wishes to talk about the words themselves or their sense. This happens, for instance, when the words of another are quoted. One's own words then first designate [*bedeuten*] words of the other speaker, and only the latter have their usual *Bedeutung*. We then have signs of signs. In writing, the words are in this case enclosed in quotation marks. Accordingly, a word standing between quotation marks must not be taken as having its ordinary *Bedeutung*.

This passage immediately precedes the one with which we opened Chapter 9 and introduced the notions of customary and indirect reference.

The ordinary use of quotation marks, the one we find in works of fiction, works of nonfiction, and newspapers, is to reproduce another's words, and to indicate that the words being reproduced are the very words used by the person under consideration. One is not, contrary to what Frege says, naming or designating the words. Let us introduce the expression *O-quotation* for ordinary quotation and *P-quotation* for the philosopher's quotation.

The difference between the two conventions is most evident when we consider iterations of quotation marks. We might, on the ordinary use, quote John's statement about Harold's statement:

John said, "Harold said, 'I accept the conditions of  
the agreement.'"(10.1)

In (10.1), we repeat John's exact words, and carry along his intention to repeat, not just paraphrase, Harold's exact words. Now contrast sentence (10.1) with

John said "I accept the conditions of the agreement.'" (10.2)

Sentence (10.2) is a paradigm example of an iteration of quotation marks on the P-quotation convention. But it has no analogue in O-quotation. We might try to read (10.2) in such a way that John is supposed to be quoting another's words, whose words, in turn, involve a specific quotation. But, as O-quotation, John is understood to be quoting the original's words directly, and only a single pair of quotation marks is really needed. With P-quotation, we can distribute outer quotation marks over inner quotation marks. Not so for O-quotation. Let us try distributing the outer quotation marks over the inner quotation marks (with '+' standing for concatenation) in (10.1). We get

John said, "Harold said," + "I accept the conditions  
of the agreement."(10.3)

But then the doubly embedded quotation very clearly collapses into a singly embedded one:

John said, "Harold said," + "I accept the conditions  
of the agreement."(10.4)

With P-quotation, the different levels of quotation count in a way they do not count for O-quotation; for when we enclose an expression in quotation marks, *we create a new name*, which can, in turn, be referred to by enclosing *it* in quotation marks. But with O-quotation, the quoted

expressions are not named. It is important to underscore how utterly alien this iteration of quotation marks is to the ordinary practice of quotation.

Now, *oratio recta* constructions, understood as O-quotations, are analogous to *oratio obliqua* constructions. Corresponding to the quotation marks in the *oratio recta* construction is the ‘that’ operator of the *oratio obliqua* construction, and on this understanding Dummett (1981*b*) appears to be exactly correct in suggesting that the iterations collapse in accordance with Principle 9.6.1. But we must be clear that on this reading, there is no referring, in the sense of naming, the object inside the appropriate clause. And so there is no functionally generated complex name to be accounted for by the Fregean function/argument algebra. The quotation marks are contextual markers, not function-expressions. Two contexts are demarcated: inside quotation marks and outside quotation marks. The same words occur in each context, but with different functions. There is no meaning attached to the iteration of contextual markers, and so there is no infinite set of semantic primitives to worry about.

Frege was correct to draw the analogy between *oratio obliqua* and *oratio recta* constructions, but he was wrong to suppose that in either case the item inside the stated context was named. Dummett’s  $\Theta$  Collapse Principle 9.6.1 is clearly plausible by analogy with O-quotation, but nothing is being named on this interpretation, so Frege’s view that the sense of a name in an indirect context stands for something is unwarranted.

On the other hand, the Principle of  $\Theta$  Collapse 9.6.1 is totally implausible if we are dealing with P-quotation. For in the case of P-quotation, the iterations very clearly do matter. We would never take as a principle of the single-quote convention that ‘‘Boston’’ is the same as ‘Boston’; and we would never, by analogy, have accepted as a principle of ‘that’ clauses that  $\Theta\Theta$  (*Boston*) must be the same as  $\Theta$  (*Boston*). O-quotation requires them to be the same; P-quotation requires them to be different. It was the inability to distinguish these two that led to the collapse of the sense/reference distinction in the last chapter.

In this chapter, we are going to look at P-quotation. We will argue that although there are infinitely many names generated, there are not infinitely many primitives to learn. The juncture is between an ordinary context and a quotation context, and there is no further problem with more deeply quoted contexts because, once we enter the quotation context, the reference of a quotation name *is* a function of the quoted interior. We will demonstrate this in Section 10.5, and speculate in Section 10.6 about the significance of context.

But before we reach that point, we will canvass some views about the single-quote convention. Frege's analysis of these contexts has had an enormous influence on philosophers seeking an understanding of the single-quote convention. Over and over again, however, the confusion between the two types of quotation has got in the way. As a result, a device that has been so easy to use, and so efficacious in facilitating error-free communication, has turned out to be a philosophical black hole.

We begin in Section 10.2 with Quine's *Structureless Name Theory* of quotation. We continue in Section 10.3 with Davidson's *Demonstrative Theory* of quotation. In Section 10.4, we examine Parsons's *Identity Theory* of quotation, which, although much closer to the Fregean structure, still fails to distinguish the two types of quotation we just identified. In the final two sections, we defend what we take to be the correct reading of a Fregean single-quote convention.

### 10.2 Quine: Structureless Names

We readily distinguish an object from its name in everyday life. But in philosophy, where our subject matter is language itself, the chances of error and confusion increase. To guard against such error, Quine prescribes vigilance, and he endorses scrupulous adherence to the single-quote convention: whenever we wish to speak about a word, we should enclose that word in quotation marks.

In a passage generations of logic students have committed to memory from Chapter 1, Section 4 of *Mathematical Logic*, entitled "Use Versus Mention," Quine (1951: 23–4) explains the convention. Considering the three sentences,

Boston is populous, (10.5)

Boston is disyllabic, (10.6)

'Boston' is disyllabic, (10.7)

he says:

The first two are incompatible, and indeed (10.5) is true and (10.6) false. Boston is a city rather than a word, and whereas a city may be populous, only a word is disyllabic. To say that the place-name in question is disyllabic we must use, not that name itself, but a name of it. The name of a name or other expression is commonly formed by putting the named expression in single quotation marks; the whole, called a *quotation*, denotes its interior. This device is used in (10.7), which, like (10.5), is true. (10.7) contains a name of the disyllabic word in question, just as (10.5) contains a name of the populous city in question. (10.7) is *about* a word

which (10.5) *contains*; and (10.5) is about no word at all, but a city. In (10.5) the place-name is *used*, and in this way the city is *mentioned*; in (10.7) a quotation is used, and the place-name is mentioned. We mention  $x$  by using a name of  $x$ ; and a statement about  $x$  contains a name of  $x$ .<sup>1</sup>

It is quite unfortunate that Quine should have wrapped these two issues together: use/mention confusion, on the one hand, and the single-quote convention, on the other. The moral of his sermon is that scrupulous adherence to the single-quote convention prevents using and mentioning the same word. But, the quotation denotes its interior, as he says explicitly, so the very same word is both used and mentioned. The devil, it seems, is literally in the details.

Quine says, “To say that the place-name in question is disyllabic we must use, not that name itself, but a name of it . . .” This echoes a comment Quine (1951: 23) makes at the beginning of the section from which the passage is extracted, “. . . a statement about an object must contain a name of the object rather than the object itself.” In the same vein, Church (1956: 61–2) says that “a word enclosed in single quotation marks is to be treated as a different word.” Tarski (1944: 16) too sounds a similar note:

[T]he fundamental conventions regarding the use of any language require that in any utterance we make about an object it is the name of the object which must be employed, and not the object itself. In consequence, if we wish to say something about a sentence, for example, that it is true, we must use the name of the sentence, and not the sentence itself.

They all hold:

PRINCIPLE 10.2.1 (QUINE NO SELF-REFERENCE) *A name must be distinct from the object it names.*

But Principle 10.2.1 is very clearly and obviously false. The standard proof of the Henkin Completeness Theorem for First-Order Logic<sup>2</sup> takes the constants of the first-order theory to stand for themselves: name and object are identical. Tarski despaired of a consistent account of truth for a natural language because he saw no way to eliminate self-reference in natural languages. So it is puzzling Tarski should believe in a “fundamental convention” that name and object be distinct. In a natural language that includes demonstratives and indexicals, it is near impossible to eliminate self-reference. Nothing prevents an individual from saying

This is a demonstrative (10.8)

where the item referred to is the word ‘this’.<sup>3</sup>

But Principle 10.2.1 is also false of the single-quote device *as characterized by Quine*. Consider the sentence

“'” is a quotation mark. (10.9)

If (10.9) succeeds in saying anything true, then the named object is not distinct at all from the name, but a very clear and perceptible part of it. We do not mean the middle quote, the status of which is a matter of considerable dispute (that is, whether the quoted material is or is not part of the name), but the surrounding quotes, which are used in this case.

Returning to the passage quoted above, let us look to what Quine says about ‘about’: “(10.7) is *about* a word which (10.5) contains; and (10.5) is about no word at all, but a city. In (10.5) the place-name is *used*, and in this way the city is *mentioned*; in (10.7) a quotation is used, and the place-name is mentioned.” This is the same sort of talk that we found to be involved in the Church-Langford Test in Section 4.4. It is also the sort of talk responsible for the misconception that a word is either used or mentioned but not both.

Quine claims that (10.5) is about no word at all. But surely, ‘about’ cannot possibly take such heavy-handed application. Why isn’t (10.5) about the word? A foreigner who is just learning English asks: “What does ‘Boston’ mean?” (10.5) would be an appropriate answer. Replying in this way, we speak to the question about the word, and in that sense, we speak about the word *as well as* the city. It is not so easy to say that (10.7) is about the city as well as the word. But that is not because of the expression used to designate the word. Rather, it is because of the predicate attributed to it: *being disyllabic* is not a feature of the city. In any event, given the looseness with which we use the word ‘about’, it does not seem correct to say that (10.7) is about the word and not the city, and it does not seem correct to say that (10.5) is about the city and not the word. Saying this, however, we do not deny that in (10.7) it is the word, not the city, that is said to be disyllabic, nor do we deny that in (10.5) it is the city, not the word, that is said to be populous.

Quine’s (1951: 26) argument that quotation-names have no significant logical structure has a distinctly Fregean cast:<sup>4</sup>

The meaning of the whole [quotation-name] does not depend upon the meanings of the constituent words. The personal name buried within the first word of the statement:

‘Cicero’ has six letters, (10.10)

e.g., is logically no more germane to the statement than is the verb ‘let’ which is buried within the last word. Otherwise, indeed, the identity of Tully with Cicero would allow us to interchange these personal names, in the context of quotation marks as in any other context; we could thus argue from the truth (10.10) to the falsehood:

‘Tully’ has six letters. (10.11)

In fact, this is just the application of the function/argument analysis we discussed in Section 2.5 as a paradigm of a Frege-style analysis.

Quine appeals to Frege’s Principles 2.3.1 and 2.3.3, and follows the procedure described in Section 2.3 for assigning a function/argument structure to the quotation-name. Because substitution does not preserve reference, Quine is unwilling to suppose that the quotation-name is a functionally constructed complex expression.<sup>5</sup> The two sentences (10.10) and (10.11) differ in truth value despite the truth of

Tully = Cicero. (10.12)

Principles 2.3.1 and 2.3.3 require that the name ‘Cicero’ that appears in (10.12) does not appear in (10.10). The default role for ‘Cicero’ in English is to refer to the man. It cannot play that role when placed in quotation marks, as in “Cicero.” So, according to Quine, that *name* is not occurring there. Quine (1953*b*: 169) terms the occurrence of ‘nine’ in “nine” an *orthographic accident*, just like the occurrence of ‘cat’ in ‘cattle’, and for this reason his theory is known as the *Structureless Name Theory*. A quotation-name turns out to be an unanalyzable, simple name.

Quine acknowledges a certain “anomalous feature” to his account of quotation-names. Although the semantic interpretation of a quotation-name is that it “denotes its [quoted] interior” (Quine 1951: 23), “from the standpoint of logical analysis each whole quotation must be regarded as a single word or sign, whose parts count for no more than serifs or syllables” (Quine 1951: 26). But one man’s anomaly is another man’s inconsistency: on the one hand, Quine claims that the quotation-name denotes its quoted interior, but, on the other, he denies that the quotation-name contains the quoted name. Quine’s position is untenable.

Frege held that it was the same word, both inside and outside the quotation context, but performing different functions in each. Quine is quite right in pointing out that the reference of the quotation-name ‘Cicero’ is not a function of the reference of its interior; but he is certainly wrong to deny complexity to a quotation-name. Doing so just lands him in inconsistency. We must agree with Frege that the same name occurs both

inside and outside the quotation marks, so there must be other forms of complexity than the function/argument complexity already noted.

### 10.3 Davidson: Demonstrative Names

Davidson (1979) is highly critical of Quine's account of the single-quote convention – not because of the inconsistency we noted, but because of Quine's seeming inability to account for how we are able to recognize in some rulelike fashion the reference of a quotation name. Davidson is particularly troubled by the infinite hierarchy of semantically primitive names Quine's theory seems to entail. Like Quine, however, he assumes a context-free framework.

The key to Davidson's (1979: 90) treatment is to give up the assumption that "the quoted material [is] part of the semantically significant syntax of a sentence." He continues:

It is natural to assume that words that appear between the boundaries of a sentence are legitimate parts of the sentence; and in the case of quotations, we have agreed that the words within quotation marks help us to refer to those words. Yet what I propose is that those words within quotation marks are not, from a semantical point of view, part of the sentence at all. It is in fact confusing to speak of them as words. What appears in quotation marks is an *inscription*, not a shape, and what we need it for is to help refer to its shape. On my theory, which we may call the *demonstrative theory* of quotation, the inscription inside does not refer to anything at all, nor is it part of any expression that does. Rather it is the quotation marks that do all the referring, and they help refer to a shape by pointing out something that has it. On the demonstrative theory, neither the quotation as a whole (quotes plus filling) nor the filling alone is, except by accident, a singular term. The singular term is the quotation marks, which may be read 'the expression a token of which is here'. Or, to bring out the way in which picturing may now be said genuinely to be involved: 'the expression with the shape here pictured'.

Davidson's handling of quotation marks is on a par with his paratactic treatment of indirect quotation. Davidson (1968–9) holds

John said that it was raining (10.13)

is to be understood as

John said that. It is raining. (10.14)

where 'that' demonstrates the proposition expressed by the sentence following. Quotation marks serve much the same function as 'that', with the addition that the words serving as part of the demonstration are the



exact words used. So the direct quotation

John said “It is raining” (10.15)

is understood as

John *literally* said that. It is raining. (10.16)

Davidson says that the quotation demonstrates an abstract shape, a token of which is presented. This has much in common with Kaplan’s (1989) account of demonstratives. Kaplan holds that when a demonstrative like *this* or *that* is accompanied by a description, as in ‘*this* man on the corner’ or ‘*that* book you were reading last week’, the accompanying description is not part of the proposition expressed, but is part of what determines the content: it is syntactically, but not semantically, part of the sentence.

Davidson (1979: 85) presents his account as an offshoot of what he calls the *Picture Theory* of quotation:

The picture theory of quotation is reminiscent of Frege’s theory of opaque (what he called oblique) contexts such as those created by ‘necessarily’, ‘Jones believed that . . .’, ‘Galileo said that . . .’, and so on . . . [T]here is the striking similarity that in both cases some linguistic device is supposed to create a context within which words play new referential roles. This concept of a context that alters reference has never been properly explained, and Frege himself was leery of it: it certainly does not lend itself to direct treatment in a theory of truth. The trouble with the picture theory, as with Frege’s treatment of opaque contexts generally, is that the references attributed to words or expressions in their special contexts are not functions of their references in ordinary contexts, and so the special context-creating expressions (like quotation marks or the words ‘said that’) cannot be viewed as functional expressions.

But Davidson (1979: 85) believes the picture theory needs revision:

In quotation, what allows us to refer to a certain expression, which we may take to be an abstract shape, is the fact that we have before us on the page or in the air something that *has* that shape – a token, written or spoken. The picture theory suggests no way to bring an inscription or utterance into the picture. This could be done only by describing, naming, or pointing out the relevant *token*, and no machinery for this purpose has been introduced.

For Davidson, the machinery for pointing out the relevant token is, as we have seen, the quotation marks themselves.

On the *Demonstrative Theory*, a quotation-name denotes something *other* than its quoted interior, something that is pictured by the quoted interior, so that one does not literally *see* what is being denoted – what one literally

sees is an unmistakably true likeness of what is being denoted. Davidson's solution has two prongs:

- *First*, he marks the distinction between the quoted interior and the denotation of the quotation-name ontologically. What is denoted is a type, not a token, and what occurs inside quotation marks is a token, not a type. But because the one is a token of the type, it is a perfect picture of what it is the quotation-name refers to. Davidson combines the fact that the two are different with an explanation for why one enables us to pick out the other.
- *Second*, the ontological difference is supplemented by a functional difference: the token occurring in the quoted interior is not serving there in its ordinary way. It is demonstrated by the quotation marks, which thereby effectively lifts it out from among the semantically significant parts of the sentence.

Both of these points are problematic. First, his distinction between type and token is ill conceived. His account is of sentence types, and so all parts of the sentence are types. We need only iterate the quotation marks once to make this clear. Consider

‘Harry’’. (10.17)

(10.17) denotes an abstract type (or shape) which is itself a quotation-name, and whose quoted interior is just as abstract as the quotation-name of which it is a part. The denotation of that quotation-name, then, must be the very same thing as its own quoted interior. It is not an ontological distinction between types and tokens that he needs, but a distinction between a type and an occurrence of a type. That is, he is marking the contextual difference as an ontological difference. This is wrong. As a result, he overlooks completely the role of the quotation marks to shift the reference of the expression occurring inside. The quotation marks function for him solely as pointers. This is inadequate. The fact that the second sentence is pointed to in (10.16) does not thereby prevent it from serving its usual role. Davidson's model for P-quotation is O-quotation, and this, as we have mentioned on several occasions, is a fatal confusion.

Davidson requires, like Quine, that name and thing named be distinct. He rejects the claim that the item inside the quotation marks is the same as the item referred to. For if they were the same, there would be no need for the quoted interior to picture anything. Let us give the name

'John' to the quoted interior in

'Harry' (10.18)

and consider the claim

John = 'Harry'. (10.19)

The standard understanding of the single-quote convention – whatever one's explanation for how it works – is that (10.19) is true. Davidson, when all is said and done, does not appear to be able to say this.

Davidson's view is incoherent. His picturing account has collapsed; the work of the paratactic device turns out to involve essentially some notion of autonomous designation, precisely the view he sought to distinguish picturing from; and his unhappiness with taking quotation marks as context-definers, but only as pointers, leaves no explanation for the exhibiting of the token, which he dearly needs to make a plausible story. It is clear, then, that to make sense of the single-quote convention, we shall have to develop this notion of single quotes as context definers.

#### 10.4 Parsons: Fregean Names

Quine and Davidson deny that the very same name occurs both inside and outside the quotation marks. Frege, by contrast, holds just the opposite view. Parsons (1982: 317) assumes the Fregean perspective that "the same word may appear in many different contexts, and what it expresses or refers to depends on the context in which it is used." But what is the role of the quotation marks? Parsons offers two suggestions: (a) the quotation marks are function-expressions; (b) neither the quotation marks nor the quotation marks together with the expression inside constitutes a referring expression. We examine these two alternatives in turn.

A number of philosophers have championed a Frege-style theory of quotation. Reichenbach (1947: 335), for example, says that quotation marks "transform a sign into a name of that sign." That is, the quotation marks demarcate a context in which expressions occur autonomously, that is, to denote themselves:

The signs of signs constructed by means of quotes are of a very peculiar kind. In them the object is employed as its own sign, and the function of the quotes consists in indicating this unusual usage. We might introduce a similar usage for the names of other physical objects; thus we might, whenever we write something about sand, put some sand in the place otherwise occupied by the word 'sand'. In order to indicate that this is not an undesired sand spot on our paper, but a

part of our language and the name of sand, we should have to put quotes left and right of the sand spot. Unfortunately such a practice, although perhaps suitable for sand, would often lead to serious difficulties, for instance if we wanted to use this method for denoting lions and tigers. It is for these technical reasons that the quotes method is restricted to the introduction of signs of signs. (Reichenbach 1947: 10)

Searle (1968: 76) argues that in a quotation “the word itself is *presented* and then talked about, and that it is to be taken as presented rather than used conventionally to refer is indicated by the quotes.”<sup>6</sup> Kaplan (1969: 119–21) also advanced a Frege-style view of the convention. This appears to coincide with the first of Parsons’s suggested readings: the quotation marks serve as *Funktionsnamen*, designating a function that maps the interior into a name of that interior. This has come to be known as the *Identity Theory* of quotation.<sup>7</sup>

On the other reading, the quotation marks serve as a form of punctuation, not as *Funktionsnamen* at all: the quotation marks plus quoted interior form a syntactical fragment (not a singular term). Both of these readings are wrong, we will argue. Parsons’s (1982) analysis suffers from a confusion of O-quotation and P-quotation. The first reading does not admit of iterations of quotation marks, and so it is an inadequate account of P-quotation. The second reading is a coherent theory of P-quotation, but in that case the quotation marks are introduced, if we may so put it, as a letter of the alphabet: they combine with the quoted expression to form a name of that expression. This is the position we think most plausible, namely, that P-quotation should be understood along the lines of what is known as the *Spelling Theory* of quotation,<sup>8</sup> where the quotation marks are not function-expressions at all. This will be formalized in the next section.

#### *Quotation Marks Are Names*

Parsons invites us to consider the sentence

Mary said ‘Joan will win’. (10.20)

‘M’ abbreviates ‘Mary said’, ‘J’ abbreviates ‘Joan will win’, and ‘Q’ abbreviates the quotation marks around the contained sentence ‘J’, so the sentence (10.20) is symbolized as ‘MQJ’. In addition, ‘ $r(\eta)$ ’ abbreviates ‘the customary reference of  $\eta$ ’, ‘ $s(\eta)$ ’ abbreviates ‘the customary sense of  $\eta$ ’, and ‘ $s_q(\eta)$ ’ abbreviates what Parsons calls ‘the quotational sense of  $\eta$ ’. Assuming the quotation marks to be function-names, Parsons (1982: 318) depicts the semantic analysis of (10.20) in Figure 10.1.<sup>9</sup>

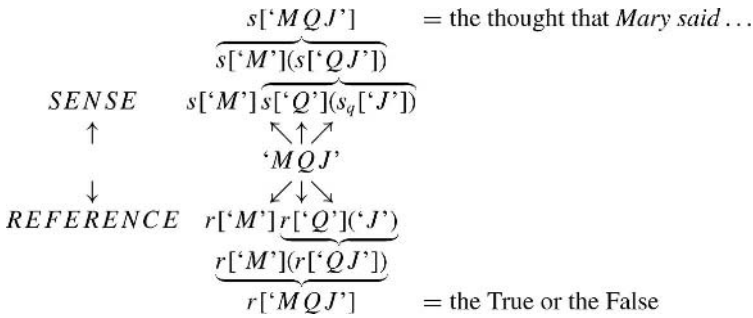


FIGURE 10.1

Here is Parsons’s (1982: 319–20) discussion of Figure 10.1:

We know that the whole sentence refers to a truth-value, i.e.  $r[MQJ]$  is a truth-value, and we know that it’s whatever truth-value  $r[M]$  maps  $r[QJ]$  to. But which function is  $r[M]$ , and which object is  $r[QJ]$ ? I know of no argument that *forces* any given answer on us here; on the other hand I can only think of one plausible and simple answer: that  $r[M]$  is the function which maps those sentences that Mary said to the True and everything else to the False, and that  $r[QJ]$  is the sentence “Joan will win”. That is, “Mary said” refers to the concept whose extension includes exactly what Mary said, and the quotation-mark name:

“Joan will win”

refers to the sentence enclosed within the quotes.

The diagram shows that  $r[Q](J) = r[QJ]$ , and we have just supposed that  $r[QJ] = J$ . So we have that  $r[Q](J) = J$ . That is  $r[Q]$  maps  $J$  to itself. But this is not a special case; the sentence diagrammed was chosen to be indicative of how all such sentences will work. So in all cases  $r[Q]$  will map any given sentence or phrase to itself. In short, the customary reference of ‘Q’ is that function which maps each and every phrase of the language to itself; it’s the identity map.

In sum, the quotation marks plus quoted interior ‘QJ’ is a complex name: the quotation marks ‘Q’ stand for a function, the quoted interior ‘J’ stands for an argument, and the complex ‘QJ’ stands for the value of that function for that argument. This is classic Fregean syntax.<sup>10</sup>

But Parsons’s claim that the customary reference of ‘Q’ is the identity map defined over the set of phrases in a language is incorrect. Let  $\mathcal{P}$  be the set of phrases of English. Then the identity map will be the function,  $f_{=} : \mathcal{P} \rightarrow \mathcal{P}$  such that for any  $x$  in  $\mathcal{P}$ ,  $f_{=}(x) = (x)$ . Let ‘I’ stand for this identity function, that is,  $r[I] = f_{=}$ . So these are both true:

$$\begin{array}{l}
 I(\text{‘Darkness tolls the knell of fading day’}) = \text{‘Darkness tolls} \\
 \qquad \qquad \qquad \text{the knell of fading day’} \qquad \qquad \qquad (10.21)
 \end{array}$$

and

$$I(\text{the first line of Gray's Elegy}) = \text{'Darkness tolls the knell of fading day'}. \quad (10.22)$$

Since

$$\text{The first line of Gray's elegy} = \text{'Darkness tolls the knell of fading day'}, \quad (10.23)$$

the function must map them both into the same thing. But Parsons's Q-function is supposed to work like this:

$$Q(\text{'Darkness tolls the knell of fading day'}) = \text{'Darkness tolls the knell of fading day''} \quad (10.24)$$

and

$$Q(\text{The first line of Gray's Elegy}) = \text{'The first line of Gray's Elegy'}. \quad (10.25)$$

The results are quite different. Q and I are, if functions, different functions.

The error in this argument is immediately visible in Figure 10.1: there is nothing that captures the role of the quotation marks as context definers. Parsons *assumes* that the sentence occurring inside quotation marks is standing for itself. But it is the quotation marks that are supposed to *make* the item inside quotation marks shift its reference.

In Figure 10.1, we find an arrow downward from the furthest right letter in 'MQJ' to '(J)'. The arrow carries us to the reference of the letter: 'J' refers to 'J'. But this is too quick: the arrow should point to  $r['J']$ , that is, the customary reference of the letter, and

$$r['J'] = 'J' \quad (10.26)$$

is false. But, of course, the quotation marks have shifted the reference. Perhaps we can introduce a new abbreviation, namely, ' $r_q[x]$ ', for 'the quotational reference of  $x$ '. This would be our notion of indirect reference that parallels the notion of indirect or quotational sense Parsons introduces. So what Parsons wants in the diagram is for the arrow down from 'J' to point to  $r_q['J']$ , for

$$r_q['J'] = 'J' \quad (10.27)$$

is true. But even with this notation, what is missing from the diagram is some indication that ‘Q’ is doing something that forces this shifting in the reference of ‘J’ from its customary reference to its quotational reference.

The problem comes to a head when we try to provide a semantic interpretation for iterated quotation marks. Consider the proper semantic interpretation of ‘QQJ’. The issue is whether we have iterated functions here or not.

ALTERNATIVE 1 We suppose that ‘J’ is referring to itself because it is flanked by quotation marks. The inner quotes stand for the identity function, so  $r[‘QJ’] = r[‘Q’](r_q[‘J’]) = r[‘Q’](‘J’) = ‘J’$ . Next, we operate on this by the outer quotes. ‘QJ’ is just standing for ‘J’, that is, ‘QJ’ = ‘J’. Thus, we find our ‘J’ standing for itself within the context of the outermost quotes, and so ‘QQJ’ = ‘QJ’ = ‘J’. In other words, the iterated quotes collapse.

ALTERNATIVE 2 Consider ‘QQJ’. What occurs inside the outermost quotation marks, since it occurs inside the quotation marks, has its quotation reference. We find that it is standing for  $r_q[‘QJ’]$ . Now this is just the expression ‘QJ’ itself. But what are we to make of the quotation marks in ‘QJ’. They are not serving there as function-expressions, since they are occurring inside quotes. This means that ‘QJ’ is not a complex at all. But, according to the claims originally made, ‘QJ’ is a complex. Thus, we are caught in a contradiction.

ALTERNATIVE 3 Consider ‘QQJ’. Since ‘QJ’ occurs inside the scope of the outermost quotation marks, it designates itself. And if it designates itself, the inner ‘Q’ is *not* serving as the identity function. The inner device is not, then, a complex name. It therefore follows that what ‘QQJ’ denotes is not a complex name, so it does not designate ‘QJ’, which is what Parsons claims it did.

These are the only alternatives we have open to us to understand the way in which iterated quotes work. Either the iterations collapse, giving the wrong answer, or if they do not collapse, they land us in a contradiction.

#### *Quotation Marks Are Not Names*

We next consider the possibility that the quotation marks are not names (either *Eigennamen* or *Funktionsnamen*). Parsons (1982: 320–1) says:

If quotation marks are not names, it is natural, then, to hold that quotation-mark names [such as “Damn”] are not names either, though they *contain* referring

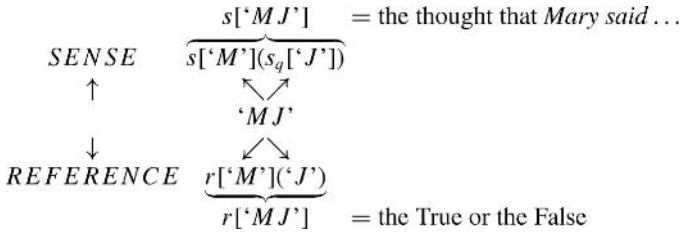


FIGURE 10.2

names as parts. (This is a curious reversal of Quine’s Theory.) In the semantic analysis of a sentence, then, we will ignore the quotation-mark names entirely; other functions will operate directly on their contents. For example, in analyzing “Mary said ‘Joan will win’” we will *not* use ‘Q’ at all. . . .

Parsons (1982: 321) pictures the situation as in Figure 10.2.

But, the idea that the quoted interior plus quotation marks is a fragment, of no more syntactic significance than

$$,Boston, \tag{10.28}$$

cannot be correct. For the quoted interior plus quotation marks invariably occupies a position that would otherwise be filled by a singular term. Could it be that the space in the sentence frame ‘\_\_\_ is a noun’ is replaceable either by a proper name, like ‘Boston’, or by a totally ungrammatical sentence fragment like (10.28)? This is implausible. If the sentence

$$“Boston” \text{ is a noun} \tag{10.29}$$

is *false*, then it is grammatically well formed; and if it is grammatically well formed, then it must be *true*, for the ability to fill precisely such sentence frames and result in grammatically well formed sentences is what constitutes our calling such an expression a noun. The *Eigenname/Funktionsname* distinction is made purely in terms of the way in which expressions go together to make up coherent, grammatically well formed wholes. Since ‘\_\_\_ is a noun’ is a *Funktionsname*, then, Fregean semantic theory demands that “Boston” be an *Eigenname*.

The diagram in Figure 10.2 misrepresents the sentence Parsons had originally given us. In the middle of the diagram, we find “MJ”, which abbreviates ‘Mary said Joan will win’. But the sentence we require an explanation for is “Mary said ‘Joan will win’.” No explanation for *this* sentence is given. We suggest, instead, Figure 10.3 as the proper interpretation. The quotation marks are not names, and this is shown by the fact that we are missing the intermediary line we have in Figure 10.1. Rather, the





quotes; on the other hand, the disabling of the inner quotes does not render *their* interior opaque to us.

### 10.5 A Formalism

Let  $\mathcal{V}$  be the vocabulary of a language  $\mathcal{L}$ , containing all the atoms out of which more complex expressions are built.  $\mathcal{V}$  includes the individual constants, variables, function symbols, predicate symbols, logical symbols, and punctuation marks, including the left quotation mark ( $lq$ ) and the right quotation mark ( $rq$ ).<sup>11</sup> Any finite sequence of elements of  $\mathcal{V}$  is an *expression* of  $\mathcal{L}$ . The set of quotation-names,  $\mathcal{E}^Q$ , is the smallest set of expressions generated by the following rule:<sup>12</sup>

**DEFINITION 10.5.1 (QUOTATION-NAME)** *For any expression  $e \in \mathcal{V}^n$ , the sequence  $\langle lq, e, rq \rangle$  is a quotation-name.*

This characterizes the set of quotation-names from the top down, as the smallest set generated using the recursive definition (10.5.1).

Alternatively, we can characterize the set from the bottom up. Let  $\mathcal{E}$  be the set of expressions that do not contain quotation-names. Each element  $e \in \mathcal{E}$  is a finite sequence,  $\langle e_1, e_2, \dots, e_n \rangle$ , with  $e_i \in \mathcal{V}$  for  $1 \leq i \leq n$ , and such that either  $e_1 \neq lq$  or  $e_n \neq rq$ . Let  $\mathcal{E}^{Q_1}$  be the set of expressions with a single pair of quotation marks around them: if  $e \in \mathcal{E}^{Q_1}$ , then  $e = \langle lq, e', rq \rangle$ , where  $e' \in \mathcal{E}$ . The set  $\mathcal{E}^{Q_1}$  contains only quotation-names formed with a single pair of quotes. But we can place a pair of single quotes around each of these to generate the elements of  $\mathcal{E}^{Q_2}$ . So ‘nine’  $\in \mathcal{E}$ , “nine”  $\in \mathcal{E}^{Q_1}$ , and “‘nine’”  $\in \mathcal{E}^{Q_2}$ . Clearly we can continue in this vein indefinitely to construct ever more complex quotation-names. Our set of quotation-names,  $\mathcal{E}^Q$ , will be the union of all these constructions:

$$\mathcal{E}^Q = \bigcup \mathcal{E}^{Q_i}, \quad \text{for } i \geq 1$$

The semantics of quotation-names appears equally well behaved. Informally, the semantics of quotation-names is usually expressed by saying that a quotation-name denotes its interior. A more formal statement of the semantic principle is:

**PRINCIPLE 10.5.1 (QUOTATION-NAME DENOTATION)** *For any expression  $e \in \mathcal{V}^n$ ,  $\langle lq, e, rq \rangle$  denotes  $e$ .*

Contrary to Russell’s (1905) famous admonition that “there is no backward road from denotation to meaning,” there is a 1–1 map  $\mathcal{Q}_1: \mathcal{E} \rightarrow \mathcal{E}^Q$  that associates each expression in  $\mathcal{E}$  with its quotation-name.

The inverse  $Q_1^{-1} : \mathcal{E}^{Q_1} \rightarrow \mathcal{E}$  is a denotation function: it maps each quotation-name into the expression it names. Again, there is a 1–1 map  $Q_2 : \mathcal{E}^{Q_1} \rightarrow \mathcal{E}^{Q_2}$  from single-quoted expressions to double-quoted expressions, and its inverse  $Q_2^{-1} : \mathcal{E}^{Q_2} \rightarrow \mathcal{E}^{Q_1}$ . There are, obviously, infinitely many of these functions, but we can unify them under one rubric. Let  $\mathcal{T}$  be the set of *terms* of  $\mathcal{L}$ .  $\mathcal{E}^Q \subset \mathcal{T}$  because each quotation-name is a *term*, even though the item it denotes might not be. We define a function

$$D_Q : \mathcal{E}^Q \rightarrow \mathcal{T}$$

from quotation-names to their denotata as follows. For any  $e \in \mathcal{E}^Q$ , there is some  $i$  such that  $e \in \mathcal{E}^{Q_i}$ , and so

$$D_Q(e) = Q_i^{-1}(e).$$

The  $Q_i$  admit composition: since

$$Q_2^{-1}(\text{‘‘nine’’}) = \text{‘nine’}$$

and

$$Q_1^{-1}(\text{‘nine’}) = \text{‘nine’},$$

we have

$$Q_1^{-1}(Q_2^{-1}(\text{‘‘nine’’})) = \text{‘nine’},$$

so  $D$  is iterable:

$$D_Q(D_Q(e)) = Q_{i-1}^{-1}(Q_i^{-1}(e)).$$

Definition 10.5.1 and the semantic interpretation for it laid down in Principle 10.5.1 clearly and adequately characterize the quotation construction for us. It provides a general method of name formation that carries just the semantic interpretation we want. The explanation given of the single-quote construction is so clear and coherent that it is a wonder anyone is puzzled by it.

## 10.6 Philosophical Remarks

Definition 10.5.1 is unproblematic. Quotation-names are characterized purely syntactically: an expression is a quotation-name if, and only if, it has a left quotation mark as the furthest left item in the sequence and a right quotation mark as the furthest right item in the sequence. Principle 10.5.1, however, requires some discussion.

'*lq*' and '*rq*' are names of very specific expressions, the left quotation mark and the right quotation mark, respectively. These expressions – left quotation mark and right quotation mark – are specific expressions in the language, but we have not specified which marks they are: 'left quotation mark' and 'right quotation mark' functionally determine the expressions, but they do not visually determine them. There are any number of ways to implement the two expressions. We might use left and right single quotes; we might use left and right double quotes; we might use left and right corner quotes; we might use left and right angle-braces. On the other hand, '*e*' is a variable that ranges over expressions. This is an instance of Principle 10.5.1:<sup>13</sup>

‘nine’ denotes ‘nine’. (10.31)

This, however, is *not* an instance of Principle 10.5.1:

‘nine’ denotes nine. (10.32)

Since '*e*' is a variable that ranges over expressions, it must be replaced by the name of an expression, not the name of a number; so 'nine' is not a substitution instance of the variable.

It is very important that we distinguish sentences like (10.31) from sentences like (10.32). The failure to do so leads to confusion and puzzlement about the single-quote convention. Sentence (10.31) is a truth of logic, or at least of logic supplemented by Principle 10.5.1. Sentence (10.32) is not a truth of logic. That the expression 'nine' *denotes*, let alone that it denotes a number and not a dog, cannot possibly be a matter of logic. That it does is entirely a matter of our conventional usage, supplemented by facts about the world.

So long as we speak about the denotation of quotation-names, there appears to be no difficulty; but if we try to link up an ordinary term (in  $\mathcal{T} - \mathcal{E}^Q$ ) with its denotation, there is no systematicity to the device.  $\mathcal{D}_Q$  is simply undefined in that case. The domain of  $\mathcal{D}_Q$  is the set of quotation-names,  $\mathcal{E}^Q$ . Its range is the set of terms,  $\mathcal{T}$ . In (10.32) we are taking an element of the set  $\mathcal{T} - \mathcal{E}^Q$  and mapping it into something that, so far, is entirely unspecified. But it is an illusion to suppose that the systematicity of Principle 10.5.2 gives this to us. We can augment  $\mathcal{D}_Q$  with a host of individual assignments,

1. 'one' denotes one
2. 'two' denotes two
3. 'three' denotes three

and so on. But we will have to make these assignments one by one; there is nothing systematic about this. The fact that the very same word occurs both inside and outside the quotation marks simply is lost. This augmentation of  $\mathcal{D}_Q$  is as unsystematic as if we were to make the following assignments:

1. Anthony denotes two
2. Charles denotes two
3. William denotes three

and so on.

So long as we insert the very same term at each of the lines in the frame,

‘ \_\_\_\_ ’ denotes \_\_\_\_ (10.33)

it would seem that we should come out with a truth. We have remarked on the obviousness of this fact on numerous occasions. Kaplan (1969: 125) actually speaks of instances of (10.33), like (10.32), as being “nearly analytic.” But certainly they cannot be analytic in the sense of replacing synonyms by synonyms to get a truth of logic, even when supplemented with Definition 10.5.1 and Principle 10.5.1. It is of course difficult to envisage a situation in which a denial of such a claim, for example,

‘Cicero’ does not denote Cicero, (10.34)

would come out true. But (10.34) seems not an analytic falsehood, but rather something more on the order of Moore’s pragmatically paradoxical

$p$  but I don’t believe it. (10.35)

This is indeed a puzzling situation. But these are puzzles that lie beyond the scope of this book.



## Appendix A

### Begriffsschrift in Modern Notation: (1) to (51)

Frege signed the preface to *Begriffsschrift* in December of 1878, and it was published the following year by George Olms. Frege did little to connect up his own work with his contemporaries, either with the logical achievements of Boole, or the mathematical investigations of Dedekind. The only explicit references are to philosophers – Aristotle, Leibniz, and Kant. His previous work, which consisted mainly of reviews, gave no indication of the direction and creativity of his thinking. Like Athena, emerging full-grown from Zeus’s brow, Frege’s remarkable work bore no evidence of the genesis and growth of the ideas presented therein. There is little surprise at the reception his contemporaries gave *Begriffsschrift*: they did not know what to make of it.

Here are some of the achievements of *Begriffsschrift*:

*First*, Frege synthesized the two otherwise opposed traditions – the Stoic logic of the propositional connectives and the Aristotelian treatment of the quantifiers – into one system, and extended the Aristotelian treatment to include relations as well as properties. His function/argument analysis of propositions supplanted the subject/predicate distinction of traditional analysis, creating one of the first extensions of mathematical forms of analysis to domains other than arithmetic and geometry.

*Second*, propositions (1), (2), (8), (28), (31), and (41), together with the Rule of Inference, Modus Ponens – and a suitable substitution rule that is employed but never precisely stated – constitute a complete and consistent axiomatization of truth-functional logic.<sup>1</sup>

*Third*, Frege provides two rules for the universal quantifier: Universal Instantiation is given in proposition (58), and Universal Generalization is given in the informal explanation of his notation for quantification. When appended to the axioms of truth-functional logic, these constitute a complete and consistent axiomatization of first-order logic.

Fourth, propositions (52) and (54), when appended to the other rules, constitute a complete, consistent axiomatization of first-order logic *with identity*.<sup>2</sup>

Fifth, higher-order quantification is introduced and used – to define the ancestral relation, a key in the definition of mathematical induction – though it is not clearly delimited from first-order quantification since proposition (58) is apparently intended to serve for any order quantifier, and not just the first-order quantifier in whose terms it is explicitly stated.

It is no small measure of the greatness of the work that his contemporaries just did not understand what he had done.<sup>3</sup> Part of the reason for the lack of understanding was his incredibly cumbersome notation. Frege's logical notation was quite idiosyncratic. It was different from the type of notation in logic used earlier as well as that which has subsequently become standard. Frege introduced a *horizontal* line to combine the parts that followed into a complete whole, a name of a thought. So,

$$\text{—— } \Delta$$

has us entertain the thought that  $\Delta$ . When the *vertical* stroke is appended in front, we get *the judgment that*  $\Delta$ :

$$\text{┆—— } \Delta$$

Frege then introduces two Boolean operations. For *It is not the case that*  $\Delta$ , we have

$$\text{┆┆—— } \Delta$$

and for *If*  $\Gamma$  *then*  $\Delta$ , we have

$$\text{┆┆┆—— } \Delta$$

$$\text{┆┆┆—— } \Gamma$$

Frege's was a parentheses-free notation, so he had to have a notation that made priority of operations clear. It does.  $\Gamma \supset \neg\Delta$  goes in as

$$\text{┆┆┆┆—— } \Delta$$

$$\text{┆┆┆┆—— } \Gamma$$

$\neg(\Gamma \supset \Delta)$  goes in as

$$\text{┆┆┆┆┆—— } \Delta$$

$$\text{┆┆┆┆┆—— } \Gamma$$

$\neg\Gamma \supset \Delta$  goes in as

$$\text{┆┆┆┆┆—— } \Delta$$

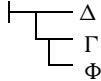
$$\text{┆┆┆┆┆—— } \Gamma$$



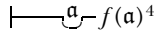
We also need to distinguish these two formulas:  $\Phi \supset (\Gamma \supset \Delta)$



and  $(\Phi \supset \Gamma) \supset \Delta$



Quantifiers are represented using the appropriate German gothic letters in the concavity and the formula. Our modern  $(\forall x)F(x)$  is represented in his notation so:<sup>4</sup>



As Schröder said, Frege’s notation appeared to copy the Japanese form of writing vertically. Frege wanted to capture the way in which a proof is written, with each step in the proof on a different line. But in his notation, he took this one step further. He broke up a *formula* so that each of its logically significant components appeared on a different line. This has proved practically unreadable. Hence this appendix, which is designed to enable the reader to understand the logical apparatus of *Begriffsschrift* by rewriting the proofs in modern notation. We include the propositions from Parts 1 and 2, only. Boolos (1985) has already provided an intelligible rendering of the important propositions from Part 3.

We use  $\neg$  for the negation sign and  $\supset$  for the material conditional, with appropriate parentheses, braces, or brackets for ease of reading. We use lower case  $a, b, c, \dots$  for propositional variables. Our substitution notation will be  $X_{x,y,z,\dots}^{a,b,c,\dots}$  which designates the result of replacing  $a$  for  $x, b$  for  $y, c$  for  $z, \dots$  in  $X$ . Each line in the proof is numbered and the proposition from which it results by substitution is written on the right. There are only two lines in each proof, because each inference is obtained by Modus Ponens.

PROPOSITION 1

$$a \supset (b \supset a)$$

PROPOSITION 2

$$(c \supset (b \supset a)) \supset ((c \supset b) \supset (c \supset a))$$

PROPOSITION 3

$$(b \supset a) \supset [(c \supset (b \supset a)) \supset ((c \supset b) \supset (c \supset a))]$$

*Proof:* by Modus Ponens from

1.  $(c \supset (b \supset a)) \supset ((c \supset b) \supset (c \supset a))$  (2)
2.  $[((c \supset (b \supset a)) \supset ((c \supset b) \supset (c \supset a))) \supset [(b \supset a) \supset ((c \supset (b \supset a)) \supset ((c \supset b) \supset (c \supset a)))]]$  (1)<sub>a,b</sub><sup>(c⊃(b⊃a))⊃((c⊃b)⊃(c⊃a)), b⊃a</sup>

**PROPOSITION 4**

$$[(b \supset a) \supset (c \supset (b \supset a))] \supset [(b \supset a) \supset ((c \supset b) \supset (c \supset a))]$$

*Proof:* by Modus Ponens from

1.  $(b \supset a) \supset [(c \supset (b \supset a)) \supset ((c \supset b) \supset (c \supset a))]$  (3)
2.  $[(b \supset a) \supset ((c \supset (b \supset a)) \supset ((c \supset b) \supset (c \supset a)))] \supset [((b \supset a) \supset (c \supset (b \supset a))) \supset ((b \supset a) \supset ((c \supset b) \supset (c \supset a)))]$  (2)<sub>a,b,c</sub><sup>(c⊃b)⊃(c⊃a), c⊃(b⊃a), b⊃a</sup>

**PROPOSITION 5**

$$(b \supset a) \supset ((c \supset b) \supset (c \supset a))$$

*Proof:* by Modus Ponens from

1.  $(b \supset a) \supset (c \supset (b \supset a))$  (1)<sub>a,b</sub><sup>b⊃a,c</sup>
2.  $[(b \supset a) \supset (c \supset (b \supset a))] \supset [(b \supset a) \supset ((c \supset b) \supset (c \supset a))]$  (4)

**PROPOSITION 6**

$$(c \supset (b \supset a)) \supset (c \supset ((d \supset b) \supset (d \supset a)))$$

*Proof:* by Modus Ponens from

1.  $(b \supset a) \supset ((d \supset b) \supset (d \supset a))$  (5)<sub>c</sub><sup>d</sup>
2.  $((b \supset a) \supset ((d \supset b) \supset (d \supset a))) \supset ((c \supset (b \supset a)) \supset (c \supset ((d \supset b) \supset (d \supset a))))$  (5)<sub>a,b</sub><sup>(d⊃b)⊃(d⊃a), b⊃a</sup>

**PROPOSITION 7**

$$(b \supset a) \supset ((d \supset (c \supset b)) \supset (d \supset (c \supset a)))$$

*Proof:* by Modus Ponens from

1.  $(b \supset a) \supset ((c \supset b) \supset (c \supset a))$  (5)
2.  $((b \supset a) \supset ((c \supset b) \supset (c \supset a))) \supset ((b \supset a) \supset ((d \supset (c \supset b)) \supset (d \supset (c \supset a))))$  (6)<sub>a,b,c</sub><sup>c⊃a, c⊃b, b⊃a</sup>

## PROPOSITION 8

$$(d \supset (b \supset a)) \supset (b \supset (d \supset a))$$

## PROPOSITION 9

$$(c \supset b) \supset ((b \supset a) \supset (c \supset a))$$

*Proof:* by Modus Ponens from

1.  $(b \supset a) \supset ((c \supset b) \supset (c \supset a))$  (5)
2.  $((b \supset a) \supset ((c \supset b) \supset (c \supset a))) \supset$   
 $((c \supset b) \supset ((b \supset a) \supset (c \supset a)))$  (8)<sub>a,b,d</sub><sup>c⊃a, c⊃b, b⊃a</sup>

## PROPOSITION 10

$$((e \supset (d \supset b)) \supset a) \supset ((d \supset (e \supset b)) \supset a)$$

*Proof:* by Modus Ponens from

1.  $(d \supset (e \supset b)) \supset (e \supset (d \supset b))$  (8)<sub>a,b</sub><sup>b,e</sup>
2.  $((d \supset (e \supset b)) \supset (e \supset (d \supset b))) \supset$   
 $((e \supset (d \supset b)) \supset a) \supset ((d \supset (e \supset b)) \supset a)$  (9)<sub>b,c</sub><sup>e⊃(d⊃b), d⊃(e⊃b)</sup>

## PROPOSITION 11

$$((c \supset b) \supset a) \supset (b \supset a)$$

*Proof:* by Modus Ponens from

1.  $b \supset (c \supset b)$  (1)<sub>a,b</sub><sup>b,c</sup>
2.  $(b \supset (c \supset b)) \supset (((c \supset b) \supset a) \supset (b \supset a))$  (9)<sub>b,c</sub><sup>c⊃b, b</sup>

## PROPOSITION 12

$$(d \supset (c \supset (b \supset a))) \supset (d \supset (b \supset (c \supset a)))$$

*Proof:* by Modus Ponens from

1.  $(c \supset (b \supset a)) \supset (b \supset (c \supset a))$  (8)<sub>d</sub><sup>c</sup>
2.  $((c \supset (b \supset a)) \supset (b \supset (c \supset a))) \supset$   
 $((d \supset (c \supset (b \supset a))) \supset (d \supset (b \supset (c \supset a))))$  (5)<sub>a,b,c</sub><sup>b⊃(c⊃a), c⊃(b⊃a), d</sup>

## PROPOSITION 13

$$(d \supset (c \supset (b \supset a))) \supset (b \supset (d \supset (c \supset a)))$$

*Proof:* by Modus Ponens from

1.  $(d \supset (c \supset (b \supset a))) \supset (d \supset (b \supset (c \supset a)))$  (12)
2.  $((d \supset (c \supset (b \supset a))) \supset (d \supset (b \supset (a \supset c)))) \supset$   
 $((d \supset (c \supset (b \supset a))) \supset (b \supset (d \supset (c \supset a))))$  (12) <sub>$c \supset a, d, d \supset (c \supset (b \supset a)), e$</sub>

**PROPOSITION 14**

$$(e \supset (d \supset (c \supset (b \supset a)))) \supset (e \supset (b \supset (d \supset (c \supset a))))$$

*Proof:* by Modus Ponens from

1.  $(d \supset (c \supset (b \supset a))) \supset (b \supset (d \supset (c \supset a)))$  (13)
2.  $((d \supset (c \supset (b \supset a))) \supset (b \supset (d \supset (c \supset a)))) \supset$   
 $((e \supset (d \supset (c \supset (b \supset a)))) \supset (e \supset (b \supset (d \supset (c \supset a))))$  (5) <sub>$b \supset (d \supset (c \supset a)), d \supset (c \supset (b \supset a)), e$</sub>

**PROPOSITION 15**

$$(e \supset (d \supset (c \supset (b \supset a)))) \supset (b \supset (e \supset (d \supset (c \supset a))))$$

*Proof:* by Modus Ponens from

1.  $(e \supset (d \supset (c \supset (b \supset a)))) \supset (e \supset (b \supset (d \supset (c \supset a))))$  (14)
2.  $((e \supset (d \supset (c \supset (b \supset a)))) \supset (e \supset (b \supset (d \supset (c \supset a)))) \supset$   
 $((e \supset (d \supset (c \supset (b \supset a)))) \supset (b \supset (e \supset (d \supset (c \supset a))))$  (12) <sub>$d \supset (c \supset a), e, (e \supset (d \supset (c \supset (b \supset a)))$</sub>

**PROPOSITION 16**

$$(e \supset (d \supset (c \supset (b \supset a)))) \supset (e \supset (d \supset (b \supset (c \supset a))))$$

*Proof:* by Modus Ponens from

1.  $(d \supset (c \supset (b \supset a))) \supset (d \supset (b \supset (c \supset a)))$  (12)
2.  $((d \supset (c \supset (b \supset a))) \supset (d \supset (b \supset (c \supset a)))) \supset$   
 $((e \supset (d \supset (c \supset (b \supset a)))) \supset (e \supset (d \supset (b \supset (c \supset a))))$  (5) <sub>$d \supset (b \supset (c \supset a)), d \supset (c \supset (b \supset a)), e$</sub>

**PROPOSITION 17**

$$(d \supset (c \supset (b \supset a))) \supset (c \supset (b \supset (d \supset a)))$$

*Proof:* by Modus Ponens from

1.  $(d \supset (c \supset (b \supset a))) \supset (c \supset (d \supset (b \supset a)))$  (8) <sub>$b \supset a, c$</sub>
2.  $((d \supset (c \supset (b \supset a))) \supset (c \supset (d \supset (b \supset a)))) \supset$   
 $((d \supset (c \supset (b \supset a))) \supset (c \supset (b \supset (d \supset a))))$  (16) <sub>$d, c, d \supset (c \supset (b \supset a))$</sub>

## PROPOSITION 18

$$(c \supset (b \supset a)) \supset ((d \supset c) \supset (b \supset (d \supset a)))$$

*Proof:* by Modus Ponens from

1.  $(c \supset (b \supset a)) \supset ((c \supset d) \supset (d \supset (b \supset a)))$  (5)<sup>b>a,c,d</sup><sub>a,b,c</sub>
2.  $((c \supset (b \supset a)) \supset ((c \supset d) \supset (d \supset (b \supset a)))) \supset$   
 $((c \supset (b \supset a)) \supset ((d \supset c) \supset (b \supset (d \supset a))))$  (16)<sup>d,d>c,c>(b>a)</sup><sub>c,d,e</sub>

## PROPOSITION 19

$$(d \supset (c \supset b)) \supset ((b \supset a) \supset (d \supset (c \supset a)))$$

*Proof:* by Modus Ponens from

1.  $(c \supset b) \supset ((b \supset a) \supset (c \supset a))$  (9)
2.  $((c \supset b) \supset ((b \supset a) \supset (c \supset a))) \supset$   
 $((d \supset (c \supset b)) \supset ((b \supset a) \supset (d \supset (c \supset a))))$  (18)<sup>c>a,b>a,c>b</sup><sub>a,b,c</sub>

## PROPOSITION 20

$$(e \supset (d \supset (c \supset b))) \supset ((b \supset a) \supset (e \supset (d \supset (c \supset a))))$$

*Proof:* by Modus Ponens from

1.  $(d \supset (c \supset b)) \supset ((b \supset a) \supset (d \supset (c \supset a)))$  (19)
2.  $((d \supset (c \supset b)) \supset ((b \supset a) \supset (d \supset (c \supset a)))) \supset$   
 $((e \supset (d \supset (c \supset b))) \supset ((b \supset a) \supset (e \supset (d \supset (c \supset a))))$  (18)<sup>d>(c>a),b>a,d>(c>b),e</sup><sub>a,b,c,d</sub>

## PROPOSITION 21

$$((d \supset b) \supset a) \supset ((d \supset c) \supset ((c \supset b) \supset a))$$

*Proof:* by Modus Ponens from

1.  $(d \supset c) \supset ((c \supset b) \supset (d \supset b))$  (9)<sup>b,c,d</sup><sub>a,b,c</sub>
2.  $((d \supset c) \supset ((c \supset b) \supset (d \supset b))) \supset$   
 $((d \supset b) \supset a) \supset ((d \supset c) \supset ((c \supset b) \supset a))$  (19)<sup>d>b,c>b,d>c</sup><sub>b,c,d</sub>

## PROPOSITION 22

$$(f \supset (e \supset (d \supset (c \supset (b \supset a)))) \supset (f \supset (e \supset (d \supset (b \supset (c \supset a)))))$$

*Proof:* by Modus Ponens from

1.  $(e \supset (d \supset (c \supset (b \supset a)))) \supset (e \supset (d \supset (b \supset (c \supset a))))$  (16)
2.  $((e \supset (d \supset (c \supset b \supset a))) \supset (e \supset (d \supset (b \supset (c \supset a)))) \supset [(f \supset (e \supset (d \supset (c \supset (b \supset a)))) \supset (f \supset (e \supset (d \supset (b \supset (c \supset a))))])$  (5)<sub>a,b,c</sub><sup>e $\supset$ (d $\supset$ (b $\supset$ (c $\supset$ a))),e $\supset$ (d $\supset$ (c $\supset$ (b $\supset$ a))),f</sup>

**PROPOSITION 23**

$$(d \supset (c \supset (b \supset a))) \supset ((e \supset d) \supset (c \supset (b \supset (e \supset a))))$$

*Proof:* by Modus Ponens from

1.  $(d \supset (c \supset (b \supset a))) \supset ((e \supset d) \supset (c \supset (e \supset (b \supset a))))$  (18)<sub>a,b,c,d</sub><sup>b $\supset$ a,c,d,e</sup>
2.  $((d \supset (c \supset (b \supset a))) \supset ((e \supset d) \supset (c \supset (e \supset (b \supset a)))) \supset ((d \supset (c \supset (b \supset a))) \supset ((e \supset d) \supset (c \supset (b \supset (e \supset a))))$  (22)<sub>c,d,e,f</sub><sup>e,c,e $\supset$ d,d $\supset$ (c $\supset$ (b $\supset$ a))</sup>

**PROPOSITION 24**

$$(c \supset a) \supset (c \supset (b \supset a))$$

*Proof:* by Modus Ponens from

1.  $(c \supset a) \supset (b \supset (c \supset a))$  (1)<sub>a</sub><sup>c $\supset$ a</sup>
2.  $((c \supset a) \supset (b \supset (c \supset a))) \supset ((c \supset a) \supset (c \supset (b \supset a)))$  (12)<sub>b,c,d</sub><sup>c,b,c $\supset$ a</sup>

**PROPOSITION 25**

$$(d \supset (c \supset a)) \supset (d \supset (c \supset (b \supset a)))$$

*Proof:* by Modus Ponens from

1.  $(c \supset a) \supset (c \supset (b \supset a))$  (24)
2.  $((c \supset a) \supset (c \supset (b \supset a))) \supset ((d \supset (c \supset a)) \supset (d \supset (c \supset (b \supset a))))$  (5)<sub>a,b,c</sub><sup>c $\supset$ (b $\supset$ a),c $\supset$ a,d</sup>

**PROPOSITION 26**

$$b \supset (a \supset a)$$

*Proof:* by Modus Ponens from

1.  $a \supset (b \supset a)$  (1)
2.  $(a \supset (b \supset a)) \supset (b \supset (a \supset a))$  (8)<sub>d</sub><sup>a</sup>

## PROPOSITION 27

$$a \supset a$$

*Proof:* by Modus Ponens from

1.  $a \supset (b \supset a)$  (1)
2.  $(a \supset (b \supset a)) \supset (a \supset a)$  (26)<sub>b</sub><sup>a ⊃ (b ⊃ a)</sup>

## PROPOSITION 28

$$(b \supset a) \supset (\neg a \supset \neg b)$$

## PROPOSITION 29

$$(c \supset (b \supset a)) \supset (c \supset (\neg a \supset \neg b))$$

*Proof:* by Modus Ponens from

1.  $(b \supset a) \supset (\neg a \supset \neg b)$  (28)
2.  $((b \supset a) \supset (\neg a \supset \neg b)) \supset ((c \supset (b \supset a)) \supset (c \supset (\neg a \supset \neg b)))$  (5)<sub>a,b</sub><sup>¬ a ⊃ ¬ b, b ⊃ a</sup>

## PROPOSITION 30

$$(b \supset (c \supset a)) \supset (c \supset (\neg a \supset \neg b))$$

*Proof:* by Modus Ponens from

1.  $(c \supset (b \supset a)) \supset (c \supset (\neg a \supset \neg b))$  (29)
2.  $((c \supset (b \supset a)) \supset (c \supset (\neg a \supset \neg b))) \supset ((b \supset (c \supset a)) \supset (c \supset (\neg a \supset \neg b)))$  (10)<sub>a,b,d,e</sub><sup>c ⊃ (¬ a ⊃ ¬ b), a, b, c</sup>

## PROPOSITION 31

$$\neg \neg a \supset a$$

## PROPOSITION 32

$$((\neg b \supset a) \supset (\neg a \supset \neg \neg b)) \supset ((\neg b \supset a) \supset (\neg a \supset b))$$

*Proof:* by Modus Ponens from

1.  $\neg \neg b \supset b$  (31)<sub>a</sub><sup>b</sup>
2.  $(\neg \neg b \supset b) \supset (((\neg b \supset a) \supset (\neg a \supset \neg \neg b)) \supset ((\neg b \supset a) \supset (\neg a \supset b)))$  (7)<sub>a,b,c,d</sub><sup>b, ¬ ¬ b, ¬ a, ¬ b ⊃ a</sup>

## PROPOSITION 33

$$(\neg b \supset a) \supset (\neg a \supset b)$$

*Proof:* by Modus Ponens from

1.  $((\neg b \supset a) \supset (\neg a \supset \neg\neg b)) \supset ((\neg b \supset a) \supset (\neg a \supset b))$  (32)
2.  $(\neg b \supset a) \supset (\neg a \supset \neg\neg b)$  (28)<sub>b</sub><sup>b</sup>

## PROPOSITION 34

$$(c \supset (\neg b \supset a)) \supset (c \supset (\neg a \supset b))$$

*Proof:* by Modus Ponens from

1.  $(\neg b \supset a) \supset (\neg a \supset b)$  (33)
2.  $((\neg b \supset a) \supset (\neg a \supset b)) \supset$   
 $((c \supset (\neg b \supset a)) \supset (c \supset (\neg a \supset b)))$  (5)<sub>a,b</sub> <sup>$\neg a \supset b, \neg b \supset a$</sup>

## PROPOSITION 35

$$(c \supset (\neg b \supset a)) \supset (\neg a \supset (c \supset b))$$

*Proof:* by Modus Ponens from

1.  $(c \supset (\neg b \supset a)) \supset (c \supset (\neg a \supset b))$  (34)
2.  $((c \supset (\neg b \supset a)) \supset (c \supset (\neg a \supset b))) \supset$   
 $((c \supset (\neg b \supset a)) \supset (\neg a \supset (c \supset b)))$  (12)<sub>a,b,d</sub> <sup>$b, \neg a, c \supset (\neg b \supset a)$</sup>

## PROPOSITION 36

$$a \supset (\neg a \supset b)$$

*Proof:* by Modus Ponens from

1.  $a \supset (\neg b \supset a)$  (1)<sub>b</sub><sup>b</sup>
2.  $(a \supset (\neg b \supset a)) \supset (a \supset (\neg a \supset b))$  (34)<sub>c</sub><sup>a</sup>

## PROPOSITION 37

$$((\neg c \supset b) \supset a) \supset (c \supset a)$$

*Proof:* by Modus Ponens from

1.  $c \supset (\neg c \supset b)$  (36)<sub>a</sub><sup>c</sup>
2.  $(c \supset (\neg c \supset b)) \supset (((\neg c \supset b) \supset a) \supset (c \supset a))$  (9)<sub>b</sub> <sup>$\neg c \supset b$</sup>



## PROPOSITION 38

$$\neg a \supset (a \supset b)$$

*Proof:* by Modus Ponens from

1.  $a \supset (\neg a \supset b)$  (36)
2.  $(a \supset (\neg a \supset b)) \supset (\neg a \supset (a \supset b))$  (8)<sub>a,b,d</sub><sup>b,¬a,a</sup>

## PROPOSITION 39

$$(\neg a \supset a) \supset (\neg a \supset b)$$

*Proof:* by Modus Ponens from

1.  $\neg a \supset (a \supset b)$  (38)
2.  $(\neg a \supset (a \supset b)) \supset ((\neg a \supset a) \supset (\neg a \supset b))$  (2)<sub>a,b,c</sub><sup>b,a,¬a</sup>

## PROPOSITION 40

$$\neg b \supset ((\neg a \supset a) \supset a)$$

*Proof:* by Modus Ponens from

1.  $(\neg a \supset a) \supset (\neg a \supset b)$  (39)
2.  $((\neg a \supset a) \supset (\neg a \supset b)) \supset (\neg b \supset ((\neg a \supset a) \supset a))$  (35)<sub>a,b,c</sub><sup>b,a,¬a⊃a</sup>

## PROPOSITION 41

$$a \supset \neg\neg a$$

## PROPOSITION 42

$$\neg\neg(a \supset a)$$

*Proof:* by Modus Ponens from

1.  $a \supset a$  (27)
2.  $(a \supset a) \supset \neg\neg(a \supset a)$  (41)<sub>a</sub><sup>a⊃a</sup>

## PROPOSITION 43

$$(\neg a \supset a) \supset a$$

*Proof:* by Modus Ponens from

1.  $\neg\neg(a \supset a)$  (42)
2.  $\neg\neg(a \supset a) \supset ((\neg a \supset a) \supset a)$  (40)<sub>b</sub><sup>¬(a⊃a)</sup>

## PROPOSITION 44

$$(\neg a \supset c) \supset ((c \supset a) \supset a)$$

*Proof:* by Modus Ponens from

1.  $(\neg a \supset a) \supset a$  (43)
2.  $((\neg a \supset a) \supset a) \supset ((\neg a \supset c) \supset ((c \supset a) \supset a))$  (21)<sub>b,d</sub><sup>a,¬a</sup>

## PROPOSITION 45

$$((\neg c \supset a) \supset (\neg a \supset c)) \supset ((\neg c \supset a) \supset ((c \supset a) \supset a))$$

*Proof:* by Modus Ponens from

1.  $(\neg a \supset c) \supset ((c \supset a) \supset a)$  (44)
2.  $((\neg a \supset c) \supset ((c \supset a) \supset a)) \supset (((\neg c \supset a) \supset (\neg a \supset c)) \supset ((\neg c \supset a) \supset ((c \supset a) \supset a)))$  (5)<sub>a,b,c</sub><sup>(c⊃a)⊃a,¬a⊃c,¬c⊃a</sup>

## PROPOSITION 46

$$(\neg c \supset a) \supset ((c \supset a) \supset a)$$

*Proof:* by Modus Ponens from

1.  $((\neg c \supset a) \supset (\neg a \supset c)) \supset ((\neg c \supset a) \supset ((c \supset a) \supset a))$  (45)
2.  $(\neg c \supset a) \supset (\neg a \supset c)$  (33)<sub>b</sub><sup>c</sup>

## PROPOSITION 47

$$(\neg c \supset b) \supset ((b \supset a) \supset ((c \supset a) \supset a))$$

*Proof:* by Modus Ponens from

1.  $(\neg c \supset a) \supset ((c \supset a) \supset a)$  (46)
2.  $((\neg c \supset a) \supset ((c \supset a) \supset a)) \supset ((\neg c \supset b) \supset ((b \supset a) \supset ((c \supset a) \supset a)))$  (21)<sub>a,b,d,c</sub><sup>(c⊃a)⊃a,¬c⊃b</sup>

## PROPOSITION 48

$$(d \supset (\neg c \supset b)) \supset ((b \supset a) \supset ((c \supset a) \supset (d \supset a)))$$

*Proof:* by Modus Ponens from

1.  $(\neg c \supset b) \supset ((b \supset a) \supset ((c \supset a) \supset a))$  (47)
2.  $((\neg c \supset b) \supset ((b \supset a) \supset ((c \supset a) \supset a))) \supset ((d \supset (\neg c \supset b)) \supset ((b \supset a) \supset ((c \supset a) \supset (d \supset a))))$  (23)<sub>b,c,d,e</sub><sup>c⊃a,b⊃a,¬c⊃b,d</sup>

## PROPOSITION 49

$$(\neg c \supset b) \supset ((c \supset a) \supset ((b \supset a) \supset a))$$

*Proof:* by Modus Ponens from

$$1. (\neg c \supset b) \supset ((b \supset a) \supset ((c \supset a) \supset a)) \quad (47)$$

$$2. ((\neg c \supset b) \supset ((b \supset a) \supset ((c \supset a) \supset a))) \supset$$

$$((\neg c \supset b) \supset ((c \supset a) \supset ((b \supset a) \supset a))) \quad (12)_{b,c,d}^{c \supset a, b \supset a, \neg c \supset b}$$

## PROPOSITION 50

$$(c \supset a) \supset ((b \supset a) \supset ((\neg c \supset b) \supset a))$$

*Proof:* by Modus Ponens from

$$1. (\neg c \supset b) \supset ((c \supset a) \supset ((b \supset a) \supset a)) \quad (49)$$

$$2. ((\neg c \supset b) \supset ((c \supset a) \supset ((b \supset a) \supset a))) \supset$$

$$((c \supset a) \supset ((b \supset a) \supset ((\neg c \supset b) \supset a))) \quad (17)_{b,c,d}^{b \supset a, c \supset a, \neg c \supset b}$$

## PROPOSITION 51

$$(d \supset (c \supset a)) \supset ((b \supset a) \supset (d \supset ((\neg c \supset b) \supset a)))$$

*Proof:* by Modus Ponens from

$$1. (c \supset a) \supset ((b \supset a) \supset ((\neg c \supset b) \supset a)) \quad (50)$$

$$2. ((c \supset a) \supset ((b \supset a) \supset ((\neg c \supset b) \supset a))) \supset$$

$$((d \supset (c \supset a)) \supset ((b \supset a) \supset (d \supset ((\neg c \supset b) \supset a)))) \quad (18)_{a,b,c}^{\neg c \supset (b \supset a), b \supset a, c \supset a}$$

## Appendix B

### Begriffsschrift *in Modern Notation: (52) to (68)*

We use  $\neg$  for the negation sign,  $\supset$  for the material conditional,  $\forall$  for the universal quantifier,  $\equiv$  for “identity of content,” with appropriate parentheses, braces, or brackets used for ease of reading. Predicates will be expressed using function notation. Our substitution notation is as in Appendix A. As before, each line in the proof is numbered and the proposition from which it results by substitution is written on the right. In the proof of proposition (55) we follow Frege’s usage: (53) <sup>$A \equiv c$</sup>  <sub>$f(A)$</sub>  means that the function  $f(x)$  is  $x \equiv c$ .

PROPOSITION 52

$$(c \equiv d) \supset (f(c) \supset f(d))$$

PROPOSITION 53

$$f(c) \supset ((c \equiv d) \supset f(d))$$

*Proof:* by Modus Ponens from

1.  $(c \equiv d) \supset (f(c) \supset f(d))$  (52)
2.  $((c \equiv d) \supset (f(c) \supset f(d))) \supset (f(c) \supset ((c \equiv d) \supset f(d)))$  (8) <sup>$f(d), f(c), c \equiv d$</sup>  <sub>$a, b, d$</sub>

PROPOSITION 54

$$c \equiv c$$

PROPOSITION 55

$$(c \equiv d) \supset (d \equiv c)$$

*Proof:* by Modus Ponens from

1.  $c \equiv c$  (54)
2.  $(c \equiv c) \supset ((c \equiv d) \supset (d \equiv c))$  (53)<sup>A=c</sup><sub>f(A)</sub>

PROPOSITION 56

$$((d \equiv c) \supset (f(d) \supset f(c))) \supset ((c \equiv d) \supset (f(d) \supset f(c)))$$

*Proof:* by Modus Ponens from

1.  $(c \equiv d) \supset (d \equiv c)$  (55)
2.  $((c \equiv d) \supset (d \equiv c)) \supset (((d \equiv c) \supset (f(d) \supset f(c))) \supset ((c \equiv d) \supset (f(d) \supset f(c))))$  (9)<sup>d=c, c=d, f(d)⊃f(c)</sup><sub>b, c, a</sub>

PROPOSITION 57

$$(c \equiv d) \supset (f(d) \supset f(c))$$

*Proof:* by Modus Ponens from

1.  $((d \equiv c) \supset (f(d) \supset f(c))) \supset ((c \equiv d) \supset (f(d) \supset f(c)))$  (56)
2.  $(d \equiv c) \supset (f(d) \supset f(c))$  (52)<sup>c, d</sup><sub>a, c</sub>

PROPOSITION 58

$$(\forall x) f(x) \supset f(c)$$

PROPOSITION 59

$$g(b) \supset (\neg f(b) \supset \neg(\forall x)(g(x) \supset f(x)))$$

*Proof:* by Modus Ponens from

1.  $(\forall x)(g(x) \supset f(x)) \supset (g(b) \supset f(b))$  (58)<sup>g(A)⊃f(A), b</sup><sub>f(A), c</sub>
2.  $((\forall x)(g(x) \supset f(x)) \supset (g(b) \supset f(b))) \supset (g(b) \supset (\neg f(b) \supset \neg(\forall x)(g(x) \supset f(x))))$  (30)<sup>f(b), g(b), (∀x)(g(x)⊃f(x))</sup><sub>a, c, b</sub>

PROPOSITION 60

$$(\forall x)(h(x) \supset (g(x) \supset f(x))) \supset (g(b) \supset (h(b) \supset f(b)))$$

*Proof:* by Modus Ponens from

1.  $(\forall x)(h(x) \supset (g(x) \supset f(x))) \supset (h(b) \supset (g(b) \supset f(b)))$  (58)<sup>h(A)⊃(g(A)⊃f(A), b</sup><sub>f(A), c</sub>

$$\begin{aligned}
2. & ((\forall x)(h(x) \supset (g(x) \supset f(x))) \supset (h(b) \supset \\
& (g(b) \supset f(b)))) \supset ((\forall x)(h(x) \supset (g(x) \supset \\
& f(x))) \supset (g(b) \supset (h(b) \supset f(b)))) \quad (12)_{a,b,c,d}^{f(b),g(b),h(b),(\forall x)(h(x)\supset(g(x)\supset f(x)))}
\end{aligned}$$

## PROPOSITION 61

$$(f(c) \supset a) \supset ((\forall x) f(x) \supset a)$$

*Proof:* by Modus Ponens from

$$\begin{aligned}
1. & (\forall x) f(x) \supset f(c) \quad (58) \\
2. & ((\forall x) f(x) \supset f(c)) \supset ((f(c) \supset a) \supset ((\forall x) f(x) \supset a)) \quad (9)_{b,c}^{f(c),(\forall x)f(x)}
\end{aligned}$$

## PROPOSITION 62

$$g(y) \supset ((\forall x)(g(x) \supset f(x)) \supset f(y))$$

*Proof:* by Modus Ponens from

$$\begin{aligned}
1. & (\forall x)(g(x) \supset f(x)) \supset (g(y) \supset f(y)) \quad (58)_{f(A),c}^{g(A)\supset f(A),y} \\
2. & ((\forall x)(g(x) \supset f(x)) \supset (g(y) \supset f(y))) \supset \\
& (g(y) \supset ((\forall x)(g(x) \supset f(x)) \supset f(y))) \quad (8)_{a,b,d}^{f(y),g(y),(\forall x)(g(x)\supset f(x))}
\end{aligned}$$

## PROPOSITION 63

$$g(y) \supset (m \supset ((\forall x)(g(x) \supset f(x)) \supset f(y)))$$

*Proof:* by Modus Ponens from

$$\begin{aligned}
1. & g(y) \supset ((\forall x)(g(x) \supset f(x)) \supset f(y)) \quad (62) \\
2. & (g(y) \supset ((\forall x)(g(x) \supset f(x)) \supset f(y))) \supset \\
& (g(y) \supset (m \supset ((\forall x)(g(x) \supset f(x)) \supset f(y)))) \quad (24)_{a,c,b}^{(\forall x)(g(x)\supset f(x))\supset f(y),g(y),m}
\end{aligned}$$

## PROPOSITION 64

$$(h(z) \supset g(y)) \supset ((\forall x)(g(x) \supset f(x)) \supset (h(z) \supset f(y)))$$

*Proof:* by Modus Ponens from

$$\begin{aligned}
1. & g(y) \supset ((\forall x)(g(x) \supset f(x)) \supset f(y)) \quad (62) \\
2. & (g(y) \supset ((\forall x)(g(x) \supset f(x)) \supset f(y))) \supset \\
& ((h(z) \supset g(y)) \supset ((\forall x)(g(x) \supset f(x)) \supset \\
& (h(z) \supset f(y)))) \quad (18)_{a,b,c,d}^{f(y),(\forall x)(g(x)\supset f(x)),g(y),h(z)}
\end{aligned}$$

## PROPOSITION 65

$$(\forall x)(h(x) \supset g(x)) \supset ((\forall x)(g(x) \supset f(x)) \supset (h(y) \supset f(y)))$$

*Proof:* by Modus Ponens from

1.  $(h(y) \supset g(y)) \supset ((\forall x)(g(x) \supset f(x)) \supset (h(y) \supset f(y)))$  (64)<sub>z</sub><sup>y</sup>
2.  $((h(y) \supset g(y)) \supset ((\forall x)(g(x) \supset f(x)) \supset (h(y) \supset f(y)))) \supset ((\forall x)(h(x) \supset g(x)) \supset ((\forall x)(g(x) \supset f(x)) \supset (h(y) \supset f(y))))$  (61)<sub>a, f(A), c</sub><sup>(\forall x)(g(x) \supset f(x)) \supset (h(y) \supset f(y)), h(A) \supset g(A), y</sup>

**PROPOSITION 66**

$$(\forall x)(g(x) \supset f(x)) \supset ((\forall x)(h(x) \supset g(x)) \supset (h(y) \supset f(y)))$$

*Proof:* by Modus Ponens from

1.  $(\forall x)(h(x) \supset g(x)) \supset ((\forall x)(g(x) \supset f(x)) \supset (h(y) \supset f(y)))$  (65)
2.  $((\forall x)(h(x) \supset g(x)) \supset ((\forall x)(g(x) \supset f(x)) \supset (h(y) \supset f(y)))) \supset ((\forall x)(g(x) \supset f(x)) \supset ((\forall x)(h(x) \supset g(x)) \supset (h(y) \supset f(y))))$  (8)<sub>a, b, d</sub><sup>h(y) \supset f(y), (\forall x)(g(x) \supset f(x)), (\forall x)(h(x) \supset g(x))</sup>

**PROPOSITION 67**

$$(((\forall x) f(x) \equiv b) \supset (b \supset (\forall x) f(x))) \supset (((\forall x) f(x) \equiv b) \supset (b \supset f(c)))$$

*Proof:* by Modus Ponens from

1.  $(\forall x) f(x) \supset f(c)$  (58)
2.  $((\forall x) f(x) \supset f(c)) \supset (((\forall x) f(x) \equiv b) \supset (b \supset (\forall x) f(x))) \supset (((\forall x) f(x) \equiv b) \supset (b \supset f(c)))$  (7)<sub>a, b, c, d</sub><sup>f(c), (\forall x) f(x), b, ((\forall x) f(x) \equiv b)</sup>

**PROPOSITION 68**

$$((\forall x) f(x) \equiv b) \supset (b \supset f(c))$$

*Proof:* by Modus Ponens from

1.  $(((\forall x) f(x) \equiv b) \supset (b \supset (\forall x) f(x))) \supset ((\forall x) f(x) \equiv b) \supset (b \supset f(c))$  (67)
2.  $((\forall x) f(x) \equiv b) \supset (b \supset (\forall x) f(x))$  (57)<sub>f(A), c, d</sub><sup>A, (\forall x) f(x), b</sup>





## Notes

### Preface

1. See Mates (1961). In each of these cases, however, the ‘sense’ component appears to be more psychological than Frege’s.
2. These appear in Frege (1980).
3. Of course, the ambiguity is virulent in the German as well. But is this something we wish to encourage?

### Chapter 1

1. The information in this biography, as in almost all of Frege’s biographies, is, with minor additions, drawn from the material in the Frege archives at the University of Münster. It was first put together in the introduction by Bynum (1972). Two recent contributions, one by Gabriel & Kienzler (1997) and the other by Lothar Kreiser (2001), shed more light on Frege’s youth in Jena.
2. Kreiser (2001) reports that Alfred was adopted after Margarete died; Beaney (1997) reports that the two had adopted Alfred shortly before her death.
3. G. E. M. Anscombe, in a personal communication, relayed that Alfred had stored a trunk of Frege’s belongings in a farmhouse somewhere in Europe, the location of which has been lost with Alfred’s demise.
4. See Sluga (1980), Dummett (1981*b*), Baker & Hacker (1984).
5. Frege begins his Political Diary with an encomium to Abbé. See Mendelsohn (1996*a*).
6. We include as appendices a discussion of Frege’s symbolism and a rendering into modern notation.
7. We know of at least one article written at the time in explanation of his system that had been submitted for publication and simply rejected.
8. Furth gives the details of the proof in his introduction to Frege (1893: xi–xii).
9. Mendelsohn (1996*a*).

## Chapter 2

1.  $f$  is a singular function, i.e., a function of one argument, but we can speak of binary functions that associate pairs of elements of  $S$  with elements of  $S'$ , and in general, of  $n$ -ary functions that associate  $n$ -tuples of elements of  $S$  with elements of  $S'$ . The generalized property-governing functions would thus be, for any  $\langle x_1, \dots, x_n \rangle$ ,  $\langle y_1, \dots, y_n \rangle$  in the domain of  $g$ :

PRINCIPLE 2.2.2 (GENERALIZED FUNDAMENTAL PROPERTY OF FUNCTIONS) *If  $x_1 = y_1, \dots, x_n = y_n$ , then  $g(x_1, \dots, x_n) = g(y_1, \dots, y_n)$ .*

2. In Sections 2.3 and 2.4, we include single quotes when appropriate even when indenting, in order to avoid use/mention confusion.
3. The function-expression might, of course, be filled by an expression other than a numeral, e.g., a complex expression like  $\langle (2 \cdot 1) + 1 \rangle$ . Inserting  $\langle (2 \cdot 1) + 1 \rangle$ , which designates the number 5, in (2.5) yields the complex expression  $\langle 2 \cdot ((2 \cdot 1) + 1) + 1 \rangle$ , which designates the number 11. Frege, however, goes still further and permits an arithmetic function-expression to be completed by any name whatsoever, e.g., 'Richard Nixon', and in such a case he takes the complex expression thus constructed to designate some arbitrarily chosen element.
4. Both principles can be generalized to function-expressions with more than one argument place. Let  $\theta(\Omega_1, \Omega_2, \dots, \Omega_n)$  be a function-expression with  $n$  argument places ( $\Omega_1, \Omega_2, \dots, \Omega_n$  not necessarily distinct). Then, corresponding to Principles 2.3.1 and 2.3.3, we have, respectively,

PRINCIPLE 2.3.4 (GENERALIZED COMPOSITIONALITY FOR REFERENCE) *For any  $n$ -place function-expression  $\theta(\Omega_1, \Omega_2, \dots, \Omega_n)$  and any names  $\alpha_1, \alpha_2, \dots, \alpha_n$ ,  $r(\theta(\alpha_1, \alpha_2, \dots, \alpha_n)) = r(\theta)[r(\alpha_1), r(\alpha_2), \dots, r(\alpha_n)]$*

and

PRINCIPLE 2.3.5 (GENERALIZED EXTENSIONALITY FOR REFERENCE) *For any  $n$ -place function-expression  $\theta(\Omega_1, \Omega_2, \dots, \Omega_n)$  and any names  $\alpha_1, \alpha_2, \dots, \alpha_n, \beta_1, \beta_2, \dots, \beta_n$ , if  $r(\alpha_1) = r(\beta_1), \dots, r(\alpha_n) = r(\beta_n)$ , then  $r(\theta(\alpha_1, \alpha_2, \dots, \alpha_n)) = r(\theta(\beta_1, \beta_2, \dots, \beta_n))$ .*

5. An example of a relation is the function  $\eta$  *conquered*  $\zeta$ , which yields the value *true* for the argument pair  $\langle \text{Caesar}, \text{Gaul} \rangle$  and the value *false* for the argument pair  $\langle \text{Hannibal}, \text{Rome} \rangle$ .
6. See Dummett (1981*a*, Chapters 4 and 15), also Geach (1962), for a discussion of the difference between names and quantifiers, and some history of the problem in logic.
7. This can be continued and complicated in interesting ways. Some functions take first-level functions as arguments, some take second-level functions, and so on; and some functions yield objects as values, some first-level functions, and so on. There is great flexibility in this structure. We discuss the structure in some detail in Section 5.3.
8. The importance of this distinction did not enter mathematical consciousness seriously until Zermelo's work on the axiomatization of set theory.
9. See the footnote to the discussion of the quantifier in Appendix A.

10. We are getting a bit ahead of ourselves here, but we have in mind the fact that Principle 2.5.2 (*Leibniz's Law*), turns out to be a special case of Principle 2.2.1 (*The Fundamental Property of Functions*).
11. Note that  $S\alpha/\beta$  is not thereby unique.
12. The details are in Section 2.5.
13. Actually, Principle 2.5.4 is still not quite right, because  $S\alpha$  might be about  $r(\alpha)$  and yet contain more than one occurrence of  $\alpha$ , not all of which conform to Principle 2.5.4. For example, again quoting from Quine, "Giorgione was called 'Giorgione' because of his size." But the necessary refinement would take us too far afield at this point, so we will make no further changes to Principle 2.5.4.

### Chapter 3

1. Michael Dummett (1981a: 125–6) also notes that the paradox can be generalized.
2. We are not assuming here that if  $\alpha = \beta$  is informative,  $S\alpha$  and  $S\alpha/\beta$  differ in cognitive value. This would commit us to holding, e.g., that  $(\alpha = \alpha) \wedge (\beta = \alpha)$  differs in cognitive value from  $(\beta = \alpha) \wedge (\beta = \beta)$ , something Frege's guidelines do not clearly imply. The principle we have adopted, and which does seem to capture the spirit of Frege's view, is this: that if  $S\alpha$  and  $S\alpha/\beta$  are interderivable on the assumption that  $\alpha = \beta$  is true, but are not interderivable without it, then if  $\alpha = \beta$  is informative,  $S\alpha$  and  $S\alpha/\beta$  differ in cognitive value. This seems to be a reasonable sufficient condition for two sentences differing in cognitive value; it is an open question whether it is a necessary condition.
3. When Frege said that '=' expressed a relation between the signs themselves, he did not mean that the signs were identical (for that would be blatantly false). He meant that the signs were equivalent in some way.
4. But this is not quite accurate. See the discussion in Chapter 6 where we go into depth about Russell's logical treatment of definite descriptions.
5. This has been particularly evident in confusions about the application of the sense/reference distinction to concept words.
6. "The sense of a proper name is grasped by everybody who is sufficiently familiar with the language or totality of designations to which it belongs; but this serves to illuminate only a single aspect of the *Bedeutung*, supposing it to have one. Comprehensive knowledge of the *Bedeutung* would require us to be able to say immediately whether any given sense attaches to it. To such knowledge we never attain" (Frege 1892c: 153). It is clear that senses too can be presented in different ways, and so even in this case comprehensive knowledge cannot be attained.
7. This is exactly the geometric example – see Section 4.2 – Frege (1879) presents to argue for informative identities. It is clear that we can have immediate acquaintance with a sense, and also that a sense can be presented in different ways.

8. This is the principle that is severely criticized by Putnam (1973). We raise some questions about it in Section 7.6 in connection with the problem of nonreferring singular terms.
9. Frege explicitly rejected Russell's idea that an object could be part of a thought. Why did he think this? Certainly, he believed that some objects, in his technical sense of 'object', were parts of thoughts. The sense of an expression is an object in this technical sense, and it can surely be part of a thought. Perhaps Frege thought that a physical object was the wrong type of object to be a part of a thought. We cannot rule this out as a factor. But it could not be a decisive factor, because Frege also denied that a logical object, like the number 2, was part of the thought expressed by ' $2 + 2 = 4$ '. So it does not appear to be the materiality that is so important. No doubt there must be something intrinsic about the sense that enables it to represent things: that would have to be a constraint. Perhaps, too, there is some further epistemological constraint involved in our being able to *grasp* the object that sanctions its role as a sense. Levine (1998) urges this point. But we have no sure answer to this question.
10. We discuss this further in Section 9.4.
11. These principles are not explicit in the text. They are distilled from the many remarks he makes about sense and reference. Our intention has been to sharpen the principles that infuse Frege's semantic theory so that their effect can be more readily discerned.
12. The argument we presented in Section 2.5 is a good example.

#### Chapter 4

1. The burden of Frege (1892c) was to clarify just this point, viz. how the sense/reference distinction was to be drawn for sentences.
2. The numbering of the sentences in the quotation has been changed to conform with ours.
3. "In our example, accordingly, the *Bedeutung* of the expressions 'the point of intersection of *a* and *b*' and 'the point of intersection of *b* and *c*' would be the same, but not their sense" (Frege 1892c: 152). "The words 'the celestial body most distant from the Earth' have a sense, but it is very doubtful if they also have a *Bedeutung*. The expression 'the least rapidly convergent series' has a sense, but demonstrably there is no *Bedeutung*..." (Frege 1892c: 153).

#### Chapter 5

1. The *Werthverlauf* – *course-of-values* or *value-range* – of a *function* is the more general notion; the *Umfang* – *extension* – of a *concept* is the more specific notion.
2. We offer a suggestion in Section 9.3.
3. Although Frege clearly thought of referring in this way, there is a problem we should note, namely, that his understanding of syntax prevents a referring function from taking both objects and functions as arguments.

4. Nor, again, is the difference between objects and functions to be found in the relation expressions bear to them.
5. This alternative is explored by Furth (1965).
6. The passage continues: “The compound thought must itself be a thought: that is, something either true or false (with no third alternative).” Clearly, the correct story about Frege’s belief in thoughts that lack truth value is extremely complicated. We just note the passage as an example of the difficulty.
7. We discuss Russell’s theory in detail in Chapter 7.
8. In his Introduction to Frege (1893: xxix).
9. Failure to observe this distinction is largely responsible for the contradiction Russell identified in *Grundgesetze*, and which we presented in Chapter 1.
10. There is a hint of this in Frege’s saying that an expression like ‘the predicate “is red”’, by explicitly identifying the predicate, deprives it of its predicative functioning.

### Chapter 6

1. This view is evident in Russell (1917).
2. See Kripke (1980) for a discussion of this view.
3. The old doctrine of distribution appears to have been an attempt to read the quantified phrases in this spirit. See Geach (1962) for a discussion and criticism of this doctrine.
4. Is there an occurrence for which this contextual treatment poses a problem? Yes, at least *prima facie*. When a definite description occurs as part of a more complex designator, there is no simple rule for eliminating it. As an exercise, the reader might contemplate unpacking *the proposition that the King of France is bald*.
5. Kaplan (1972) captures this aspect of Russell’s theory.
6. The term is introduced by Neale (1990).
7. The difference between these two is evident when it comes to quantified modal logic. Whereas Russell’s treatment has been of vital importance to the development of quantified modal logic, Frege’s treatment, and the persistent confusion of Russell’s treatment with Frege’s, has actually had the most unfortunate consequences. See Fitting & Mendelsohn (1998, Chapter 1).
8. In contrast with Russell (1903*b*) and Meinong (1904), both of whom thought they did. See Fitting & Mendelsohn (1998) for a discussion of these views.
9. “‘The King of France is not bald’ is false if the occurrence of ‘the King of France’ is *primary*, and true if it is *secondary*. Thus all propositions in which ‘the King of France’ has a primary occurrence are false; the denials of such propositions are true, but in them ‘the King of France’ has a secondary occurrence.” Russell (1905: 53)
10. It is important to bear in mind that Russell is not appealing to Frege’s Substitution Principle 2.5.1. If you look at the wording of the puzzle, you will note that he is substituting one term for another *inside the proposition*, which means, in effect, that his substitution is being carried out at Frege’s level of sense. At that level, of course, there is no term in the proposition

corresponding to the description ‘the author of *Waverley*’, and the substitution is blocked.

11. This appeal to the scope distinction is somewhat different from the appeal in the other two puzzles because it operates in a context that is not truth functional. On the large-scope reading, the identification of the object is external to the proposition believed; on the small-scope reading, the identification of the object is internal to the proposition believed. For the other two cases, even though there is a scope distinction to be drawn, the identification of the object is in each case internal to the proposition expressed. Our terminology is similar to, but our interpretation is quite different from, Forbes (1987).
12. But not exactly the same. Russell’s framing of the paradox appears to appeal to Frege’s Substitution Principle for Reference 2.5.1, but he is really appealing to Frege’s *Begriffsschrift* Substitution Principle 3.3.1. His substitution is being carried out, as noted in the previous note, at Frege’s level of sense.
13. They do not, of course, agree on the *Bedeutung* of declarative sentences. Russell does not believe that the notion of *Bedeutung* applies to sentences, only to names and descriptions. See, e.g., the letter we quote from at the beginning of Section 8.6.
14. ‘The round square’ is a meaningful expression, so there must be such an object as this round square.
15. For a full discussion of this paradox, see Fitting & Mendelsohn (1998).
16. Frege does not appear to have had as clear a connection between syntax and semantics as Russell did. Nowhere is this more evident than in the handling of identity in Frege (1879) as well as in Frege (1892c). Although he proclaims identities informative, and cites the need to prove identities as reason to include an identity sign in logic, he feels no need to provide any syntactic machinery for representing these proofs.
17. This corrects the characterization in Fitting & Mendelsohn (1998: 250).
18. The discussion is found in Frege (1893, Sections 28–31). His ‘proof’, of course, fails: *Grundgesetze* is inconsistent, as Russell later showed.
19. Russell’s treatment offers no comfort here. On his view, both turn out to be false.
20. For a discussion of these issues, see Fitting & Mendelsohn (1998).

## Chapter 7

1. For discussion of this principle, see Wright (1983) and Dummett (1991).
2. Bennett (1974) actually speaks of a “Kant-Frege view on existence.”
3. There is a second difference in the presentation of the two paradoxes. For the Paradox of Identity, the direct reference assumption is salient, because the difference between  $a = a$  and  $b = b$  is important. For the Paradox of Nonbeing, both  $a$  does not exist and  $b$  does not exist appear self-defeating, and the issue of the difference in informativeness between the two pales by comparison.

In the literature, Nonbeing is a paradox but Identity is a puzzle. For the reasons just noted, this distinction is unwarranted. We consider both arguments to be paradoxes.

4. This is especially true of mathematical practice.
5. Salmon (1998: 253–4) notes this *Begriffsschrift*-like analysis, and cites other, later, passages in which Frege adopts this view.
6. This is not to imply any essential connection in classical logic between existence and identity. To define ‘ $x$  exists’ we only require a predicate true of everything. ‘ $Fx \vee \neg Fx$ ’, for example, supplies us with the definition

$$x \text{ exists} \leftrightarrow Fx \vee \neg Fx$$

This accords nicely with intuitions about existence: to say that something exists is to say that something is true of it, i.e., that either ‘ $F$ ’ is true of it or ‘ $\neg F$ ’ is true of it.

7. Or, as in (7.1), ‘something’.
8. Lockwood (1975) gleaned from Frege’s (1892*b*) justification for the distinction between an ‘is’ of identity and an ‘is’ of predication yet another redundancy theory, which we might call the Redundancy Theory of Identity:

*Where  $a$  and  $b$  are singular terms and ‘is’ the usual ‘is’ of predication,  $a$  is identical with  $b$  if, and only if,  $a$  is  $b$ .*

See Mendelsohn (1987) for a discussion of this view.

9. In Frege’s terminology, this is to regard ‘the present King of France’ as an *Eigenname*. In Russell’s terminology, this is to regard it as a genuine or logically proper name.
10. See Russell (1917) for a discussion of this notion.
11. See Fitting & Mendelsohn (1998) for details of this analysis of the paradox.
12. Of course, he did provide arguments elsewhere. See, e.g., Russell (1918: 232–4).
13. It isn’t, of course.
14. “Since nothing falls under the concept ‘not identical with itself’, I define nought as follows:

o is the Number which belongs to the concept ‘not identical with itself’.”

(Frege 1884*b*: 87).

15. See, e.g., Perry (1970 and 1978).
16. The notion of number plays no front-stage role in this argument. The analogy is directly between identity and existence. If we are right, however, the story about cardinal number needs to be revised.

It is not clear whether Bennett ever noted the conflict between his position in Bennett & Alston (1984) about number, on the one hand, and his approval in Bennett (1974) of the Fregean position about existence.

17. Frege (1884*a*: 65–6) sums up:

In general one can lay down the following: If you want to assign a content to the verb ‘to be’, so that the sentence ‘ $A$  is’ is not pleonastic and self-evident, you will have to allow circumstances under which the negation of ‘ $A$  is’ is possible; that is to say, that there are subjects of which being must be denied. But in that case the concept ‘being’ will no longer be suitable for providing a general explanation of ‘there are’ under which ‘there are  $B$ s’ means the same as ‘something that has being falls under the concept  $B$ ’; for if we apply this explanation to ‘There are subjects of which being

must be denied', then we get 'Something that has being falls under the concept of not-being' or 'Something that has being is not'. There is no way of getting over this once a content of some kind – it doesn't matter what it is – is agreed to the concept of being. If the explanation of 'there are *Bs*' as meaning the same as 'Something that has being is *B*' is to work, we just have to understand by being something that goes entirely without saying.

18. Frege finally did make an amendment to syntax and introduce a definite description operator in *Grundgesetze* Section 11, which we gave above as Definition 6.5.1. But, as we noted in Section 6.5, it is still an *Eigenname*, and so it fails to get around the problems with singular denials of existence we have been detailing all these pages.
19. The story is not this stable: there is a certain amount of waffling about whether these concepts are denoted or expressed, and whether they are intensional or extensional.
20. Here is how he continues the paragraph:

This is an important precursor of the view of Frege that any legitimate existential statement must be built out of propositional atoms of the form 'There is an *F*', where *F* stands for a determining predicate. According to this Kant-Frege view, the real form of 'Tigers exist' is not like that of 'Tigers growl', but rather like that of 'There are tigers', or 'The concept of tigerhood is instantiated'. Granted that Kant's arguments fall far short of proving this hypothesis, they do at least illustrate and elucidate it; and the hypothesis itself is a philosophical contribution which deserves attention and which may even be true.

## Chapter 8

1. "What is a fact? A fact is a thought that is true" (Frege 1918: 342).
2. It is in this context that Frege (1892*c*) introduced the example of the sentence 'Odysseus was set ashore at Ithaca while sound asleep' that lacked a truth value because it contained a name, 'Odysseus', that failed to refer to anything. Frege took this as *confirming* his claim that it is the truth value of a sentence that is compositionally related to the reference of its constituent singular term.
3. Gödel (1944) is the source of all these sharpenings.
4. Unless one thinks that definite descriptions are not referring expressions, something that never occurred to Frege because definite descriptions are the paradigm of complex referring expression on which he based his logical grammar and its metaphysical connection to objects and properties. But cf. Gödel (1944), Neale (1995), and Donaho (1998).
5. This confirms that he regarded the identification of the truth values as referents of sentences as the main result of the essay.
6. We develop this point in Section 10.1.
7. We provide these details in Section 9.2.
8. This is just because the proper name '9' is taken to designate the same object in every possible world, while the description 'the number of the planets' is not. The description might designate different numbers in different possible worlds, because there might have been a different number of planets than there in fact is.



9. The great virtue of Russell's technical treatment of definite descriptions is that it provides a syntax for marking this distinction. Fitting & Mendelsohn's (1998) predicate abstract notation is a bit easier to read. The *de re* reading of (8.15) is  $\langle \lambda x. \Box(x > 7) \rangle(9)$ .  $\Box$  attaches to the predicate  $x > 7$  to form the complex predicate  $\Box(x > 7)$ , which is applied to the name. (We affirm of the number 9 that it has the property *being necessarily greater than 7*.) The *de dicto* reading of (8.15) is  $\Box(\lambda x. x > 7)(9)$ .  $\Box$  attaches to the closed sentence  $\langle \lambda x. x > 7 \rangle(9)$ . (We affirm it is necessary *that* the number 9 is greater than 7.)
10. For example, on the *de re* reading, (8.12) cannot follow from (8.11) in the manner indicated in Quine's argument. If ' $q$ ' is true in this world, ' $\{x|x = x \wedge q\}$ ' in this world designates  $\{x|x = x\}$  and this object is, in every possible world, identical with  $\{x|x = x\}$ . But this does not mean that ' $q$ ' is true in every possible world. If the terms were both *rigid designators* – and ' $\{x|x = x \wedge q\}$ ' would be rigid if ' $q$ ' were necessarily true – the *de re* reading would be logically equivalent to the *de dicto* reading. But the argument is constructed with a nonrigid designator, ' $\{x|x = x \wedge q\}$ '.
11. For further discussion of these matters, see Fitting & Mendelsohn (1998).
12. See Fitting & Mendelsohn (1998, Chapter 1) for discussion of this point.
13. Not quite, of course, since Frege admits sentences that lack a reference. But we will not worry about these cases here.
14. This point has also been noted by Hochberg (2003: 182–6).
15. Church (1956) sets forth propositional logic with the two primitives, ' $\supset$ ' (for *if, then*) and ' $\perp$ ' (for *false*), and then defines ' $\neg p$ ' as ' $p \supset \perp$ '.
16. Frege's account of judging is largely metaphorical. Frege (1892c: 159) says that "judgments can be regarded as advances from a thought to a truth value." On another occasion, Frege (1918: 513) characterizes a judgment as "the recognition of the truth of a thought."
17. This is in Frege (1892b). For a discussion, see Mendelsohn (1978).
18. This holds for Frege. Russell, who is not similarly committed to the argument in Section 8.2, is not compelled to accept the truth values as object. His view is, as one would expect, quite different from Frege's. For Russell, 'true' is a concept word, and *truth* is to be explained in terms of *correspondence*.
19. This would indicate, however, that (8.24) says that the thought *stands for* the True. There is a small problem that should be noted. Recall Frege's example of a sentence that expresses a thought even though it lacks a truth value,

Odysseus was set ashore at Ithaca while sound asleep. (8.28)

If (8.28) lacks a truth value, then so must

It is true that Odysseus was set ashore at Ithaca  
while sound asleep. (8.29)

Now, if we understand (8.29) to be an identity,

Odysseus was set ashore at Ithaca while sound  
asleep = the True. (8.30)

then we can easily affirm that (8.30) lacks truth value. The constituent sentence (8.28), by hypothesis, lacks a reference, and so the complex in which it is embedded, viz. (8.30), must also lack a reference. But if we understand (8.29) to ascribe a property to the thought expressed by (8.28), i.e., as

That Odysseus was set ashore at Ithaca while sound asleep  
denotes the True, (8.31)

then we shall have to say that (8.29) is false. The sentence is meaningful, and so expresses a thought, but it does not designate a truth value. Hence these two do not quite yield the same analysis.

20. This is the source of one version of what has come to be known as *the Redundancy Theory of Truth*, which goes roughly as follows: to say *It is true that p* is to say no more nor less than just *p*, and to say *It is false that p* is to say no more nor less than just  $\neg p$ . Incidentally, this is not to say that Frege endorsed any particular definition of truth: his view remained that truth is indefinable.
21. For further discussion, see Ricketts (1986).

### Chapter 9

1. A given reference can be picked out by different senses, so there is no function that takes us from a reference to *the* sense that picks it out.
2. We have borrowed much from this analysis of the hierarchy.
3. Parsons (1981), for example, counts himself in the same camp. Forbes (1987) contrasts his own position with Dummett's, but counts himself within the broader camp of rigid hierarchists.
4. This is Dummett's argument. Note its ancestor in Church's (1954) point we discussed in Section 3.6.
5. A number of commentators, in particular, Hylton (1990) and Kremer (1994), have given us reason to believe that Russell (1905) had identified a problem for his own semantic theory of the same name because of the demands of direct reference and the Principle of Acquaintance. In this regard, we should also call the reader's attention to Kaplan (1989, footnote 23). The reader might wish to look at the manuscripts "On Fundamentals," "On Meaning and Denotation," "On the Meaning and Denotation of Phrases," and "Points About Denoting," which have been published in Urquhart (1994), in which we find Russell struggling with a sense/reference distinction of his own before coming up with his 1905 Theory of Descriptions. The argument of the present chapter, however, renders Russell's claim that he is criticizing Frege's theory much more plausible.
6. This way of diagramming Frege's semantic theory is borrowed from Parsons (1981).
7.  $s_i(t)$  and  $r_i(t)$  will, then, be the appropriate sense and reference, respectively, of  $t$  when embedded in  $i$  'that' clauses.
8. We have mentioned this many times in the text. See also Dummett (1981a: 158–9).

9. Dummett (1981*a*) argues for this view. Parsons (1981), who takes the sense of a function-expression to be a *function*, and not just incomplete in some way analogical to the reference of function-expressions, does not appear to hold this view. The structured proposition view has been ably argued by Richard (1990); the idea that propositions are unstructured has been argued by Stalnaker (1984).
10. It is instructive in this regard to recall some comments in Searle's (1957: 342) very influential defense of Frege against Russell's (1905) objections to the sense/reference theory:

Russell's arguments suffer from unclarity and minor inconsistencies throughout and I have tried to restate them in a way which avoids these. But even in their restated form, they are faulty. Their faults spring from an initial mis-statement of Frege's position, combined with a persistent confusion between the notions of *occurring as a part of a proposition* (being a constituent of a proposition) and *being referred to by a proposition*. The combination of these two leads to what is in fact a denial of the very distinction Frege is trying to draw and it is only from this denial, not from the original thesis, that Russell's conclusions can be drawn.

Searle is quite right to underscore Frege's desire to distinguish *being a constituent of a proposition* and *being referred to by a (part of a) proposition*. But, as we have argued, it is not obvious that Frege had successfully made the distinction when propositions themselves were the subject of discourse. These are just the cases that exercised Russell, and it is no misreading on his part to point this out. Russell, after all, believed that one cannot treat a sentence as a *name*, but only as expressing a proposition.

11. We have changed the references to conform to our own numbering of the examples and the format of the text to the original German edition. Linsky (1983) and Parsons (1981) have drawn our attention to this passage.
12. See Linsky (1967) for a criticism of Carnap.
13. See Burge (1979), Parsons (1981), and Forbes (1987) for good discussions of the principles needed for the hierarchy.
14. Kripke (1980) calls a term that designates the same object in every possible world in which the object exists a 'rigid designator'. See Fitting & Mendelsohn (1998, Chapter 10) for further discussion of this notion.
15. There are problems understanding how Principle 9.6.1 is supposed to work if we understand expressions like *the sense of t* and *the indirect sense of t* as function-expressions, let alone function-expressions that iterate. Are we to suppose that

the sense of 'the sense of t' = the sense of t?

But, this says that t and 'the sense of t' have the same sense. This cannot be. If they had the same sense, they would have the same reference. But, let t be 'Giorgione was so-called because of his size'. The reference of t is a truth value; the reference of 'the sense of t' is a proposition. Are we to suppose that

the sense of the sense of t = the sense of t?

But, as we remarked earlier, a sense can have more than one sense. Let  $t$  be ‘Giorgione was so-called because of his size’. Now the sense of  $t$  will be a proposition, *the proposition that Giorgione was so-called because of his size*. But there is no unique object that is the sense of this (i.e., the sense of the sense of  $t$ ): this is why (9.12) and (9.13) differ in cognitive value.

16. The steps in the argument are more readily accessible when put into symbols:

$$‘Fa’ \text{ expresses } \Theta Fa \quad (9.41)$$

$$‘Fa’ \text{ expresses } \Theta F\Theta a \quad (9.42)$$

$$‘F\Theta a’ \text{ expresses } \Theta F\Theta a \quad (9.43)$$

$$‘F\Theta a’ \text{ expresses } \Theta F\Theta\Theta a \quad (9.44)$$

$$\Theta\Theta\alpha = \Theta\alpha \quad (9.45)$$

Now (9.41) and (9.43) are truisms; (9.42) and (9.44) are each sanctioned by Principle 9.2.3; (9.45) is sanctioned by Principle 9.6.1. By (9.45), the items to the right of ‘expresses’ are the same in (9.42) and (9.44), so the items on the left, in each case – which become (9.39) and (9.40) – express the same proposition.

This argument has an ancestor in Burge (1979). See Parsons (1991).

## Chapter 10

1. We have changed the numbering of the examples to conform with our own.
2. See, e.g., Mendelson (1987).
3. Kripke (1975: 693) makes this point.

Let ‘Jack’ be a name of the sentence ‘Jack is short’, and we have a sentence that says of itself that it is short. I can see nothing wrong with “direct” self-reference of this type. If ‘Jack’ is not already a name in the language, why can we not introduce it as a name of any entity we please? In particular, why can it not be a name of the (uninterpreted) finite sequence of marks ‘Jack is short’? (Would it be permissible to call this sequence of marks “Harry,” but not “Jack”? Surely prohibitions on naming are arbitrary here.) There is no vicious circle in our procedure, since we need not *interpret* the sequence of marks ‘Jack is short’ before we name it. Yet if we name it “Jack,” it at once becomes meaningful and true.

4. The numbering of the sentences has been changed to conform with our own.
5. This is why Davidson (1979: 81) calls Quine’s the *Proper Name Theory*: “a quotation, consisting of an expression flanked by quotation-marks, is like a single word, and is to be regarded as logically simple.”
6. Searle, however, for reasons we find unconvincing, denies that this presenting itself is to be regarded as naming itself. Searle claims that the only reason for naming something is if it is not present, and that it therefore makes no sense to name something that is presenting itself. This does not seem right, however. We are reminded of Geach’s vivid translation for Frege’s *Begriffsschrift* view of identity: the names flanking an identity sign appeared *in propria persona* (Black & Geach 1952: 56).

7. The Identity Theory has also been championed recently by Washington (1992).
8. See Bennett (1980) for discussion and references.
9. We have made ever-so-slight changes in the diagram to make it just a bit more perspicuous and in conformity with the diagramming in Parsons (1981).
10. There is an oddity in this view. The name “Boston” names ‘Boston’, and it does so *twice*: the quoted interior stands for itself (since it is occurring autonomously) *and* the quoted interior plus quotation marks is standing for the quoted interior. But there are deeper problems to plumb.
11. Each of these sets of elements of the vocabulary is pairwise disjoint. So, in particular, the left and right quotation marks are not among the function symbols.
12. This is adapted from Richard (1986: 360). Our discussion of quotation-names has been strongly influenced by this article.
13. We drop the sequencing braces and just write the expression as per usual.

### Appendix A

1. In the preface, Frege notes that he could reduce the number of his axioms further by combining (31) and (41) as a biconditional. But this sort of reduction carries no intellectual interest, for it is the number of axioms he is after, not the quality, viz. the independence. Note that (8) is derivable from the rest.
2. They are not clearly quantificational as opposed to truth functional because Frege used the identity sign both for identity and the biconditional.
3. This is evident from the reviews of the work, some of which have been included in Bynum’s (1972) edition of *Begriffsschrift*.
4. Frege (1879: 69) says that “this signifies the judgment that the function is a fact whatever may be taken as its argument.” In particular, on that same page, he says “if the Gothic letter occurs as a functional symbol, account must be taken of this circumstance.” In effect, then, his proposition 58

$$\begin{array}{l} \vdash \text{---} f(c) \\ \quad \vdash \text{---} f(a) \end{array}$$

i.e., Universal Instantiation, is supposed to hold for second-order quantifiers as well.



## Bibliography

- Anderson, C. A. (1980), "Some New Axioms for the Logic of Sense and Denotation: Alternative (o)," *Noûs* **14**, 217–34.
- Arnauld, A. (1662), *La Logique, ou L'Art de Penser*. Translated by James Dickoff and Patricia James as *The Art of Thinking*, Bobbs-Merrill Company, Indianapolis, 1994.
- Baker, G., & Hacker, P. M. S. (1984), *Frege: Logical Excavations*, Oxford University Press, New York.
- Baumann (1868), *Die Lehren Von Zeit, Raum und Mathematik*, 2 Vols., Berlin.
- Beaney, M., ed. (1997), *The Frege Reader*, Blackwell, Oxford.
- Bennett, J. (1974), *Kant's Dialectic*, Cambridge University Press, Cambridge.
- Bennett, J. (1980), "Quotation," *Noûs* **22**, 399–418.
- Bennett, J., & Alston, W. (1984), "Identity and Cardinality: Geach and Frege," *Philosophical Review* **93**, 553–68.
- Black, M. (1954), "Frege on Functions," in *Problems of Analysis*, Cornell University Press, Ithaca, pp. 229–54. Reprinted in Klemke (1968), pp. 223–48.
- Black, M., & Geach, P. (1952), *Translations from the Philosophical Writings of Gottlob Frege*, Blackwell, Oxford.
- Boolos, G. (1985), "Reading the *Begriffsschrift*," *Mind* **94**, 331–44.
- Burge, T. (1979), "Frege and the Hierarchy," *Synthese* **40**, 265–81.
- Bynum, T. W., ed. (1972), *Conceptual Notation and Related Articles*, Clarendon Press, Oxford.
- Carnap, R. (1937), *The Logical Syntax of Language*, Routledge & Keegan Paul, London.
- Carnap, R. (1947), *Meaning and Necessity: A Study in Semantics and Modal Logic*, University of Chicago Press, Chicago.
- Chomsky, N. (1957), *Syntactic Structures, Janua Linguarum IV*, Mouton, The Hague.
- Church, A. (1943), "Review of Carnap, *Introduction to Semantics*," *Philosophical Review* **52**, 298–304.

- Church, A. (1950), "On Carnap's Analysis of Statements of Assertion and Belief," *Analysis* **10**, 97–9.
- Church, A. (1951), "A Formulation of the Logic of Sense and Denotation," in H. M. K. P. Henle & S. K. Langer, eds., *Structure, Method, and Meaning: Essays in Honor of Henry M. Sheffer*, Liberal Arts Press, New York.
- Church, A. (1954), "Intensional Isomorphism and Identity of Belief," *Philosophical Studies* **5**, 65–73.
- Church, A. (1956), *Introduction to Mathematical Logic*, Vol. 1, Princeton University Press, Princeton.
- Church, A. (1973), "Outline of a Revised Formulation of the Logic of Sense and Denotation, I," *Noûs* **7**, 24–33.
- Davidson, D. (1968–9), "On Saying That," *Synthese* **19**, 130–46. Reprinted in Davidson (1984), pp. 93–108.
- Davidson, D. (1969), "True to the Facts," *Journal of Philosophy* **66**, 748–64. Reprinted in Davidson (1984), pp. 37–54.
- Davidson, D. (1979), "Quotation," *Theory and Decision* **11**, 27–40. Reprinted in Davidson (1984), pp. 79–92.
- Davidson, D. (1984), *Inquiries into Truth and Interpretation*, Oxford University Press, Oxford.
- Descartes, R. (1641), *Meditations on First Philosophy*, in R. S. Haldane & G. R. T. Ross, eds., *The Philosophical Works of Descartes*, Vol. 1, Cambridge University Press, Cambridge, 1911, pp. 131–200.
- Donaho, S. (1998), "Are Declarative Sentences Representational?" *Mind* **107**, 33–58.
- Dummett, M. (1955), "Frege on Functions: A Reply," *Philosophical Review* **64**, 96–107. Reprinted in Klemke (1968), pp. 268–83.
- Dummett, M. (1981a), *Frege, Philosophy of Language*, 2d Edition, Harvard University Press, Cambridge.
- Dummett, M. (1981b), *The Interpretation of Frege's Philosophy*, Harvard University Press, Cambridge.
- Dummett, M. (1991), *Frege: Philosophy of Mathematics*, Harvard University Press, Cambridge.
- Engel, S. M. (1963), "Kant's 'Refutation' of the Ontological Argument," *Philosophy and Phenomenological Research* **23**, 20–35.
- Evans, G. (1982), *The Varieties of Reference*, Oxford University Press, Oxford.
- Fitting, M. & Mendelsohn, R. L. (1998), *First-Order Modal Logic*, Kluwer, Dordrecht.
- Forbes, G. (1987), "Indexicals and Intensionality," *Philosophical Review* **96**, 3–33.
- Frege, G. (1879), *Begriffsschrift, Eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*, Verlag von L. Nebert, Halle. Preface and Part 1 translated in Beaney (1997), pp. 47–78.
- Frege, G. (1884a), "Dialog mit Pünjer über Existenz," in Hermes et al. (1969), pp. 60–75. Translated as "Dialogue with Pünjer on Existence" in Frege (1979), pp. 53–67.
- Frege, G. (1884b), *Die Grundlagen der Arithmetik, Eine logisch-mathematische Untersuchung über den Begriff der Zahl*, Verlag von Wilhelm Koenner, Breslau. Translated by J. L. Austin as *The Foundations of Arithmetic: A Logico-Mathematical Enquiry into the Concept of Number*, Blackwell, Oxford, 1959.



- Frege, G. (1891), "Funktion und Begriff," Hermann Pohle, Jena. Translated as "Function and Concept" in Beaney (1997), pp. 130–48.
- Frege, G. (1892a), "Review of G. Cantor: Zur Lehre vom Transfiniten," *Zeitschrift für Philosophie und philosophische Kritik* **100**, 269–72.
- Frege, G. (1892b), "Über Begriff und Gegenstand," *Vierteljahrsschrift für wissenschaftliche Philosophie* **16**, 195–205. Translated as "On Concept and Object" in Beaney (1997), pp. 181–93.
- Frege, G. (1892c), "Über Sinn und Bedeutung," *Zeitschrift für Philosophie und philosophische Kritik* **100**, 25–50. Translated as "On Sinn and Bedeutung" in Beaney (1997), pp. 151–71.
- Frege, G. (1893), *Grundgesetze der Arithmetik*, Vol. 1, Pohle, Jena. Partially translated by Montgomery Furth as *The Basic Laws of Arithmetic: Exposition of the System*, University of California Press, Berkeley and Los Angeles, 1964.
- Frege, G. (1894), "Review of E. Husserl: Philosophie der Arithmetik," *Zeitschrift für Philosophie und philosophische Kritik* **103**, 313–32. Extracts translated in Beaney (1997), pp. 224–5.
- Frege, G. (1906), "Einleitung in die Logik," in Hermes et al. (1969), pp. 201–12. Translated as "Introduction to Logic" in Frege (1979), pp. 185–96.
- Frege, G. (1915), "Meine grundlegenden logischen Einsichten," in Hermes et al. (1969), pp. 271–2. Translated as "My Basic Logical Insights" in Frege (1979), pp. 251–2.
- Frege, G. (1918), "Der Gedanke: Eine logische Untersuchung," *Beiträge zur Philosophie des deutschen Idealismus* **1**, 58–77. Translated as "Thought" in Beaney (1997), pp. 325–45.
- Frege, G. (1923), "Logische Untersuchungen, Dritter Teil: Gedankengefüge," *Beiträge zur Philosophie des deutschen Idealismus* **3**, 36–51. Translated as "Compound Thoughts" in Klemke (1968), pp. 537–58.
- Frege, G. (1971), "On the Foundations of Geometry," in E.-H. W. Kluge, ed., *Gottlob Frege, On the Foundations of Arithmetic and Formal Theories of Geometry*, Yale University Press, New Haven.
- Frege, G. (1979), *Gottlob Frege: Posthumous Writings*, University of Chicago Press, Chicago. Translation by Peter Long and Roger White of Hermes et al. (1969).
- Frege, G. (1980), *Philosophical and Mathematical Correspondence*, University of Chicago Press, Chicago. Translation by Hans Kaal of an abridged version of Gabriel et al. (1976).
- Furth, M. (1965), "Two Types of Denotation," in N. Rescher, ed., *Essays in Logical Theory*, American Philosophical Quarterly Monograph, 1965.
- Gabriel, G., Hermes, H., Kambartel, F., Thiel, C., & Veraart, A., eds. (1976), *Wissenschaftlicher Briefwechsel, Meiner, Hamburg*.
- Gabriel, G., & Kienzler, W., eds. (1997), *Frege in Jena. Beiträge zur Spurensicherung*, *Kritisches Jahrbuch der Philosophie*, Vol. 2, Königshausen & Neumann, Würzburg.
- Geach, P. (1962), *Reference and Generality*, Cornell University Press, Ithaca.
- Geach, P. (1967), "Identity," *Review of Metaphysics* **21**, 3–12.
- Gödel, K. (1944), "Russell's Mathematical Logic," in P. A. Schilpp, ed., *The Philosophy of Bertrand Russell*, Tudor Publishing, New York, pp. 123–54. Reprinted in H. Putnam and P. Benacerraf, eds., *Philosophy of Mathematics, Selected Readings*, Prentice-Hall, Englewood Cliffs, 1964, pp. 211–32.

- Heidelberger, H. (1975), "Review of Dummett, *Frege, Philosophy of Language*," *Metaphilosophy* **6**, 35–43.
- Hermes, H., Kambartel, F., & Kaulbach, F., eds. (1969), *Gottlob Frege: Nachgelassene Schriften*, Felix Meiner, Hamburg.
- Hochberg, H. (2003), *Introducing Analytic Philosophy: Its Sense and Nonsense, 1879–2002*, Dr. Hansel-Hohenhausen, Frankfurt.
- Hylton, P. (1990), *Russell, Idealism and the Emergence of Analytic Philosophy*, Clarendon Press, Oxford.
- Kant, I. (1781), *Immanuel Kant's Critique of Pure Reason*. Translated by Norman Kemp Smith, Macmillan, London, 1964.
- Kaplan, D. (1969), "Quantifying In," in D. Davidson & J. Hintikka, eds., *Words and Objections: Essays on the Work of W. V. Quine*, D. Reidel, Dordrecht, pp. 178–214. Reprinted in Linsky (1971), pp. 112–44.
- Kaplan, D. (1972), "What Is Russell's Theory of Descriptions?" in D. F. Pears, ed., *Bertrand Russell: A Collection of Critical Essays*, Anchor Books, Garden City, pp. 227–44.
- Kaplan, D. (1989), "Demonstratives," in J. Almog, J. Perry, & H. Wettstein, eds., *Themes from Kaplan*, Oxford University Press, Oxford, pp. 481–564.
- Klemke, E. D., ed. (1968), *Essays on Frege*, University of Illinois Press, Urbana.
- Kneale, W. & Kneale, M. (1962), *The Development of Logic*, Clarendon Press, Oxford.
- Kreiser, L. (2001), *Gottlob Frege, Leben, Werk, Zeit*, F. Meiner, Hamburg.
- Kremer, M. (1994), "The Argument of 'On Denoting'," *Philosophical Review* **103**, 249–98.
- Kripke, S. (1975), "Outline of a Theory of Truth," *Journal of Philosophy* **72**, 690–716.
- Kripke, S. (1980), *Naming and Necessity*, Harvard University Press, Cambridge.
- Levine, J. (1998), "Acquaintance, Denoting Concepts, and Sense," *Philosophical Review* **107**, 415–45.
- Linsky, L. (1967), *Referring*, Humanities Press, New York.
- Linsky, L. (1983), *Oblique Contexts*, University of Chicago Press, Chicago.
- Linsky, L., ed. (1971), *Reference and Modality*, Oxford University Press, London.
- Lockwood, M. (1975), "On Predicating Proper Names," *Philosophical Review* **84**, 471–98.
- Marsh, R. C., ed. (1956), *Logic and Knowledge: Essays 1901–1950 by Bertrand Russell*, Allen & Unwin, London.
- Marshall, W. (1953), "Frege's Theory of Functions and Objects," *Philosophical Review* **62**, 374–90. Reprinted in Klemke (1968), pp. 249–67.
- Mates, B. (1952), "Synonymity," in L. Linsky, ed., *Semantics and the Philosophy of Language*, University of Illinois Press, Urbana, pp. 111–38.
- Mates, B. (1961), *Stoic Logic*, University of California Press, Berkeley and Los Angeles.
- Meinong, A. (1904), "On The Theory of Objects." Translated in R. M. Chisholm, ed., *Realism and the Background of Phenomenology*, Free Press, New York, 1960, pp. 76–117.
- Mendelsohn, R. L. (1978), "Frege on Predication," *Midwest Studies in Philosophy* **3**, 69–82.

- Mendelsohn, R. L. (1987), "Frege's Two Senses of 'Is'," *Notre Dame Journal of Formal Logic* **28**, 139–60.
- Mendelsohn, R. L., tr. (1996*a*), "Diary: Written by Professor Dr Gottlob Frege in the Time from 10 March to 9 May 1924," *Inquiry* **39**, 303–42.
- Mendelsohn, R. L. (1996*b*), "Frege's Treatment of Indirect Reference," in M. Schirn, ed., *Frege: Importance and Legacy*, de Gruyter, Berlin, pp. 410–37.
- Mendelson, E. (1987), *Introduction to Mathematical Logic*, 3d Edition, Wadsworth & Brooks/Cole, Pacific Grove.
- Mill, J. S. (1843), *A System of Logic, Ratiocinative and Inductive, Being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation*, London.
- Moore, G. E. (1936), "Is Existence a Predicate?" *Proceedings of the Aristotelian Society*, Supplementary Vol. XV, 275–88. Reprinted in *Philosophical Papers*, Allen & Unwin, London, pp. 114–25.
- Neale, S. (1990), *Descriptions*, Blackwell, Oxford.
- Neale, S. (1995), "The Philosophical Significance of Gödel's Slingshot," *Mind* **104**, 761–825.
- Ockham, William of (1323), *Summa Logicae*. Part 1 translated by Michael J. Loux as *Ockham's Theory of Terms*, University of Notre Dame Press, Notre Dame.
- Parsons, T. (1981), "Frege's Hierarchies of Indirect Senses and the Paradox of Analysis," *Midwest Studies in Philosophy* **6**.
- Parsons, T. (1982), "What Do Quotation Marks Name? Frege's Theories of Quotations and That-Clauses," *Philosophical Studies* **42**, 315–28.
- Parsons, T. (1991), "Frege's Hierarchies of Indirect Sense." Handout at the Frege Conference, Munich, 1991.
- Perry, J. (1970), "The Same F," *Philosophical Review* **79**, 181–200.
- Perry, J. (1978), "Relative Identity and Number," *Canadian Journal of Philosophy* **8**, 1–14.
- Putnam, H. (1973), "Meaning and Reference," *Journal of Philosophy* **70**, 699–711. Reprinted in Schwartz (1977), pp. 119–32.
- Quine, W. V. O. (1948), "On What There Is," *Review of Metaphysics* **2**. Reprinted in Quine (1961), pp. 1–19.
- Quine, W. V. O. (1951), *Mathematical Logic*, Revised Edition, Harper & Row, New York.
- Quine, W. V. O. (1953*a*), "Reference and Modality," in Quine (1961), pp. 136–59.
- Quine, W. V. O. (1953*b*), "Three Grades of Modal Involvement," in *The Ways of Paradox and Other Essays*, Random House, New York, pp. 156–74.
- Quine, W. V. O. (1960), *Word and Object*, MIT Press, Cambridge.
- Quine, W. V. O. (1961), *From a Logical Point of View*, 2d Revised Edition, Harper & Row, New York.
- Ramsey, F. P. (1931), *The Foundations of Mathematics*, Methuen, London and New York.
- Reichenbach, H. (1947), *Elements of Symbolic Logic*, Macmillan, New York.
- Richard, M. (1986), "Quotation, Grammar, and Opacity," *Linguistics and Philosophy* **9**, 383–403.
- Richard, M. (1990), *Propositional Attitudes: An Essay on Thoughts and How We Ascribe Them*, Cambridge University Press, Cambridge.

- Ricketts, T. G. (1986), "Objectivity and Objecthood: Frege's Metaphysics of Judgment," in J. Hintikka & L. Haaparanta, eds., *Frege Synthesized*, D. Reidel, Dordrecht, pp. 65–96.
- Russell, B. (1903*a*), Correspondence with G. Frege. Translated in Gottfried Gabriel et al, eds., *Gottlob Frege, Philosophical and Mathematical Correspondence*, University of Chicago Press, Chicago, 1980.
- Russell, B. (1903*b*), *The Principles of Mathematics*. Reprinted by Norton, New York, 1938.
- Russell, B. (1905), "On Denoting," *Mind* **14**, 479–93. Reprinted in Marsh (1956), pp. 39–56.
- Russell, B. (1917), "Knowledge by Acquaintance and Knowledge by Description (1911)," in *Mysticism and Logic*, Allen & Unwin, London, pp. 152–67.
- Russell, B. (1918), "The Philosophy of Logical Atomism." Reprinted in Marsh (1956), pp. 175–282.
- Russell, B. (1919), *Introduction to Mathematical Philosophy*, Allen & Unwin, London.
- Ryle, G. (1932), "Systematically Misleading Expressions," *Proceedings of the Aristotelian Society* **32**, 139–70. Reprinted in G. Ryle, *Collected Papers*, Vol. 2, pp. 39–62, Barnes & Noble, New York, 1971.
- Salmon, N. (1998), "Nonexistence," *Noûs* **32**, 277–319.
- Schwartz, S. P., ed. (1977), *Naming, Necessity, and Natural Kinds*, Cornell University Press, Ithaca.
- Searle, J. R. (1957), "Russell's Objections to Frege's Theory of Sense and Reference," *Analysis* **17**, 137–43. Reprinted in Klemke (1968), pp. 337–45.
- Searle, J. R. (1968), *Speech Acts*, Cambridge University Press, Cambridge.
- Sluga, H. (1980), *Gottlob Frege*, Routledge & Kegan Paul, Boston.
- Smullyan, A. F. (1948), "Modality and Description," *Journal of Symbolic Logic* **13**, 31–7.
- Stalnaker, R. (1984), *Inquiry*, MIT Press, Cambridge.
- Strawson, P. F. (1950), "On Referring," *Mind* **59**, 320–44.
- Tarski, A. (1944), "The Semantic Conception of Truth," *Philosophy and Phenomenological Research* **4**, 341–75. Reprinted in L. Linsky ed., *Semantics and the Philosophy of Language*, University of Illinois Press, Urbana, 1952, pp. 11–47.
- Urquhart, A., ed. (1994), *Collected Papers of Bertrand Russell*, Vol. 4, Routledge, London.
- Washington, C. (1992), "The Identity Theory of Quotation," *Journal of Philosophy* **89**, 582–605.
- Whitehead, A., & Russell, B. (1910), *Principia Mathematica*, Vol. 1, Cambridge University Press, Cambridge.
- Wiggins, D. (1965), "Identity Statements," in R. J. Butler, ed., *Analytic Philosophy*, second series, Blackwell, Oxford, pp. 40–71.
- Wittgenstein, L. (1922), *Tractatus Logico-Philosophicus*, *Annalen der Naturphilosophie*. Translated by D. F. Pears and B. F. McGuinness, Routledge & Kegan Paul, London, 1961.
- Wittgenstein, L. (1953), *Philosophical Investigations*, Blackwell, Oxford.
- Wright, C. (1983), *Frege's Conception of Numbers as Objects*, Scots Philosophical Monographs Number Two, Aberdeen University Press, Aberdeen.

# Index

- Abbé, E., 1, 203  
about, 21, 54, 129, 131  
active/passive, 46  
actual, 35  
Alston, W., 112, 209  
Anderson, C. A., 143  
*Anschauung*, 34, *see also* experience;  
intuition  
Anscombe, G. E. M.,  
argument-expression, 13  
Arnauld, A., xv  
assertion, 134  
assertion sign, 134
- Baker, G., 203  
Baumann, J., 3  
Beaney, M., xviii, 203  
*Bedeutung*, xv, xvii–xviii, 67, 93, *see also*  
reference  
*Begriff*, 16, *see also* concept  
*Begriffsschrift* semantic theory, 43, 114–18  
*Begriffswort*, 77, 120, *see also* concept  
word
- Bennett, J., 112, 123, 124, 208, 209, 215  
*Bestimmungsweise*, 47–9  
Black, M., xviii, 55, 71, 133, 214  
Boole, G., 2  
Boolos, G., 187  
Brentano, F., 111  
Burge, T., 213, 214  
Bynum, T., 203, 215
- Cantor, G., 3  
cardinal number, 64, 111  
Carnap, R., 6, 52, 141, 152, 154  
categorical grammar, 68  
Chomsky, N., 26, 69
- Church, A., 40, 41, 125–6, 143, 167, 211,  
212  
Church-Langford Translation Test, 52, 168  
cognitive value, 28–9, 32, 44, 92  
compositionality, 11, 37, 98, 125, 128, 145,  
146, 148  
concept, 16, 63, 64, 73, 81, 82, 97, 111, 120,  
144  
extension of, 79, 206  
first-level, 80  
second-level, 18  
concept *horse*, 73  
concept word, 63, 117  
concept/object, 64, 73, 83, 123  
content, 37, 43, 114, 115–18, 125  
cognitive, 28, 30, 33  
conceptual, 42, 43, 44, 45, 46, 47, 59  
Context Principle, 3, 100, 139  
Correspondence Theory of Truth, 138–9  
course-of-values, 206
- Davidson, D., 126, 137, 141, 170, 171, 214  
*de dicto*, 107, 129  
*de re*, 107, 129  
attribution of a property, 98  
Dedekind, R., 3  
definite descriptions, 85, 96, 133  
disguised, 85, 93, 103  
Frege's Theory, 96–7  
Russell's Theory, 85–7  
Descartes, R., 104  
direct acquaintance, 108  
direct reference, 92, 115, 117, 153  
directly referential, 94, 156  
Donaho, S., 210  
Dummett, M., xix, 7, 65, 81, 113, 142, 155,  
156, 158, 165, 203, 204, 205, 208, 213

- Eigennamen*, 65, 120, 132, *see also* proper name
- Engel, S. M., 122
- Erkenntniswerth*, 28, *see also* cognitive value
- Evans, G., 121
- existence, xv, 18, 79, 94, 100, 107, 108, 111, 112, 116, 118, 119, 121
- being, 94
- claim, 106
- experience, 34
- extensionality, 12, 17, 37, 131, 145, 148
- fact, 210
- Fischer, K., 1
- Fitting, M., 89, 207, 208, 209, 211, 213
- Forbes, G., 154, 208, 212, 213
- formal concept, 82
- formal mode/material mode, 23–6, 52, 67
- Free Logic, 98
- Frege, Alfred, 1, 203
- Frege, Auguste, 1
- Frege, Karl Alexander, 1
- Frege/Russell Description Theory of Names, 85
- Fritsche, F., xix
- function, 66, 150
- first-level, 18, 70, 204
- second-level, 18, 204
- function-expression, 11, 63, 66
- function-name, 67
- first-level, 69, 77
- second-level, 69, 80
- function/argument, 9–13, 14, 18, 26, 44, 64, 136, 149, 165
- function/object, 8–9, 63, 72
- Funktionsname*, 67, *see also* function-expression; function-name
- Furth, M., 41, 60, 203, 207
- Gabriel, G., 203
- Geach, P., xviii, 55, 112, 204, 207, 214
- Gedanke*, 43, *see also* thought
- Gegenstand*, 8, 65, *see also* object
- Gödel, K., 130, 137, 210
- Hacker, P., 203
- Heidelberger, H., 143
- Hilbert, D., 5
- Hochberg, H., 211
- Husserl, E., 3
- Hylton, P., 212
- idea, 35, 36
- identity of content, 42, 45, 47, 61–2
- Indiscernibility of Identicals, 19, *see also* Leibniz's Law
- infinite hierarchy, 141, 154, 170
- nonrigid, 142
- rigid, 142
- Inhalt*, xvi, 7, *see also* content
- begrifflichen*, 43, *see also* content: conceptual
- beurteilbar*, 44, *see also* judgment: possible content of
- Inhaltsgleichheit*, 42, *see also* identity of content
- intuition, 47
- judgment, 134
- possible content of, 44
- Kant, I., 100, 113, 121
- Kaplan, D., 57, 150, 156, 171, 174, 183, 207, 212
- Kerry, B., 73
- Kienzler, W., 203
- Kneale, M., 41, 50
- Kneale, W., 41, 50
- Kreiser, L., 203
- Kremer, M., 212
- Kripke, S., 207, 213, 214
- Law of Excluded Middle, 90, 98
- Law of Noncontradiction, 20
- Leibniz's Law, 19–23, 26
- Levine, J., 206
- Lieseberg, Margarete, 1, 203
- Linsky, L., 41, 50, 53–8, 154, 213
- Lockwood, M., 209
- Lotze, H., 1
- Marshall, W., 71
- Marty, A., 111
- Mates, B., 40, 203
- Meinong, A., 94, 111, 207
- Mendelsohn, R. L., 89, 207, 208, 209, 211, 213
- Mendelson, E., 214
- Mill, J. S., xv, 93
- Moore, G. E., 105, 183
- Neale, S., 207, 210
- number-name, 11
- O-quotation, 164
- object, 8–9, 65, 66, 73, 82, 108
- nonexistent, 90
- objective, 35
- Ockham, W. of, xv
- Ontological Argument, 110
- oratio obliqua*, 128, 142, 165
- oratio recta*, 128, 165

- P-quotation, 164  
 Paradox of Identity, 28, 42, 84, 91, 92, 101, 208  
 Paradox of Nonbeing, 94, 96, 101, 106  
 Parsons, T., 142, 151, 173, 174, 175, 177, 178, 212, 213, 214, 215  
 part, 38  
 part/whole, 44, 64, 149  
 Peano, G., 5  
 Perry, J., 209  
 Pohle, H., 4  
 predicate  
   first-level, 102  
   second-level, 102  
 predicate abstract, 89, 211  
 presuppose, 96  
 primary/secondary, 87, 207  
 Principle of Acquaintance, 93  
 proper name, 66, 96, 144  
   genuine or logically, 33, 85, 103  
 property  
   first-level, 110  
   second-level, 110  
 proposition, 115  
   object-dependent, 89  
   singular, 33, 44  
   structured, 149  
 propositional function, 97  
 Propositionalist, 23  
 Pünjer, B., 101, 115–18  
 Putnam, H., 206
- quantified modal logic, 128, 207  
 quantifier, 18, 103, 110, 111  
 Quine, W. V. O., xix, 13, 16, 20, 22, 52, 104, 105, 128, 129, 166, 167, 168, 169, 179
- quotation marks  
   Demonstrative Theory,  
   Identity Theory, 174  
   iteration of, 177  
   Picture Theory, 171  
   Proper Name Theory, 214  
   Spelling Theory, 174  
   Structureless Name Theory, 169  
 quotation name, 180
- Ramsey, F., 83, 84  
 Realism, 139  
 Redundancy Theory of Existence, 105, 122  
 Redundancy Theory of Identity, 209  
 Redundancy Theory of Truth, 105, 212
- reference, 32, 33, 93, 115  
   of a *Begriffswort*, 120  
   context-free, 162  
   context-sensitive, 162  
   customary, 40, 140, 147  
   of a declarative sentence, 38, 125  
   indirect, 40, 140, 147, 153  
   of a name, xv, 144  
   of a singular term, 33  
   referentially opaque, 129  
   referentially transparent, 130  
 Reichenbach, H., 173  
 relation, 8, 16, 18  
 relative identity, 112  
 Richard, M., 213, 215  
 Ricketts, T., 212  
 rigid description, 156  
 rigid designator, 213  
 Russell, B., xvi, xvii, 5, 31, 32–3, 41, 58, 84, 85, 86, 87, 88, 89, 90, 91, 93, 94, 97, 98, 102, 103, 106, 108, 109, 128, 132, 133, 141, 142, 153, 154, 158, 180, 207, 209, 212, 213  
 Russell's contradiction, 5  
 Ryle, G., 108
- Salmon, N., 209  
 Schering, E., 1  
 Schröder, E., 2  
 scope distinction, 87, 208  
 Searle, J., 174, 213  
 self-reference, 153, 167  
 semantic ascent, 47  
 sense, 32, 33, 36, 115  
   customary, 40, 140, 153, 156  
   of a declarative sentence, 38, 146  
   of a function-expression, 64  
   grasp of, 36  
   indirect, 40, 140, 153, 156  
   of a name, xv, 144  
   of a singular term, 33  
 Sense-datum theorist, 23  
 sense/reference, 33, 36, 37, 47–8, 63–4, 96, 112, 114, 119, 127, 161, 165  
 simple/complex, 10, 12–13  
 single-quote convention, 163  
 single-quote translation, 53  
 singular term, 17, 117  
   nonreferring, 118  
*Sinn*, xv, 93, 145, *see also* sense  
 Sluga, H., 203  
 Smullyan, A., 89  
 Stalnaker, R., 213  
 Strawson, P. F., 96  
 substitution principle, 91, 146  
   *Begriffsschrift*, 29, 44  
   for reference, 19, 38  
   for sense, 39–40

- Tarski, A., 137, 167  
 that, 147  
 'that' clause, 39–40, 127, 146, 148,  
 149, 155, 161  
 'that' operator, 141, 165  
 thought, 37, 38, 39, 45, 70, 120, 140,  
 144, 146, 149  
   object-dependent, 121  
   objects part of, 206  
 truth-functional, 16, 88, 128, 131  
 truth value, 15, 19, 32, 38, 125  
   the True and the False, 125,  
 139
- Umfang*, 206, *see also* concept: extension of  
 use/mention, 60, 167, 168
- Vorstellung*, 35, *see also* idea
- Washington, C., 215  
*Werthverlauf*, 8, 64, 206  
 Whitehead, A. N., 87, 88, 89, 97  
 Wiggins, D., 41, 58–9  
 Wittgenstein, L., xvii, 5, 82  
 Wright, C., 208
- Zermelo, E., 5, 204