

Meaning and Truth

Topics in Contemporary Philosophy

Editors

Joseph Keim Campbell
Michael O'Rourke
David Shier

Editorial Board Members

Kent Bach
Michael Bratman
Nancy Cartwright
Richard Feldman
John Martin Fischer
Nick Gier
P.J. Ivanhoe
Michael McKinsey
John Perry
James Rachels
Stephen Schiffer
Harry Silverstein
Brian Skyrms
Peter van Inwagen

Meaning and Truth

Investigations in Philosophical Semantics

Edited by

Joseph Keim Campbell
Washington State University

Michael O'Rourke
University of Idaho

David Shier
Washington State University



SEVEN BRIDGES PRESS, LLC

CHATHAM HOUSE PUBLISHERS

NEW YORK • LONDON

Seven Bridges Press, LLC
135 Fifth Avenue, New York, NY 10010-7101

Copyright © 2002 by Seven Bridges Press, LLC

All rights reserved. No part of this book may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without prior permission of the publisher.

Executive Editor: Jehanne Anabtawi
Managing Editor: Katharine Miller
Composition: Bytheway Publishing Services
Cover Design: Darby Downey, DD Creative Services
Printing and Binding: CSS Publishing Company, Inc.

Library of Congress Cataloging-in-Publication Data

Meaning and truth : investigations in philosophical semantics / edited
by Joseph Keim Campbell, Michael O'Rourke, David Shier.

p. cm. — (Topics in contemporary philosophy)

Includes bibliographical references and index.

ISBN 1-889119-55-5 (pbk.)

1. Semantics (Philosophy)—Congresses. I. Campbell, Joseph Keim,
1958— II. O'Rourke, Michael, 1963— III. Shier, David, 1958— IV. Series.

B840 .M456 2001

121'.68—dc21

2001002390

Manufactured in the United States of America

10 9 8 7 6 5 4 3 2 1

Contents

Foreword by Stephen Schiffer	vii
Editors' Acknowledgments	ix
Introduction	
<u>JOSEPH KEIM CAMPBELL, MICHAEL O'ROURKE, AND DAVID SHIER</u>	
1. Investigations in Philosophical Semantics: A Framework	1
Part I. Cognition and Linguistic Meaning	
<u>KENT BACH</u>	
2. Seemingly Semantic Intuitions	21
<u>MICHAEL MCKINSEY</u>	
3. The Semantic Basis of Externalism	34
<u>ROBIN JESHION</u>	
4. Acquaintanceless <i>De Re</i> Belief	53
<u>STEPHEN SCHIFFER</u>	
5. Meanings	79
Part II. Linguistic Meaning and the World	
<u>NATHAN SALMON</u>	
6. Mythical Objects	105

vi	MEANING AND TRUTH	
	<u>MARIAN DAVID</u>	
7.	Truth and Identity	124
	<u>KIRK LUDWIG</u>	
8.	What Is the Role of a Truth Theory in a Meaning Theory?	142
	<u>JONATHAN SUTTON</u>	
9.	A New Argument against Modesty	164
	<u>ROBERT CUMMINS</u>	
10.	Truth and Meaning	175
	Part III. Aspects of Linguistic Meaning	
	<u>HERMAN CAPPELEN AND ERNIE LEPORE</u>	
11.	Insensitive Quantifiers	197
	<u>EMMA BORG</u>	
12.	Deferred Demonstratives	214
	<u>LENNY CLAPP</u>	
13.	What Unarticulated Constituents Could Not Be	231
	<u>ANNE BEZUIDENHOUT</u>	
14.	Generalized Conversational Implicatures and Default Pragmatic Inferences	257
	<u>KENT BACH AND ANNE BEZUIDENHOUT</u>	
15.	Distinguishing Semantics and Pragmatics	284
	Index	311

Foreword

Stephen Schiffer

New York University

THE PAPERS IN this volume are article versions of selected talks given at the third annual Inland Northwest Philosophy Conference, on Truth and Meaning, held in Moscow, Idaho, and Pullman, Washington, March 24–26, 2000. This was the first year the conference was funded to bring in participants from all over the United States, and if, as I expect, future colloquia in the series meet the same high standards, the annual INPC will occupy an important place in American philosophical life. As a high-quality annual colloquium, it will quickly gain the prestige and attention now held by only two other such philosophy colloquia in the United States, the one at Chapel Hill and the one at Oberlin.

I was honored by the invitation to be the keynote speaker at the colloquium, but I had little idea of what to expect from a philosophy of language colloquium in Moscow, Idaho. Happily for me, it turned out to be one of the best-run and most stimulating philosophy conferences I have ever attended in any area of philosophy. The editors of this volume, Joseph Keim Campbell, Michael O'Rourke, and David Shier, who must be thanked for conceiving the series and actually getting it to happen, organized and ran the conference with near-awesome skill. The Universities of Idaho and Washington State are to be commended for their generous and wholehearted support, thereby making this new philosophical institution possible, one that will bring to those universities each year a level and excitement of philosophical activity enjoyed at very few other universities.

The collection of papers published in this volume, aptly subsumed under the wide-ranging rubric *Meaning and Truth*, covers most, if not all, of the topics in the philosophy of language that are currently of most concern. The papers by Lenny Clapp, Robert Cummins, Marian David, Kirk Ludwig, Michael McKinsey, Jonathan Sutton, and myself deal with foundational questions about the nature of meaning, of meaning theories for particular languages, and the

analytical relations between meaning and truth. The papers by Emma Borg, Herman Cappelen and Ernie Lepore, Robin Jeshion, and Nathan Salmon bring specific questions about reference and quantification to bear on more general questions about the nature of meaning. And the contributions by Kent Bach and Anne Bezuidenhout concern the semantics/pragmatics distinction. (More detailed introductory comments on these papers is provided in the editors' introduction.) I hope readers of these papers will be as stimulated and informed by them as we, the participants of the conference, were by the talks on which they were based.

Editors' Acknowledgments

THE ESSAYS IN this volume are descendants of some of the papers delivered at the third Inland Northwest Philosophy Conference (INPC), held March 24–26, 2000, in Moscow, Idaho, and Pullman, Washington. For their financial and administrative support of the conference, we are grateful to the Philosophy departments of Washington State University (Michael Neville, Chair) and the University of Idaho (Kathryn Paxton George, Chair), the College of Liberal Arts at Washington State University (Barbara Couture, Dean), the College of Letters and Science at the University of Idaho (Kurt Olsson, Dean), and the Research offices at the two universities. The INPC Steering Committee, consisting of Mary Bloodsworth, Ray Dacey, Robert Epperson, David Garrison, Harry Silverstein, and Kurt Torell, played a big role in making the conference happen, and we are grateful to them for their time and energy. Special thanks are also due to Jean McIntire, Dee Dee Torgeson, and Angie Whitney. In addition, we are grateful to the Franklin J. Matchette Foundation for a grant to sponsor the INPC Keynote Lecture delivered by Stephen Schiffer, and to the Idaho Humanities Council, a state-based affiliate of the National Endowment for the Humanities, for its consistent and unflagging support of the conference.

After the INPC was held, difficult decisions had to be made about which of the many excellent conference papers would be included in this volume. We were aided in this task by a number of quick and conscientious referees, to whom we owe a substantial debt of gratitude. Here is that list of creditors: Paddy Blanchette, Mary Bloodsworth, Eros Corazza, Andrew Cortens, James Edwards, Reinaldo Elugardo, Robert Epperson, Heimir Geirsson, Nick Gier, Michael Losonsky, Peter Ludlow, Marga Reimer, Samuel Rickless, Harry Silverstein, Rob Stainton, Russell Wahl, Leora Weitzman, and Jonathan Westphal. We thank Seven Bridges Press for their help on this book project in all of its stages. Also, the contributing authors have been wonderful to work with on this project, and we are grateful to them.

Finally, we thank our families for their support during conference preparations and the editing process. Extra special thanks go out to Delphine Keim Campbell, Rebecca O'Rourke, and Phyllis Shier, the three people to whom this book is dedicated.

CHAPTER 1

Investigations in Philosophical Semantics: A Framework

Joseph Keim Campbell, Michael O'Rourke, and David Shier

FRAMING THE THEORY OF MEANING

The Big Picture

MEANING IS EVERYWHERE—in our thoughts, in our words, in our actions, in the world. Wherever we turn, it is there. Each of us crafts a life around the meaning we find, setting goals, acting and reacting according to what we take this meaning to be. There is, of course, nothing new in this observation. It describes our experience in a way that collects together a varied set of features that do not form a natural category. As such, it may motivate a theoretical investigation into the nature of meaning, but it will not ground one. A ground for a theory of meaning can only be recovered if this observation is focused. There are a number of ways to do this. We might focus, for instance, on those aspects that are regarded as generally meaningful, as opposed to those that have meaning for specific individuals or select groups. Alternatively, we might focus on meaningful human actions, attending to other appearances of meaningfulness only when they are relevant to an understanding of action.

Still another way is to focus on those aspects of meaning that are grounded in representation, or roughly, things that have the function of standing for or signifying something beyond themselves. For instance, photographs, diagrams, realistic drawings, sentences, discourses, and so on. Attempts at isolating these for investigation reveal that other similar things do not count as representations—doodles, say, or random strings of letters. Thus, the first question that arises for one who chooses this approach is *What is the difference between representations and similar things that are not representations?* An answer to this question should make clear the properties that distinguish representations from non-representational objects.

The discipline of semantics, broadly construed, can be understood as devoted to the investigation of the representational character of things that have these properties, or what is the same, to their significance as representations. The focus in this volume is the *semantics of language*, or that branch of the broader discipline that concerns linguistic representations—i.e., words, phrases, and sentences that function as parts of language. These are recovered from auditory disturbances and ink stains, noises and marks. As is generally true of representations, these noises and marks resemble other noises and marks that are not linguistic representations. For instance, consider:

- (1) Scalia is a textualist.
- (2) cliasa si a xttlstiuae

Despite similar group size and similar letter distribution across groups, only (1) counts as a linguistic representation, while (2) is alphabet soup. The working hypothesis behind the semantics of language is that (1) has certain properties that underwrite its representational character, and (2) lacks them; further, these properties account for the place of (1) in a systematic language that people understand and employ. The semantics of language is in the business of accounting for these properties.

Dimensions of the Semantics of Language

The goal of the semantics of language is to construct a theory of the meanings of linguistic representations. Theory construction in this domain is no different than theory construction in other disciplines where the world contributes the phenomena. We start with observations drawn from our experience with the phenomena and then develop a conceptual model, under empirical pressure, that subsumes the phenomena under systematic generalizations. Empirical and conceptual elements interact synergistically—the empirical data initially inspire conceptual construction and then force conceptual revision along the way to a theory, while the conceptual structures that are built delimit what counts as empirical data. Thus, investigations in the semantics of language must be responsive to empirical evidence as well as conceptual considerations.

Relative to this characterization, we can identify two broad dimensions of the resulting theory, viz., the *conceptual dimension* and the *empirical dimension*. The former concerns the internal structure of the theoretical framework, comprising the concepts themselves and their interrelation, while the latter concerns the relation between this structure and the world of our experience.

Along the conceptual dimension, we find considerations bearing on both the nature of linguistic meaning and our knowledge of it. The first set of considerations are metaphysical and can be classified with the help of two fundamental questions:

What are linguistic representations?

What explains the fact that linguistic representations are meaningful?

The first question motivates us to reflect on the nature of language. Is it a set of abstract entities, such as symbol types or utterance types, or is it rather a set of cognitive representations, realized neurophysiologically? More narrowly, the first question raises the issue of the semantic hierarchy of linguistic representations. Should words or sentences be regarded as the primary semantic building blocks, or should we look to uses or utterances instead? A popular way to answer the second question is in terms of meanings that are associated with linguistic representations in some complex fashion. These could be things that are already lying around, such as objects in the world or ideas in the mind, although serious obstacles lie in the way of this purely referential approach. We might, instead, take these representations to be associated with abstract entities, such as functions or propositions, that explain their significance as representations. Or perhaps we combine these two proposals in the classic Fregean style. A wholly different approach takes meaningfulness to be explained by the manner in which linguistic representations are used. Here the leading idea is that linguistic representations are meaningful because of the specific roles they play in the economy of human interaction. They have meaning because of the roles they play, setting this approach apart from the first, which holds that they can fill these roles *because* of their meanings.

The importance and value of language are anchored in its meaningfulness, but only because we have knowing access to this meaningfulness. As with intellectual investigations generally, inquiry into the nature of language throws light on our knowledge of language while it illuminates its object. But in contrast with many other inquiries, such as, say, astronomy, linguistic inquiry does not obviously take as its object something that lies outside the mind of the knower. Indeed, there are those who argue that we must direct the light of our inquiry at this knowledge in order to illuminate language, since the language we speak and our knowledge of it are inextricably bound up with one another. Whether or not this is correct, it establishes epistemic considerations as central to the study of language generally, and so to semantic theory in particular. The epistemic aspect of semantic theory can be isolated with the help of three questions:

In what does our knowledge of linguistic meanings consist?

How do we come to know linguistic meanings?

How do we use our semantic knowledge?

In answering the first, we must attend to the cognitive realization of the language and the suite of abilities it underwrites, as it is through these that the relevant epistemic conditions are met. The second question directs us to consider the circumstances in which we enrich our semantic knowledge, a ground rich

with clues about the nature of linguistic meaning. The final question is related to perhaps the most fundamental aspect of our semantic knowledge—the interpretation and production of significant utterances. Our semantic knowledge accounts for our ability to participate in communicative exchanges as producers and interpreters, and this is arguably the most important employment of language.

Conceptual labors shape the theoretical framework, but the tenability of the theory is dependent on its relationship to empirical data extracted from observation of actual linguistic practice. The theory is about worldly phenomena, so we must look to those phenomena to inspire and guide us through the judgments they pass on our efforts. Hence the importance of the empirical dimension of semantic theory. We can look at two specific sources of data as especially relevant to semantic theorizing, both of which influence much of what is done in this volume. First, there are the facts about how natural languages are actually used by members of the relevant speech communities. It is in these observations that we find data, such as deferred demonstratives or nonsentential assertions, that require treatment by the theories we build. Thus, this source supplies the pool of data to be explained by those theories. Second, there are the semantic intuitions of those who speak the languages. Just as grammaticality intuitions are an important touchstone for syntacticians, so too are semantic intuitions a touchstone for semanticists. Although their importance is disputed, it is certainly clear that these intuitions guide the theorist in a provisional way and inspire theoretical judgments. These two sources supply facts about how we act with language and how we react to it, facts that must be explained by any adequate theory of linguistic meaning.

Modeling Linguistic Meaning

To carry out the task of explaining the epistemological and metaphysical facts associated with the significance of language, semantic theory might be asked to do either of two things. First, it might be asked to lay bare the central semantic concepts and the interrelations between them. Second, there is the business of yielding a meaning for each meaningful sentence of a language. Both are legitimate goals, but they correspond to semantic theory at different levels of abstraction. Pursuit of the former generates a conceptual framework that supports the assignment of meaning required by the latter, while pursuit of the latter ensures that this conceptual framework is connected to the data in a direct way. In aiming to construct this hierarchy of theories, semantic theorists will typically focus on some context or set of contexts in which meaningful language can be studied. One such context is interpersonal communication. Communicative discourse is sustained by the subtle interplay of meanings, and so investigation of it can shed light on both the crucial conceptual components of a theory of meaning and on the data this theory must explain. Thus, communicative dis-

course can serve as a source for both the conceptual and the empirical evidence needed to construct such a theory.

With this in mind, we offer an analysis of a fairly typical episode of interpersonal communication. By focusing on an exchange involving speaker, spoken sentence, and listener, we identify elements that can be used to construct a model of communicative discourse. This model is spare and provisional, requiring modification in light of additional episodes, puzzles, pathological cases, and so forth, but it serves us by highlighting relationships among concepts that influence the framing of semantic theory.

Consider, then, an episode in which two people are conversing, one in a chair (*S*) and one by an open window (*L*). At some time *t* during the conversation, *S* utters the sentence, “I’m cold,” intending to cause *L* to shut the window. Focus first on *S*, and in particular on her thoughts. At the time *S* utters the sentence, she is in a certain overall cognitive state, determined in part by her beliefs about herself, her listener, the conversation in which she has participated, her desires regarding herself and the future of the conversation, and whatever intentions she might have concerning her future contributions to the episode. These cognitive representations are related in complicated ways to the world, being effects of worldly causes, causal factors of effects in the world, and representations of the way the world is or might be. They *structure* the experience for *S*, framing her contributions and her interpretations. Within this overall state, there are specific representations that motivate *S* to utter the sentence, “I’m cold.” These representations include her beliefs about her current comfort level and her desire to warm up, and they causally influence in some complicated way the utterance of the sentence. These particular representations *trigger* the utterance. The speaker’s thoughts, then, causally influence her verbal performance in two different but related ways. In addition, the sentence uttered can be seen as *expressing* the content of certain thoughts that figure into the triggering cause of the utterance. In some cases, the literal meaning of the sentence may exhaustively express the content of these thoughts, but in many if not most cases, the thought expressed outstrips the literal content of the sentence.

Second, consider the sentence uttered. This sentence, like the thoughts that led to its production, stands in a complex relationship to the world. It borrows from the world, taking both a referent (the person picked out by ‘I’) and a time. It also purports to describe a state of affairs, viz., how things stand with the felt temperature of *S*. Finally, it is produced by a speaker for a reason, and so is a window of sorts on her thoughts. The overall meaning of the sentence as uttered on this occasion is influenced by each of these relationships. Attending to them aids us in making a few preliminary distinctions among aspects of linguistic meaning. First, there is what we can call *sentential meaning*, or the conventional linguistic meaning that is associated with the sentence type and so follows it from utterance to utterance. This includes conventional meanings of nonindex-

ical terms and phrases, referential features (or “characters”) for indexical elements of the sentence, both articulated and suppressed, and the syntactical relationships among these. We might think of it as something like a rule or a function that determines which proposition is expressed by any token of the sentence in a context. In this case, the sentential meaning could be given by something like, *<The speaker of this sentence> is cold at <the time of the utterance>*, where we use descriptive phrases to specify the referential features for ease of exposition. Second, there is what we can call *utterance meaning*, or the meaning typically associated with the utterance type, individuated by the sentence produced and the circumstances of its production. Utterance meaning will comprise at least the sentential meaning plus elements drawn from the context in accordance with the referential features of the indexical constituents, but it may also comprise meanings that are not anchored in any explicit part of the sentence. Here, the utterance meaning is given by something like *S is cold at t*. Finally, there is the *speaker meaning*, or the meaning the speaker intended the listener to take from the utterance in that context. This could simply be the utterance meaning, or it could be the utterance meaning plus additional implicated meanings, or perhaps it could be the implicated meanings alone without the utterance meaning, as in cases of irony. In this episode, the speaker meaning would be given by something like, *S is cold at t and S wants L to close the window*. Together, these three levels of meaning enable speakers to make claims about the world, and in so doing, further their practical goals.

Finally, we turn to the listener. *L* receives the utterance, processing its semantic characteristics below the level of consciousness. This is done against a background of cognitive states that structure the process for *L*. Beliefs, desires, and intentions shape the way in which *L* interprets this utterance. In doing so, *L* will likely aim to determine what *S* intended by the remark, assuming he wishes to remain a party to the conversation. Did *S* mean only to describe her comfort level, or was there something more to her utterance? In the context as described, *L* will likely take it to be an implicit request and do so without even thinking about it, leaving him with a decision to make about closing the window. Here we have a successful episode: *S* has intended *L* to understand her needs by way of her utterance, and *L* has understood them.

One way to develop semantic theory is to begin with this model as a foundation and then construct a theory that explains the nature of its various elements and the relationships among them. To be adequate, though, the theory must of course extend beyond this one episode. In the course of theory construction, it will be necessary to consider a wide variety of additional episodes, both successful and pathological, that involve a range of sentences, discourse contexts, and participants. Among the semantic complexities to be addressed are issues of compositionality, quantification, scope, modality, singular reference, and propositional attitude contexts. This all too brief and rather random list of

issues points to the complexity of the explanatory task that confronts the semanticist. The assumption behind investigations into the meaning of natural language is that nomological principles bring the richly varied data into systematic connection, supporting explanatory generalizations at the level of specific languages and at the level of language in general. The modest aim of semantic theory is the identification of these principles and the generalizations they support.

SEMANTIC INVESTIGATIONS

Semantic theory aims to account for the concepts and conceptual relationships that explain the meaningfulness of language, and thereby make possible the specification of meanings for the significant pieces of particular natural languages. The model outlined in the previous section relates many of the most crucial elements, thereby delineating one general way of developing the theoretical framework for semantics. We use this framework to structure the contents of this volume, relying specifically on three general relationships that figure into the model. First, there is the relation between the contents of thoughts, whether in the speaker or the listener, and the meanings of linguistic items. Second, we have the semantically complex relation between the sentence uttered and the world. Finally, there are the similarly complex relations among the aspects of meaning that can be associated with the utterance. Each of these serves as a principle for organizing the essays in this volume. We consider each principle and its associated essays in turn.

Cognition and Linguistic Meaning

'Content' is a term of art, introduced to capture what it is about representations that makes them capable of representing anything in the first place. By their very nature, representations *present* information about states of affairs, or perhaps present states of affairs themselves, for consideration, and so it is not unreasonable to think of them as metaphorically *containing* this information or these states of affairs. Cognitive and linguistic representations serve as exemplars here. These representation types are linked not only because they bear content, but also causally, through the production and interpretation of utterances, and logically, through similarity of structure. A full inquiry into the nature of linguistic meaning requires that we attend to these linkages, as such an account should explain the causal-functional role of linguistic meaning in communication. We begin our volume with a cluster of essays devoted to aspects of this cognitive/linguistic nexus.

In the first essay, "Seemingly Semantic Intuitions," Kent Bach addresses methodological concerns surrounding the use of so-called semantic intuitions in our theorizing. As we have noted, we gain clues about semantic facts through our intuitions about the meaningfulness of linguistic items, including uttered sentences, and these intuitions also serve to constrain theory construction in this

domain. But how prominent a role should they play? One influential recent answer is that we should *preserve* our intuitions in our theorizing (Recanati 1993). Bach, however, believes that while intuitions serve an important role as data, they do not in general reveal the semantic facts. Relying on the semantic-pragmatic distinction, the “Syntactic Correlation Constraint,” and the notion of “sentence nonliterality,” Bach advances a series of arguments and marshals a series of examples intended to show that our semantic intuitions can be sensitive to nonsemantic factors and insensitive to semantic ones, and so should not be trusted to lead us to semantic truths.

With the second essay, “The Semantic Basis of Externalism” by Michael McKinsey, we turn our attention to cognitive content. The appearance of Hilary Putnam’s “Twin Earth” thought experiments in the literature brought externalism about linguistic content into view as a viable theoretical alternative. Putnam was concerned with making a semantic point about certain terms, but externalism about content spread naturally to cognitive representations. In this domain externalism takes on various forms, but one prominent form is characterized by McKinsey as follows: “Many . . . predicates of the form ‘is thinking that p’ express properties that are wide, in the sense that possession of such a property by an agent logically implies the existence of contingent objects of a certain sort that are external to that agent” (this volume, p. 35). This view about the contents of cognitive representations has important implications for their role in the cognitive economy. For instance, McKinsey has argued that it is incompatible with privileged access to the contents of one’s own thoughts, on the grounds that privileged access and externalism jointly entail that one can have *a priori* knowledge of the existence of external contingent objects, an evident absurdity (McKinsey 1991). But there are those, such as Brian McLaughlin and Michael Tye, who hold out for the compatibility of externalism and privileged access. In his essay, McKinsey argues that such compatibilism is maintained only by adopting an interpretation of externalism that undercuts the original semantic motivations for the view.

The nature of cognitive content remains the focus of the third essay, “Acquaintanceless *De Re* Belief,” by Robin Jeshion. One can have a belief about an object by believing that whatever has some property *P* is thus-and-so; alternatively, one can have a belief about an object by believing of a *particular* object that one has in mind that it is thus-and-so. The former type of belief is known as *de dicto* belief, while the latter is *de re* belief. Paradigm cases of *de re* belief are underwritten by direct connections between the believer and the object, such as perceptual and, in particular, visual contact. Traditional accounts of *de re* belief require that the believer have such a connection with the object in order to have the belief, and the term ‘acquaintance’ is introduced to stand for this type of connection. Jeshion contends that while acquaintance is present in the paradigm cases, it is not a necessary condition on *de re* belief. In her view,

de re belief is belief that plays a certain role in cognition. She argues that upon introducing a certain type of name into the language—viz., a name whose referent has been fixed by a felicitous act of descriptive stipulation, such as ‘Neptune’—one can have beliefs about the referent of the name that play the *de re* cognitive role and so count as *de re* beliefs, even when the believer is unacquainted with the referent.

The preceding essay examines one way in which linguistic meaning can influence our views of cognitive content. In the last essay in this section, Stephen Schiffer examines the implications that cognitive content, and in particular, belief content, can have on linguistic meaning. In *Remnants of Meaning*, Schiffer argued that there is no positive theory of meaning for language and thought that is true. In “Meanings,” he rejects this conclusion and replaces it with a deflationary account of meanings. He holds that the things we mean, assert, believe, and so on are propositions of a particular kind—what he calls pleonastic propositions. This is an allusion to the “something-from-nothing transformations” that introduce propositions into our conceptual scheme, such as the one that takes us back and forth between ‘Fido is a dog’ and ‘That Fido is a dog is true’ (more colloquially, ‘It’s true that Fido is a dog’), which contains the new singular term ‘that Fido is a dog’ whose reference is the proposition that Fido is a dog. The characterization of pleonastic propositions in terms of something-from-nothing transformations then gives way to a more precise characterization in terms of its being a conceptual truth that if such-and-such is the case, then the proposition that such-and-such is true. But it’s the unique sort of relation which that-clauses in propositional-attitude reports (e.g., ‘that Fido is a dog’ in ‘Ralph believes that ‘Fido is a dog’) bear to the proposition to which they refer that fills out the account of pleonastic propositions. In the end, Schiffer argues (a) that if the pleonastic propositions to which that-clauses refer are structured, in a certain technical sense of that term, then the doctrine of pleonastic propositions is a version Fregeanism, wherein the propositions we mean and believe are structured entities composed of “concepts” (albeit pleonastic concepts) of the objects and properties our beliefs and statements purport to be about; (b) that it’s at best indeterminate whether pleonastic propositions are structured; and (c) that nothing of theoretical interest turns on this indeterminacy.

Linguistic Meaning and the World

In the second cluster of essays, our attention is directed toward the metaphysical side of semantic theory. The relevance of linguistic meaning to the general run of human beings is grounded in the connections that anchor language to the world. Characterizations of linguistic and cognitive content often require mention of the world, since aspects of content typically stand in certain relationships with elements of the world. The nature of these relationships is the subject of intense debate, as are the natures of those elements of the world that

are involved in such characterizations. Principal among these relationships is *truth*, a concept that is dismissed as superfluous by some and held up as the cornerstone of semantic theory by others. Much of what we can say about the nature of meaning in our semantic inquiry will hinge on how these metaphysical debates turn out. In this section, we have five contributions to these debates that examine various aspects of the metaphysics of meaning.

Among the metaphysical concerns is the worry about what sorts of things, if anything, are the referents of fictional and mythical terms. In “Mythical Objects,” Nathan Salmon takes up the topic of mythical terms, arguing that our discourse and beliefs about myths commit us to the existence of mythical objects—genuine, though abstract, entities that are created by belief in myths. On Salmon’s view, for instance, the Loch Ness monster is a real thing, though not a real creature. It is an abstract entity that came into being as a result of people’s mistaken beliefs. It’s what those people believe (mistakenly) to inhabit the Loch. Salmon mobilizes various linguistic phenomena in support of his surprising view, including the infamous puzzle raised by Peter Geach concerning sentences such as ‘Hob thinks a witch has blighted Bob’s mare, and Nob wonders whether she (the same witch) killed Cobb’s sow’. The trouble is that there seems to be no formal analysis of the sentence corresponding to the intuitively possible reading on which the speaker is committed to Hob’s and Nob’s thoughts having a common focus (a posited witch) but not to the existence of a real witch. Salmon rejects several proposed solutions, arguing that relational analyses overcommit the speaker to a witch, or at least a real person thought to be a witch, while the notional ones undercommit the speaker by failing to ensure the proper sort of common focus between Hob’s and Nob’s thoughts. Salmon argues that the puzzle can be solved neatly with a fully relational analysis on which the speaker is committed to the existence of a *mythical* witch—a real, abstract entity created by Hob’s and Nob’s mistaken beliefs.

In “Truth and Identity,” Marian David addresses the longstanding issue of the nature of truth. According to the traditional correspondence theory, often contrasted with coherence theories and pragmatic theories of truth, a true proposition is one that corresponds to a fact. David addresses another alternative, the identity theory, according to which true propositions don’t correspond to facts, but simply *are* facts. He argues that this theory has problematic consequences, and points out that since the correspondence theory is sometimes alleged to “collapse into” the identity theory, correspondence theorists will need to block this collapse charge. David reconstructs the chain of reasoning that appears to underlie the collapse charge, and suggests a response on behalf of the correspondence theorist. He then considers two versions of the correspondence theory and the different ways in which they can offer the suggested response to the collapse charge.

The concept of truth is obviously central to understanding linguistic mean-

ing, since the meaning of an indicative sentence together with the way the world is determines whether or not it is true. While traditionally the concept of meaning has been thought to be a relatively richer concept than that of truth, Donald Davidson famously argues in his essay, “Truth and Meaning,” that we should construct a theory of meaning for a language out of a Tarski-style truth theory, or “T-theory” (Davidson 1984). A T-theory has base clauses specifying the satisfaction conditions for predicates and referents for referring terms, and recursive clauses specifying the satisfaction conditions for recursive constructions in the language in terms of the satisfaction of their parts. Such a theory must satisfy Tarski’s Convention T, which requires that an adequate truth theory for a language be formally correct and entail every theorem of the form

(T) S is true-in-L iff P

for all sentences of the object language, where S is replaced by a description of a sentence in the object language in terms of its construction from its primitive significant elements and P by a metalanguage translation of S. This theorem is the *T-sentence* for the object language sentence S. One question that emerges about this proposal is whether T-theories are supposed to *replace* meaning theories, or whether they represent a novel way of pursuing a meaning theory traditionally conceived. While many, aided by Davidson’s own words, seem to favor the former answer, Kirk Ludwig argues for the latter in his essay, “What Is the Role of a Truth Theory in a Meaning Theory?”

Ludwig’s goal in this essay is to “say what knowledge we could have about a truth theory for a language that would enable us to use it to interpret speakers of *that language*” (this volume, p. 146). This knowledge would count, then, as a meaning theory. He proceeds in two steps. First, he argues that the axioms that supply the reference or satisfaction conditions in a formally correct T-theory must use metalanguage terms that mean the same thing as the object language terms whose reference or satisfaction conditions they give. Such a theory will generate all the T-sentences for the object language among its theorems and do so in a way that reveals, in their proofs, their compositional structure. However, because the theory may have theorems that have the right form but are not T-sentences, a second step is required. In this step, Ludwig spells out a “canonical proof procedure” that generates T-sentences with proofs drawing solely on the content of the axioms for the theory. If we have a T-theory that has a canonical proof procedure and axioms satisfying the first constraint and we also know what the axioms of that theory mean, then we can use the recursive machinery of the theory to generate theorems that support interpretation of speakers of the object language. A meaning theory, then, comprises this knowledge. With this in hand, Ludwig proceeds to argue that while the meaning theory *exploits* the T-theory to reveal the compositional structure of the language

and generate true meaning theorems, it does not *embed* the T-theory and so avoids the semantic defects, such as paradox and vagueness, that frustrate more traditional versions of truth-theoretic semantics.

As is true of influential theories, Davidson's is not without its critics. One prominent criticism leveled at the Davidsonian program is due to Michael Dummett (1975), who argues that T-theories fail to serve as adequate theories of meaning for natural languages. On Dummett's view, a theory of meaning must be such that knowledge of it would be knowledge of the language. Thus, theories of meaning must be "full-blooded," supplying explanations for the concepts necessary to understand the primitive terms of the language. Dummett argues that a T-theory fails to supply such explanations and so is "excessively modest" and therefore inadequate as a theory of meaning. Dummett's case for this view, however, is generally regarded as unpersuasive because it is not clear that semantic knowledge is required for possession of these concepts.

In "A New Argument Against Modesty," Jonathan Sutton seeks to shore up Dummett's conclusions that (a) a theory of meaning must explain the possession of certain concepts, and (b) the Davidsonian T-theory approach fails to do this. Sutton begins by making explicit a general methodological constraint on theory construction. He calls this the "Competition Constraint," and it specifies that if "a certain kind of theory is of the right *form* to be a theory of meaning for a natural language, then debates about the semantics of various constructions should be describable as debates over which particular theories of that form are correct" (this volume, pp. 166–67). Given this, if the Davidsonian program supplies the proper form of a theory of meaning, we should be able to characterize debates about the semantics of expressions as debates about various T-theories. Sutton begins by describing two different ways to conceive of proper names, a difference that could support a semantic debate. He then argues that this debate cannot be modeled with T-theories unless the metalanguages used in specifying the T-sentences are enhanced to include information about possession of the different concepts of proper names. However, this amounts to saying that a T-theory is adequate only if it is full-blooded, which is precisely Dummett's conclusion.

The received, truth-conditional approach to semantics inspired by Davidson is also challenged in the final essay of this section, "Truth and Meaning." In this essay, Robert Cummins argues that the core Davidsonian principle that to know the meaning of an expression is to know its satisfaction condition is far from obviously true. Indeed, according to Cummins, truth is hardly even relevant to what's involved with speaking and understanding a language. Cummins identifies the "communicative meaning" of a term as whatever it is that one has to have in one's head in order to understand the term. And what that is, he continues, is a concept—or, what is the same according to

Cummins, a body of knowledge about its instances. Therefore, a concept is not a mental representation, but is rather a kind of theory about things of that *sort*. However, theories cannot be combined in the sorts of ways necessary for Tarskian combinatorics; i.e., Tarskian processes don't take theories as their inputs. Therefore, on this view of concepts and hence of communicative meanings, Tarskian truth theory could not be an adequate theory of communicative meaning. Drawing on ideas from connectionist theory, Cummins then sketches a possible alternate approach to a theory of communicative meaning, one that could construe sentences as, in effect, recipes for assembling representations in the heads of others and on which, contrary to Davidsonian semantics, the communicative capacity of language could far outstrip its expressive power.

Aspects of Linguistic Meaning

Assessment of a typical episode of conversation reveals connections between linguistic content and cognitive content in all participants, as well as connections between these contents and the world. These are the extrinsic relations into which linguistic meaning enters, and the essays in the previous sections attend to these. No account of the meaning of language would be complete, however, without some discussion of its *intrinsic* relations, i.e., the relations among the various aspects of the total significance of the conversational contributions. Earlier, we distinguished between sentential meaning, utterance meaning, and speaker meaning. How the total significance of an utterance is modeled depends on decisions about the legitimacy of these categories and, if legitimate, the conditions that constrain and relate them to sentences, utterances, and speakers. The essays in Part III address this cluster of questions.

We begin this section with three essays devoted to the analysis of specific linguistic constructions involving quantifiers, deferred demonstratives, and putative unarticulated constituents. As we noted in Part I, theory construction in natural language semantics depends on a type of methodological synergy between empirical observation and rational reconstruction. Semantic theory must explain linguistic constructions as they appear in actual discourse, so attention to data about the use of specific constructions is critical to success. These essays demonstrate how this is done, and they do so with three semantically significant linguistic constructions.

Consider the sentence, 'No one is absent today', as uttered by a professor to her class. On the standard reading that one could find in any logic textbook, the sentence would be interpreted as expressing the proposition that there isn't a single absent individual anywhere. Surely this isn't what the professor intends to communicate. But how are we to construe this situation? On the one hand, one might try to retain the relatively simple, standard semantics for the sentence, and hold that what's actually said or asserted in this case departs from the proposition semantically expressed. One might maintain, for example, that although

the professor utters a sentence that expresses the false proposition that no one anywhere is absent, she is nonetheless intending to communicate the information that no one *in the class* is absent. She communicates this by means of uttering the false sentence and relying on her audience to work it out from features of the context. On the other hand, in light of such cases where the standard semantic account doesn't seem to match up with what people say, one might argue that the standard semantic account is inadequate and should be revised to be sensitive to contextual features and to thus reflect what people are saying.

In the first essay in this section, "Insensitive Quantifiers," Herman Cappelen and Ernie Lepore take up the issue of whether the semantics of quantifiers should be rendered as context-sensitive. A number of philosophers and linguists, appealing to intuitions about such sentences as 'No one is absent today', have argued for context-sensitive semantics for quantifiers. In particular, it has been claimed that the domains of quantifiers should be interpreted as semantically restricted to narrower subsets by contextually relevant features. Thus, the professor's 'no one' might, according to the semantics, apply to the domain of just the people in the class, and thus her sentence would actually express a true proposition. However, Cappelen and Lepore build a three-part case against context-sensitive semantics for quantifiers. First, they argue that such semantics would make incorrect predictions about some sorts of utterances—for example, that certain sentences could be used to express necessary truths when intuitively they could not. Second, they argue that the intuitions used to motivate the context-sensitive semantics cannot genuinely support that treatment, on the grounds that speakers would have these intuitions even if the semantics were explicitly context-insensitive. Finally, Cappelen and Lepore raise an objection to the context-sensitive treatment based on the observation that nothing in the context provides a unique determination of the restricted domain.

Similar issues arise in connection with demonstrative pronouns. An editor of this volume was once riding to school with his wife when she pointed at a newly vacant lot that formerly contained a large building familiar to them both. "Look," she said, "that's not there anymore!" The word 'that' would normally be treated as a demonstrative and so, on the widely held view, treated as a genuine referring term. But there is an obvious problem in this case—to wit, the absence of a building for her to demonstrate. What she was saying, or trying to say, seems true, but the very fact that would make it true also seems to make it impossible for her to refer to the intended object! Similarly, when a victim points to a mug shot and says, "That's the attacker," it doesn't seem appropriate at all to reply, "You were attacked by a photo?" Yet if we treat the word as an ordinary (referential) demonstrative, the sentence would indeed express a proposition about the object pointed to, i.e., the photo. In her chapter "Deferred Demonstratives," Emma Borg considers cases of demonstrative utterances in which the speaker uses a demonstrative to pick out an object that stands in some

relation to the thing actually pointed to in the demonstration. It is tempting to analyze such cases as, in effect, equivalent to utterances of certain descriptions, and a number of philosophers have done just that. For instance, the first example above might be treated as equivalent to an utterance of ‘The building that was on that lot is gone.’ Again, as with the earlier question about the context-sensitivity of quantifiers, one might hold either that such descriptive content is semantically associated with deferred demonstratives or that the standard referential semantics for demonstratives applies and that the descriptive content is somehow conveyed pragmatically, through the extrasemantic exploitation of contextual features. Borg considers both these descriptivist approaches to deferred demonstratives, and argues that each faces formidable difficulties. She uses a number of tests, however, which involve considerations of modality, reference failure, and scope, to show that deferred demonstratives behave not as disguised descriptive phrases but as genuine referring expressions. In the remainder of the essay, she develops her own positive account on which deferred demonstratives are treated as semantically referential and which, she argues, surmounts the difficulties standardly faced by referential accounts by separating the notions of ostensive gesture (e.g., pointing) and demonstration.

In his essay, “What Unarticulated Constituents Could Not Be,” Lenny Clapp considers a set of issues involving putative counterexamples to truth-conditional compositionality, i.e., the principle that the truth conditions of an utterance are a function of the logical form of the utterance together with the meanings of its words. Such compositionality has been the received view since Frege, and lies at the heart of traditional semantic programs such as those advanced by Davidson and by Montague, but it has come under pressure from several directions. Clapp discusses four distinct sorts of *prima facie* counterexamples to compositionality in the literature, i.e., examples involving truth conditions that vary across contexts, but which contain no relevant context-sensitive elements. The four sorts of cases involve quantifiers with restricted domains, comparative adjectives, propositional attitude reports, and nonsentential assertions. He then considers two strategies for defending truth functional compositionality from these apparent counterexamples, both of which involve the positing of unarticulated constituents, i.e., propositional constituents that are not the semantic values of any phonetically realized sentence element. The first is the “pragmatic ellipsis” strategy, and the second is the “hidden indexical” strategy. Clapp argues that both of these strategies face versions of the same problem, in that there is an overabundance of candidates for what the relevant unarticulated constituents could be, and no principled way for choosing among them. In such a situation, he argues, we have no grounds for accepting *any* of the unarticulated constituent analyses, and thus we should regard the counterexamples to compositionality as genuine. If this is correct, it would appear to undermine received semantic programs and, as Clapp notes, it would put tremendous strain on the

traditional Gricean distinction between semantics and pragmatics. If semantics is concerned with “what is said,” as Grice (1989) put it, and if what is said is in part a function of contextual features, as the examples discussed by Clapp are purported to show, then semantics must be concerned in part with such contextual features—the very sorts of features held to be paradigmatically pragmatic on the Gricean model.

This result points to a larger question about the meaning associated with an utterance, viz., whether the three-part distinction between sentential, utterance, and speaker meaning is even sustainable. Stephen Levinson (2000) has argued that it is, urging us to embrace a level of utterance-type meaning that fits between the conventional meaning of sentences and the “one-off” type of speaker meaning that is highly dependent on special features of a particular context of utterance. As such, this level comprises what Grice called “generalized conversational implicatures,” which are implicatures associated in typical circumstances with certain types of expressions (Grice 1989, 37). Levinson takes these to be “default pragmatic inferences” that yield preferred interpretations that can be cancelled but typically are not. These are generated by applications of default heuristics and inference rules in the production and interpretation of utterances, and they supplement sentential meaning while remaining distinct from speaker meaning. In “Generalized Conversational Implicatures and Default Pragmatic Inferences,” Anne Bezuidenhout disputes Levinson’s case for this view, arguing that it does not have the desirable features it purports to have. In particular, its default heuristics and inference rules do not always yield the correct results, and when they do, the results do not appear different from particularized conversational implicatures of the sort that constitute speaker meaning. Further, the framework of heuristics and inference rules often produces multiple interpretations, and it isn’t obvious that any one of these can claim to be the *default* interpretation. For these reasons, Bezuidenhout advocates rejecting Levinson’s three-part distinction in favor of a two-part “Relevance Theoretic” account that is grounded in a single principle, the “Communicative Principle of Relevance.” She agrees with him that generalized conversational implicatures are not speaker meaning, but contends that they are not different in kind from sentential meaning; instead, they combine with sentential meaning to form the meaning of the uttered sentence, a unified level of meaning that is the proper object of semantic inquiry.

The final chapter, “Distinguishing Semantics and Pragmatics,” contains further work on the levels of meaning associated with an utterance episode, and in particular on the distinction between the content due to the uttered sentence and that due to the speaker. In this chapter, Kent Bach and Anne Bezuidenhout offer contrasting views on this distinction, debating whether semantics should be restricted to content encoded in sentences (Bach) or whether it should accommodate additional content due to highly contextualized pragmatic infer-

ences (Bezuidenhout). In “Semantic, Pragmatic,” Bach argues that the semantic-pragmatic distinction concerns “two types of information associated with an utterance.” Semantic information is encoded in the sentence uttered, relative to what he calls “narrow context,” and so is independent of a speaker’s communicative intentions. By contrast, pragmatic information is made relevant by the act of uttering the sentence and pertains to the audience’s identification of the content of the speaker’s communicative intentions. Thus, anything not governed by sentence grammar is not a part of semantics. This contrasts significantly with the view defended in “Radical Pragmatics,” where Bezuidenhout argues for *pragmatic intrusion*, or the intrusion of pragmatic elements into the semantic content of sentences, at least if that is identified with “what is said” by utterances. (Note that she speaks of the contents of *utterances*, whereas Bach speaks of the contents of *sentences*.) Pragmatic intrusion expands what counts as semantic by making the content of an uttered sentence much more dependent on utterance context and communicative intentions than Bach will allow. As a result, much of what counts as strictly pragmatic on Bach’s view qualifies as semantic according to Bezuidenhout. Theory construction in this domain requires that we determine the reach of the relevant concepts and the relationship among them. Like the others in this volume, this chapter seeks to establish some of the distinctions necessary to achieve this goal.

REFERENCES

- Davidson, D. 1984. *Inquiries Into Truth and Interpretation*. Oxford: Oxford University Press.
- Dummett, M. 1975. “What is a Theory of Meaning? (I).” In S. Guttenplan (ed.), *Mind and Language*. Oxford: Oxford University Press.
- Geach, P. 1967. “Intentional Identity.” *Journal of Philosophy* 64: 627–32.
- Grice, H. 1989. *Studies in the Ways of Words*. Cambridge, Mass.: Harvard University Press.
- Levinson, S. 2000. *Presumptive Meanings: The Theory of Generalized Conversational Implicatures*. Cambridge, Mass.: MIT Press.
- McKinsey, M. 1991. “Anti-Individualism and Privileged Access.” *Analysis* 51: 9–16.
- McLaughlin, B. and M. Tye. 1998. “Externalism, Twin Earth, and Self-Knowledge.” In C. MacDonald, B. Smith, and C. Wright (eds.), *Knowing Our Own Minds*. Oxford: Oxford University Press.
- Putnam, H. 1975. “The Meaning of ‘Meaning.’” In K. Gunderson (ed.), *Language, Mind, and Knowledge*. Minneapolis: University of Minnesota Press.
- Recanati, F. 1993. *Direct Reference: From Language to Thought*. Oxford: Blackwell.
- Schiffer, S. 1987. *Remnants of Meaning*. Cambridge, Mass.: The MIT Press.

PART I
Cognition and Linguistic Meaning

CHAPTER 2

Seemingly Semantic Intuitions

Kent Bach

San Francisco State University

FROM ETHICS TO epistemology to metaphysics, it is common for philosophers to appeal to “intuitions” about cases to identify counterexamples to one view and to find support for another. It would be interesting to examine the evidential status of such intuitions, snap judgments, gut reactions, or whatever you want to call them, but in this chapter I will *not* talk about moral, epistemological, or metaphysical intuitions. I will focus on semantic ones. In fact, I will focus on semantic intuitions about sentences, not individual words (although the contributions of individual words may ultimately be at issue in some of these cases), and on closely related intuitions about what is said in utterances of those sentences. Such intuitions play an important role in the philosophy of language. For example, intuitions about the informativeness of identity statements give rise to Frege’s problem; intuitions about the failure of substitution in attitude contexts are used to impose a constraint on an adequate theory of attitude ascriptions; and intuitions about sentences containing definite descriptions used referentially or incomplete definite descriptions have been relied upon to cast doubt on Russell’s theory of descriptions.

Although I have my doubts about such appeals to intuition with regard to these longstanding problems in the philosophy of language, I will not state them here. The cases I will consider are philosophically less interesting, but they are the sorts of examples that provide data for answers to the controversial question of how to draw the line between semantics and pragmatics (for my answer see Bach 1999). Such examples are often cited, e.g., by François Recanati in his paper, “The Pragmatics of What Is Said” (1989), and by the psycholinguists Ray Gibbs and Jessica Moise in their paper, “Pragmatics in Understanding What Is Said” (1997), to undermine the orthodox view that, to put it crudely, semantics provides input to pragmatics without any feedback from pragmatics. Following are a few examples thought to provide evidence for so-called pragmatic intrusion, in which pragmatic factors allegedly contribute to semantic interpretation:

(1) Billy will get promoted if he works hard.

which is intuitively understood to mean that Billy will get promoted if *and only if* he works hard;

(2) Mary has three cars.

which is intuitively understood to mean that Mary has *exactly* three cars;

(3) Bobby hasn't taken a bath.

which is intuitively understood to mean that Bobby hasn't taken a bath *lately*; and

(4) Molly got infected and went to the hospital.

which is intuitively understood to mean that Molly got infected *and then, because of her infection*, went to the hospital.

As I see it, these are all examples of what I call (Bach 1994a) conversational "implicature" (as opposed to Grice's implicature). In each case, I claim, what the speaker means is distinct from what he is saying, because what he means includes an implicit qualification on what he is saying, something that is not really part of what is said. However, people's intuitions tend not to be sensitive to the difference, at least not until they're sensitized. That's because they tend to ignore what I'll call the "Syntactic Correlation Constraint," as expressed by Paul Grice's stipulation that what is said must correspond to "the elements of [the sentence], their order, and their syntactic character" (1989, 87).

Since this is pretty much all that Grice says about what is said, I should add a few things about how I understand it. What is said is determined compositionally by the semantic contents of the constituents ("elements") of the sentence as a function of their syntactic relationship. To allow for the presence of tense and sometimes indexicals, we should add that what is said in a context is the semantic content of the sentence relative to that context. Notice that I do not speak of the semantic contents of *utterances*. If "utterance" means what is uttered, then an utterance is just a sentence. And if "utterance" means an act of uttering, then the content of an utterance is really the content of the speaker's communicative intention, which can depart in various ways from semantic content. Also, although what is said is sometimes described as the "proposition expressed" by the sentence (relative to the context), this mistakenly assumes that every sentence expresses a complete proposition. In fact, the syntactic requirements on well-formed sentences do not exclude the case of sentences whose semantic interpretation is not a complete proposition. There are many sorts of

sentence that do not express a complete proposition, not even relative to a context (see (15) and (16) on page 29 and Bach 1994a, sec. 2).

At any rate, the Syntactic Correlation Constraint entails that if any element of the content of an utterance, i.e., of what the speaker intends to convey, does not correspond to any element of the sentence being uttered, it is not part of what is *said*. So when people intuitively think that what is said includes such elements, their intuitions are illicitly including something that just isn't there. Now Grice did not think that people's intuitions deserve cavalier dismissal. He could have just explained them away by appealing to the distinction between what is said and what is implicated in uttering a sentence or to the distinction between what a sentence means and what a speaker means in uttering it, but he was worried by what struck him as a kind of paradox:

We must of course give due . . . weight to intuitions. . . . For in order that a nonconventional implicature should be present in a given case, my account requires that a speaker shall be able to utilize the conventional meaning of a sentence. . . . This . . . seems to lead to a sort of paradox: if we, as speakers, have the requisite knowledge of the conventional meaning of sentences we employ to implicate, when uttering them, something the implication of which depends on the conventional meaning in question, how can we, as theorists, have difficulty with respect to just those cases in deciding where conventional meaning ends and implicature begins? If it is true, for example, that one who says that A or B implicates the existence of non-truth-functional grounds for A or B, how can there be any doubt whether the word 'or' has a strong or weak sense? I hope that I can provide the answer to this question, but I am not certain that I can. (1989, 49)

In my opinion, Grice accorded intuitions too much respect. In fact, there is nothing at all paradoxical about how theorists can disagree about matters on which people's intuitions tend to agree.

However, Grice's paradox will seem troubling to anyone who supposes that the central aim of semantics is to account for such intuitions, especially ones about the truth-conditions of sentences. But this anxiety, like many others, is unfounded. It is the central aim of semantics to account for semantic facts, not intuitions. People's spontaneous judgments or "intuitions" provide data for semantics, but it is an open question to what extent they reveal semantic facts and should therefore be explained rather than explained away. Since, as I am suggesting, they are often responsive to nonsemantic information, to what is implicit in what is said but not part of it, they should be treated cautiously. They should certainly not be given the respect accorded to them by Recanati's so-called Availability Principle, which prescribes that intuitions about what is said

be “preserved” in our theorizing. Nor should they be taken as seriously as they are by Gibbs and Moise, who “examined people’s intuitions” and claimed to find that data about what people say about what is said “lend support to theories of utterance interpretation [according to which] pragmatics strongly influences people’s understanding of what speakers say [as well as what they] communicate” (1997, 51).

Aside from the question of the reliability of such intuitions and their relevance to semantics and its relation to pragmatics, there is the question of what role, if any, they play in the process of communication. It seems their role is marginal at best. In the course of speaking and listening to one another, we generally don’t consciously reflect on the semantic contents of the sentences we hear or on what is said in their utterance. We are focused on what we are communicating or on what is being communicated to us, not on what is said. Moreover, we don’t have to be able to make accurate judgments about what information is semantic and what is not in order to be sensitive to semantic information. To “preserve intuitions” in our theorizing about what is said would be like relying on the intuitions of unsophisticated moviegoers about the effects of editing on a film. Although people’s cinematic experience is dramatically affected by such factors as cuts and camera angles, there is no reason to suppose that their intuitions are reliable about what produces what effects. Intuitions about what is said may be similarly insensitive to the difference between the contribution that is made by the semantic content of a sentence and that made by extralinguistic factors to what an utterance communicates. So, I say, what worried Grice was not a real paradox but just an ordinary philosophical problem.

Let’s get down to cases. In discussing them, I should stress that I am not claiming that semantic intuitions are totally unreliable and shouldn’t be trusted at all. Rather, they should be relied upon judiciously, and only after being fed an ample diet of examples, including contrasting ones. For instance, although it might intuitively seem that (1),

(1) John will get promoted if he works hard.

says or at least entails that John won’t get promoted if he doesn’t work hard, this apparent entailment can be explicitly canceled without contradiction:

(1x) John will get promoted if he works hard, though he might get promoted even if he doesn’t work hard.

And intuitively it does not seem redundant to utter the strengthened version of (1),

(1+) John will get promoted if *and only if* he works hard.

Here are a few more examples of faulty intuitions. In each case, rather than take the intuition at face value, we can describe what is going on in a way that explains both the occurrence of the intuition and its falsity.

- (5) Jack and Jill went up the hill.
- (6) Jack and Jill are engaged.
- (7) Jill got married and became pregnant.

Although these sentences express complete propositions, in uttering them a speaker is likely to have meant something more specific, a qualified version of what he said:

- (5+) Jack and Jill went up the hill *together*.
- (6+) Jack and Jill are engaged *to each other*.
- (7+) Jill got married and *then* became pregnant.

where the italicized words, which are *not* part of the original sentence, indicate part of what the speaker meant in uttering (5), (6), or (7). He would have to utter those words (or roughly equivalent words—the exact words don't matter) to make what he meant more fully explicit (let's call the fuller version an "expansion" of the original and what is left out its "implicit qualification"). Utterances of sentences like (5) – (7) illustrate what I call "sentence nonliterality" (Bach, 1994b, 69–72), as opposed to constituent nonliterality, since no expression in the sentence is being used nonliterally. What the speaker means is not the exact proposition, as compositionally determined, that is expressed by the sentence, and the difference between the two propositions is not attributable to any particular constituent of the sentence. A speaker who utters (6), for example, is not *saying* that Jack and Jill are engaged to each other, any more than he would be saying this if he uttered "Jack and his sister Jill are engaged." That he means they're engaged to each other is implicit in what he is saying or, more precisely, in his saying of it. It is not part of what is said, since it passes Grice's test of cancellability; that is, it may be taken back without contradiction. There is no contradiction in uttering (6x),

- (6x) Jack and Jill are engaged but not to each other.

or, for that matter, (5x) or (7x),

- (5x) Jack and Jill went up the hill but not together.
- (7x) Jill got married and became pregnant but not in that order.

The implications of utterances of (5), (6), and (7) are all cancellable.

Even so, for many people the apparent content of each of sentences (5)–(7) includes something that is not predictable from the compositional semantics of the uttered sentence. Taking semantic intuitions seriously would make life miserable for semanticists, even more miserable than it already is. It would require doing semantics from the top down: starting with the supposed meaning of a sentence and working down to the meanings of its constituents. Covert constituents would have to be posited to provide the “residue” of meaning not accounted for by the overt ones. Or, alternatively, special meanings, hence ambiguity, would have to be attributed to certain of the overt constituents, insofar as their ordinary meanings seem not to make the right contribution to what is said. All this can be avoided if we don’t take people’s seemingly semantic intuitions too seriously.

Why are such intuitions unreliable about the semantic contents of sentences like those we’ve considered? Part of the reason is that *typical* utterances of them involve sentence nonliterality. Unlike cases of metaphor or metonymy, there are no constituents that intuitively are being used nonliterally. Moreover, there is a recurrent pattern of nonliterality associated with such sentences. Phenomenologically, their nonliteral use seems literal, at least insofar as our intuitions are insensitive to the difference between conventionalization and mere standardization (Bach 1998). As with what Grice called “generalized” conversational implicature, where there must be specific contextual reasons for supposing that an implicature is *not* present (for a monumental study of generalized conversational implicature, see Levinson 2000), in the above cases the sentence is typically used to communicate something that is not predictable from its meaning alone. So it’s no wonder that when people are asked for their intuitions about such a sentence, they will tend to imagine it uttered in a normal context and count its typical implicit qualification as part of its content. They tend to attribute something to the conventional meaning of the sentence that in fact is attributable only to typical utterances of it.

Recent experiments by Gibbs and Moise (1997) have sought to establish the reliability of semantic intuitions, as with examples like (8):

- (8) Martha gave John her key and he opened the door.

People judge that part of what is said is that John opened the door with the key Martha gave him. However, Gibbs and Moise’s experimental design was clearly flawed. For one thing, it imposed a false dichotomy on their subjects by forcing them to choose between what is said and what is implicated. Subjects weren’t offered the in-between category of implicit qualification (that which is implicit in the saying of what is said). Also, they were not given the opportunity to make cancellability judgments or comparative judgments about what is said by ex-

plicitly qualified utterances as opposed to unqualified ones. Gibbs and Moise didn't ask subjects if there is a contradiction in a sentence like (8x),

(8x) Martha gave John her key and he opened the door, but not with the key she gave him.

Gibbs and Moise predict that subjects would find a contradiction here; I predict that they wouldn't. Similarly, they didn't ask subjects to compare (8) with (8+),

(8+) Martha gave John her key and he opened the door with the key she gave him.

and to judge whether they say the same thing. Gibbs and Moise predict that subjects would judge that (8) and (8+) do say the same thing; I predict that they wouldn't.

Even if Gibbs and Moise are right about people's untutored intuitions about the original examples, it would be easy to sensitize their intuitions about what is said to Grice's cancellability test for what is not said. Just present them with sentences like (5)–(7) followed by cancellations of what is not explicit in the utterance, as in (5x)–(7x) above. Ask them if they sense a contradiction or just a clarification. Or ask them, with a stress on “say,” whether what a speaker *says* in uttering explicitly qualified versions of (5), (6), or (7), i.e., (5+)–(7+) above, is the same as what a speaker *says* with (5), (6), and (7) themselves, and they are likely to discern the difference. If so, this contradicts the intuition that the implicit qualifications are part of what is said in the original utterances. So the verdict of intuition is reversed when we appeal to people's cancellability judgments and their comparative judgments about what is said by explicitly qualified versus unqualified utterances.

I could discuss Gibbs and Moise's experiments in detail (in fact, Nicolle and Clark [1999] have done so, and report that their own experiments often delivered different results, sometimes with people deeming clear cases of implicature to count as what is said), but the main difficulty with their research, which shows how misguided it was, is that it tested for the wrong thing. Nicolle and Clark thought they could get data about what is said, and thereby test the empirical validity of Recanati's Availability Principle, by asking people what is said by a given utterance, or by asking them whether something that is conveyed by a given utterance is implicated or merely said. Evidently they assumed that what people *say* about what is said is strongly indicative of what *is* said. In fact, what it is indicative of is how people apply the phrase “what is said” and perhaps of what they mean by the word “say.” It tells us little about what is said, much less about the cognitive processes whereby people understand utterances.

To appreciate how small a role semantic intuitions play in utterance comprehension, consider the case of ambiguity. There are many ambiguous sentences one of whose meanings is far more likely to be operative than the other. The following headlines illustrate what I mean:

SURVIVOR OF SIAMESE TWINS JOINS PARENTS
PROSTITUTES APPEAL TO POPE
PANDA MATING FAILS; VETERINARIAN TAKES OVER
STUD TIRES OUT
BRITISH LEFT WAFFLES ON FALKLAND ISLANDS
TEACHER STRIKES IDLE KIDS
SQUAD HELPS DOG BITE VICTIM
ENRAGED COW INJURES FARMER WITH AXE
STOLEN PAINTING FOUND BY TREE

We find these headlines funny because we notice their unintended meanings, but evidently the meanings weren't noticed by the editors who allowed the headlines to run. Similarly, we are not likely to notice the ambiguity of (9) considered by itself,

(9) Bill scratched the car with an umbrella.

or the ambiguity of (10), considered by itself,

(10) Bill scratched the car with a broken tail light.

but their ambiguity is obvious as soon as we compare them with each other or with the obviously ambiguous sentence,

(11) Bill scratched the car with a broken antenna.

Similarly, by itself neither of the following sentences is likely to seem ambiguous,

(12) The soldiers exchanged their arms for food.

(13) The soldiers used their arms to protect their faces.

but their ambiguity is evident once we compare them with each other or with the obviously ambiguous (14),

(14) The soldiers celebrated by waving their arms in the air.

Interestingly, it has long been known (Lackner and Garrett 1973) that people unconsciously access irrelevant meanings of ambiguous words, but only very briefly, and only immediately (about 150 msec.) after hearing them. People's failure consciously to "intuit" the irrelevant meaning has no bearing on the process of utterance comprehension, and no one would seriously claim that their failure to do so is strong evidence against the existence of the ambiguity.

I have stressed that people's semantic intuitions tend not to respect the Syntactic Correlation Constraint. Now I'd like to consider some objections to that constraint and to the so-called "minimalist" conception of semantic content that goes with it.

There are two objections that I'll mention just briefly. The first notes the fact mentioned earlier, that some sentences, like (15) and (16), do not express complete propositions, not even relative to a context.

(15) Bonnie is ready. (for what?)

(16) Clyde is finished. (doing what?)

In such cases, there is something not semantically specified that is needed to yield a complete proposition. However, so the objection goes, what is said must be a complete proposition. Therefore, what is said in such cases is not a projection of the syntax of the sentence; it includes some element that does not correspond to any constituent, or feature of a constituent, of the sentence. However, why must what is said be a complete proposition? What's wrong with using (15_{IQ}) and (16_{IQ}) to report, and report fully, what a speaker *says* in uttering (15) or (16)?

(15_{IQ}) S said that Bonnie is ready.

(16_{IQ}) S said that Clyde is finished.

It may be true that a speaker, in using a sentence to *communicate* something, must communicate a complete proposition, but it hardly follows that any sentence used to communicate a complete proposition must itself express one. Sentences that are syntactically complete but semantically incomplete do not. To understand such utterances the hearer must figure out how the speaker intends what is said to be turned into a complete proposition. I call this process "completion." Utterances requiring completion, like those requiring expansion (utterances with an implicit qualification on what is said), carry implicatures along with what is said.

A second objection is based on resisting what Robyn Carston has called "the compulsion to treat all pragmatically derived meaning as implicature" (1988, 176). Once it is recognized that the contribution of pragmatic processes is not

limited to the determination of implicatures, “there is no reason,” according to Carston, “why pragmatics cannot contribute to the explicature, the truth-conditional content of the utterance,” which she equates with what is said. However, Carston is implicitly assuming that if something is not implicated, it is part of what is said. Since what she calls the explicature, which needn’t be fully explicit, can be the truth-conditional content of the result of an expansion or completion of the utterance, it cannot be identified with what is said.

A third objection is based on the fact that on minimalism what is said is often false even in cases when the utterance of the sentence in question is true. Sentence (3), for example,

- (3) Bobby hasn’t taken a bath.

though likely to be used to convey that Bobby hasn’t taken a bath lately, can itself be true only in the unlikely event that Bobby has never taken a bath. Similarly, (17) and (18), though used to convey truths, are likely to be literally false by minimalist standards:

- (17) That car doesn’t look expensive—it is expensive.
 (18) Nobody goes there any more—it’s too crowded. (once uttered by Yogi Berra)

Why should the prediction that these sentences are literally false lead to an objection to minimalism? Because intuitively they are true. Well, that’s one consideration, but it has little weight once we invoke the distinction between what is said and what is meant, and remember that intuitions tend to be insensitive to that distinction and to be responsive to implicit qualifications, as explicitly included in (17+) and (18+),

- (17+) That car doesn’t *merely* look expensive—it is expensive.
 (18+) Nobody *important* goes there any more—it’s too crowded.

Another basis for the objection is that, given this distinction, it can only be the obvious falsity of what is said that explains the hearer’s inference to what the speaker means. Evidently, the objection assumes that minimalism must treat these cases as Gricean quality implicatures. In those cases, the hearer’s inference is triggered by his recognition that the utterance, if taken at face value, violates Grice’s first maxim of quality, “Do not say what you believe to be false” (1989, 27). However, the obvious falsity of these sentences has nothing to do with how the hearer figures out what the speaker is conveying (or with how the speaker intends him to do so). The cases we’re concerned with are quite unlike an utterance of, say, “I could eat a million of those potato chips,” which conveys how

irresistible they are. In that case it is the obvious falsity of what is said that triggers the hearer's search for something other than what is said. But in the examples in question it is not obvious falsity that does that. Consider one more example. Suppose a child is crying because of a tiny cut and his mother tries to calm him by uttering (19),

(19) You're not going to die.

Obviously the mother is not assuring the kid of his ultimate immortality, but the operative pragmatic anomaly here is not obvious falsity but lack of relevant specificity. This is clear if we consider a positive version of the same utterance. An oncologist could say to a cancer patient who demands a frank prognosis, "I'm sorry to tell you, but there is nothing I can do. You're going to die." Presumably the patient won't take the doctor to mean that he, like anyone else, is mortal. But it's not the obvious truth of what is said that enables him to understand the doctor, it's the presumption that the doctor is telling him something relevant to his medical condition. Similarly, utterances of the negations of the previous examples would typically have the same implicit qualifications as utterances of those sentences themselves. Obvious truth, like obvious falsity, has no bearing on the hearer's inference in these cases.

Another objection to minimalism claims that even if we accept the strict, minimalist conception of what is said, what is said in that sense can have no psychological reality unless it is something that a hearer must identify before inferring the speaker's communicative intention. In other words, for what is said to matter psychologically, the hearer must identify what is said before identifying what is meant. But, so the objection goes, introspectively at least it seems that in many cases the first proposition one arrives at is not the "minimal" proposition (as Recanati 1989 calls it), the proposition which, according to the Syntactic Correlation Constraint, comprises what is said. Even if this is so, that is no objection to semantic minimalism. The process of utterance comprehension is obviously a very interesting topic for psychology, but it's hard to see why facts about *hearers'* cognitive processes should be relevant to what a *speaker* says. How could the fact (if it is a fact) that what is said sometimes has no psychological reality for the hearer show that it is a mere abstraction? All this shows is that hearers can infer what a speaker is conveying without first identifying what the speaker is saying. The semantic notion of what is said pertains to the character of the information available to the hearer in the process of identifying what the speaker is conveying, not to what goes on in this process (Bach and Harnish, 1979, 91–93).

Moreover, suppose that it is true that what is said, in the minimalist sense, is sometimes not consciously accessed. It is still consciously accessible. This is evident from the fact that people recognize, as we saw with examples (1)–(4),

that implicit qualifications on these utterances are cancelable. Furthermore, even if in some cases the minimal proposition is not actually computed and plays no role in the interpretation process as it actually occurs, because of “local processing” on constituents of the sentence, it can still play some role. Even if a hearer doesn’t explicitly represent what is said by the utterance of a sentence, hence does not explicitly reject it, still he makes the *implicit* assumption that it is not what is meant. Implicit assumptions are an essential ingredient in default reasoning in general (Bach 1984) and in the process of understanding utterances in particular. Communicative reasoning, like default reasoning in general, is a case of jumping to conclusions without explicitly taking into account all alternatives or all relevant considerations. Even so, to be warranted such reasoning must be sensitive to such considerations. This means that such considerations can play a dispositional role even when they do not play an explicit role. They lurk in the background, so to speak, waiting to be taken into account when there is special reason to do so.

I conclude that intuitive and related cognitive considerations do not undermine a minimalist conception of what is said. As Jerry Fodor says, “No doubt, intuitions deserve respect, . . . [but] informants, oneself included, can be quite awful at saying what it is that drives their intuitions. . . . It is *always* up for grabs what an intuition is an intuition of” (1998, 86). In the case of seemingly semantic intuitions, they are largely irrelevant to determining what is said. They are influenced by semantically irrelevant information, they tend to be insensitive to relevant distinctions, and they are likely to be biased in favor of understandings corresponding to things that people are relatively likely to communicate. Or so it seems to me, at least intuitively.

REFERENCES

- Bach, K. 1984. “Default Reasoning: Jumping to Conclusions and Knowing When to Think Twice.” *Pacific Philosophical Quarterly* 65: 37–58.
- Bach, K. 1994a. “Conversational Implicature.” *Mind & Language* 9: 124–62.
- Bach, K. 1994b. *Thought and Reference*. Paperback edition, revised with postscript. Oxford: Oxford University Press.
- Bach, K. 1998. “Standardization Revisited.” In A. Kasher (ed.), *Pragmatics: Critical Concepts, Volume IV*. London: Routledge. Online at <http://userwww.sfsu.edu/~kbach/standard.html>.
- Bach, K. 1999. “The Semantics-Pragmatics Distinction: What It Is and Why It Matters.” In K. Turner (ed.), *The Semantics-Pragmatics Interface from Different Points of View*. Oxford: Elsevier. Online at <http://userwww.sfsu.edu/~kbach/spd.htm>.
- Bach, K., and R. Harnish. 1979. *Linguistic Communication and Speech Acts*. Cambridge, Mass.: MIT Press.
- Carston, R. 1988. “Implicature, Explicature, and Truth-theoretic Semantics.” In R. M. Kempson (ed.), *Mental Representations: The Interface Between Language and Reality*. Cambridge: Cambridge University Press.
- Fodor, J. 1998. *Concepts: Where Cognitive Science Went Wrong*. Oxford: Oxford University Press.
- Gibbs, R., and J. Moise. 1997. “Pragmatics in Understanding What Is Said.” *Cognition* 62: 51–74.
- Grice, P. 1989. *Studies in the Way of Words*. Cambridge, Mass.: Harvard University Press.

- Lackner, J. R., and M. Garrett. 1973. "Resolving Ambiguity: Effects of Biasing Context in the Unattending Ear." *Cognition* 1: 359–72.
- Levinson, S. 2000. *Presumptive Meanings: The Theory of Generalized Conversational Implicatures*. Cambridge, Mass.: MIT Press.
- Nicolle, S., and B. Clark. 1999. "Experimental Pragmatics and What Is Said: A Response to Gibbs and Moise." *Cognition* 69: 337–54.
- Recanati, F. 1989. "The Pragmatics of What Is Said." *Mind & Language* 4: 294–328.

The Semantic Basis of Externalism

Michael McKinsey
Wayne State University

THE PRIMARY EVIDENCE and motivation for externalism in the philosophy of mind is provided by the semantic facts that support direct reference theories of names, indexical pronouns, and natural kind terms. But many externalists have forgotten their semantic roots, or so I shall contend here. I have become convinced of this by a common reaction among externalists to the main argument of my 1991 paper, "Anti-Individualism and Privileged Access." In that argument, I concluded that externalism is incompatible with the principle that we can have privileged, nonempirical knowledge of the contents of our own thoughts. The reaction in question amounts to a dismissive denial of one of my argument's main premises. This premise, which I defended at length in the paper, is that an externalist thesis regarding a cognitive property should hold that possession of the property by a person *logically*, or *conceptually*, implies the existence of objects external to that person.

Externalists who defend the compatibility of their view with privileged access by denying this premise of mine usually insist that externalist theses are about *a posteriori metaphysical*, as opposed to logical or conceptual, dependency relations between cognitive properties and external objects.¹ This insistence, I maintain, is a sign of deep confusion. The dependence on external facts, or wideness, of the property expressed by a given cognitive predicate is a function, not of some mysterious *a posteriori* metaphysical relation that the property (due to its "nature" perhaps) happens to bear to external objects; rather, it is a function of the wide *meaning* of, or the wide semantic contribution made by, a crucial component of the predicate, such as a name, an indexical, or a natural kind term. The fact that a given cognitive predicate expresses a wide property, therefore, is a fact about that predicate's meaning and logical form. When such a predicate is used, we do not just ascribe a property that happens, on independent *a posteriori* grounds, to bear some dependency

relation to external objects. Rather, the meaning and logical form of the predicate are such that, when we use it, we actually *say, assert*, that the person bears a given mental relation to external things. Any such assertion, of course, will logically, not just metaphysically, imply the bearing of relations to external things.

I shall call the compatibilist's strategy of claiming that externalist theses are about *a posteriori* metaphysical, as opposed to logical or conceptual, dependency relations, "the metaphysical strategy."² As we shall see, use of this strategy requires one to drive a wedge between what is actually said by use of a wide cognitive predicate, on the one hand, and the kinds of external facts, relation to which makes the property expressed by the predicate a wide property, on the other. But if there is such a wedge, then the wide meanings of the words contained in cognitive predicates cease to have any semantic bearing on the wideness of the properties expressed by the predicates. Thus, the compatibilist's metaphysical strategy effectively removes the semantic basis of externalism, and so takes away whatever reason there was to believe externalism in the first place.³

As a means of exposing the bankruptcy of the metaphysical strategy, I will concentrate on the recent attempt by Brian McLaughlin and Michael Tye (1998) to use this strategy to defend the compatibility of externalism and privileged access against my (1991a) argument. Their discussion is unusually forthright and detailed, and so, I will argue, their use of the metaphysical strategy clearly reveals its inconsistency with externalism's semantic motivation.

I WILL BEGIN with a brief restatement of my original *reductio* argument against compatibilism. I took the principle of privileged access to say that we each necessarily have the capacity to acquire *a priori* knowledge of the *contents* of our own thoughts:

Privileged Access to Content (PAC)

It is necessarily true that if a person *x* is thinking that *p*, then *x* can in principle know *a priori* that he himself, or she herself, is thinking that *p*.

Here, by '*a priori* knowledge' I mean knowledge that can be obtained, as we say, "from the armchair" or "just by thinking", as opposed to knowledge that is obtained by perceptual observation or empirical investigation. I argued that PAC is inconsistent with the following externalist thesis:

Semantic Externalism (SE)

Many *de dicto*-structured predicates of the form 'is thinking that *p*' express properties that are wide, in the sense that possession of such a

property by an agent logically implies the existence of contingent objects of a certain sort that are external to that agent.⁴

My argument considered an instance of ‘is thinking that p’ that contains the natural kind term ‘water’. Suppose it is true that Oscar is thinking that water is wet. Then by PAC it follows that

- (1) Oscar can know *a priori* that he is thinking that water is wet.

But given that ‘is thinking that water is wet’ expresses a logically wide property, it also follows that

- (2) The proposition that Oscar is thinking that water is wet logically implies the proposition E,

where E is some “external proposition” that asserts or implies the existence of objects external to Oscar. E might, for instance, be the proposition that water exists. Now the conjunction of (1) and (2) is clearly absurd. If Oscar can know *a priori* the proposition that he is thinking that water is wet, and this proposition that he knows *a priori* logically implies E, then Oscar could just *deduce* E from something he knows *a priori*, and so he could know E itself *a priori*. But this consequence is just absurd. For by assumption, E is a proposition such as the proposition that water exists, a proposition that asserts or implies the existence of contingent objects external to Oscar, and Oscar obviously cannot know such propositions from the armchair. So if ‘is thinking that water is wet’ expresses a logically wide property, then contrary to PAC, no one can know *a priori* that he or she is thinking that water is wet.⁵ And in general, PAC and SE are inconsistent.

MCLAUGHLIN AND TYE agree that the general *form* of my argument (or “McKinsey’s recipe,” as they call it) is correct, but they claim that the form has no correct instances. That is, they claim that there is *no* externalist thesis, and *no* case of a thought that p, such that *both* (i) having a thought that p, on that externalist thesis, logically implies some external proposition E, *and* (ii) a relevant principle of privileged access implies that one can know *a priori* that one has a thought that p (1998, 291). Thus, in the particular instance of the argument just stated, McLaughlin and Tye would just reject premise (2).

Naturally, I think that McLaughlin and Tye are wrong to claim that my argument has no correct application. For there are in fact some clear examples of externalist theses that obviously refute this claim. One such thesis is discussed at length by McLaughlin and Tye. This is the thesis that some thought contents are singular propositions, propositions that contain concrete individuals as constituents. McLaughlin and Tye agree with this thesis. Moreover, they correctly

point out that my argument against compatibilism can be deployed to show that no one can ever have privileged access to the fact that one's thought has a given singular proposition as its content (1998, 291–292). Certainly, I would agree that my argument has this consequence. However, this consequence surely *seems* inconsistent with the sort of principle of privileged access that McLaughlin and Tye themselves endorse, namely, PAC. So here is a case where they themselves seem committed to my argument's applicability.

Yet McLaughlin and Tye go on to insist that no one has ever held that we can have privileged access to the contents of our own thoughts, when these contents are singular propositions (p. 292). The question is, how can they consistently say this, when they themselves endorse PAC, and PAC certainly seems to imply this very consequence which, they claim, no one has ever held?

Suppose that

- (3) Jones is thinking that Cicero is an orator,

where, we assume, the sentence 'Cicero is an orator' expresses a singular proposition with Cicero as a constituent, so that (3) ascribes to Jones a thought that has this proposition as its content. It is an immediate consequence of the conjunction of PAC and (3) that

- (4) Jones can know *a priori* that he is thinking that Cicero is an orator.

But since (3) ascribes to Jones a thought whose content is the singular proposition that Cicero is an orator, and since—as McLaughlin and Tye assert—having such a thought logically requires the existence of Cicero, it surely seems to follow that

- (5) The proposition that Jones is thinking that Cicero is an orator logically implies that Cicero exists.

Yet as we've seen, and as McLaughlin and Tye agree, (4) and (5) are straightforwardly inconsistent. So the question is, How can McLaughlin and Tye themselves avoid this inconsistency? After all, they are explicitly committed to PAC, and they certainly also seem committed to the externalist thesis that entails (5).

It is quite clear that McLaughlin and Tye endorse PAC, and also that, as a consequence, they would endorse (4) (see p. 299, for instance). So they must deny (5). Now it is true that they seem to come very close to asserting (5), but actually, what they assert is not (5) but rather

- (6) "It is a conceptual truth that having a thought with the singular proposition that Cicero is an orator as its content requires the existence of Cicero" (1998, 292).

Apparently, then, McLaughlin and Tye think that they can consistently endorse (4) while denying (5) because they think that they can distinguish the proposition expressed by (3) (i.e., the proposition that Jones is thinking that Cicero is an orator) from the proposition expressed by

- (7) Jones has a thought whose content is the proposition that Cicero is an orator.

Thus, McLaughlin and Tye seem to be saying, while (7) ascribes a property that logically implies that Cicero exists, so that *this* property is one to which no one has privileged access, (3) by contrast ascribes a *different* property of the form 'is thinking that p', and this property, according to PAC, is one to which we *can* have privileged access.

Now this is a peculiar position for an externalist to take. Externalists generally agree that sentences like (3) containing small-scope proper names ascribe wide cognitive properties. But their *reason* for saying this, of course, is that the names contained in such sentences are directly referential, and thus contribute their referents to the proposition expressed by the whole cognitive ascription. The standard view is that an ascription like (3) says that the agent bears the relation of thinking to the (singular) proposition that Cicero is an orator.⁶ The ascribed cognitive property is wide on this account simply because it is relational with respect to the man Cicero. Thus, on this view, (3) ascribes to Jones and Cicero the relation that any objects x and y bear to each other just in case x bears the relation of thinking to the proposition that y is an orator.

In other words, the standard externalist explanation of *why* a sentence like (3) ascribes a wide property is just that (3) says exactly what (7) does, and thus logically implies that Jones has a thought whose content is a proposition involving the man Cicero. But we have just seen that McLaughlin and Tye must *deny* that (3) and (7) say the same thing. Again, they must hold that (3) does *not* say that Jones has a thought whose content is the singular proposition that Cicero is an orator. They never do make clear what, on their view, a sentence like (3) *does* say; but at least it seems clear that on their view, whatever (3) says, it does *not* say that Jones has a thought whose content is the proposition that Cicero is an orator.

But having denied the basic reason why most externalists believe that sentences like (3) ascribe wide cognitive properties, the question arises, why do McLaughlin and Tye themselves believe that such sentences ascribe such properties? The answer seems oddly inconsistent: they believe that (3) ascribes a wide property because it ascribes a thought whose content is the singular proposition that Cicero is an orator! (see p. 299). How can we make sense of this? Perhaps as follows: McLaughlin and Tye hold that (3) does not *say* or logically *imply* that Jones has a thought whose content is the proposition that Cicero is an orator.

Nevertheless, they seem to think that if Cicero does in fact exist, then (3) ascribes to Jones the having of a thought of a certain type T such that, to have a thought of type T one *must*, as a matter of fact, have a thought whose content is the proposition that Cicero is an orator (p. 299). In short, McLaughlin and Tye first separate the relevant wide content logically from what (3) says or implies, only to later reconnect it in some other way, to preserve the appearance of externalism.

This suggestion raises many difficult questions. What exactly is the type T of thought ascribed by (3)? Why *must* thoughts of type T have the wide content in question? And what is the force of the ‘must’ here? Since it cannot be the logical ‘must’, the modality is apparently that of metaphysical necessity. But how is having a thought of type T supposed to be “metaphysically,” but not logically, related to the proposition that Cicero is an orator, and what is the nature of this mysterious metaphysical connection anyway? I think it is very unlikely that these questions have any cogent answers.

AT ONE POINT in their paper, McLaughlin and Tye suggest that they would endorse a two-factor theory of the thought types ascribed by a cognitive predicate like ‘is thinking that Cicero is an orator’. One factor is the truth-condition, or proposition, expressed by the imbedded sentence; the second factor is a more fine-grained semantic feature, which they call a “mode of presentation” of that proposition (p. 294). But this two-factor idea doesn’t really help answer any of the questions I have raised about their view.

On the two-factor view, a sentence like (3) (‘Jones is thinking that Cicero is an orator’) will ascribe a type T of thought that either involves both factors, or involves just one of the two. If T involves both the singular proposition that Cicero is an orator and a mode of presentation of that proposition, or T involves just the singular proposition alone, then it follows that (3) says or logically implies that Jones has a thought whose content is the singular proposition that Cicero is an orator. But of course, McLaughlin and Tye explicitly deny this consequence.

On the other hand, the type T of thought ascribed by (3) might involve just a mode of presentation of the proposition in question. This alternative avoids logical implication of the wide content all right, but now it certainly seems to follow that the cognitive property ascribed by (3) would be a purely *narrow* property, a consequence that of course is also contrary to McLaughlin and Tye’s view. After all, one would assume, a mode of presentation of the proposition that Cicero is an orator would be only *contingently* related to that proposition. For instance, the mode of presentation M in question might involve a mode of presentation of Cicero as, say, “the greatest Roman senator during the Republic.” In another possible world in which Brutus rather than Cicero satisfies this description, M would pick out the different proposition that Brutus is an orator. So if the type T involves only a mode of presentation like this, then surely, the property of having a thought of type T is narrow.

To avoid this consequence, McLaughlin and Tye would have to insist that the relevant mode of presentation is somehow “metaphysically” but not logically related to the proposition that Cicero is an orator. But again, what is the nature of this mysterious metaphysical connection supposed to be? And how exactly could a mode of presentation of a proposition be metaphysically but not logically connected to that proposition? Again, these questions seem unanswerable. If McLaughlin and Tye were to take this line, they would just be trading inconsistency for irremediable obscurity.

Look at what has happened. In order to avoid my *reductio* argument, McLaughlin and Tye have had to suggest a view on which the wide meanings of the words contained in the that-clause of a cognitive ascription are strictly irrelevant to what ends up being *said* by that ascription, or to what cognitive property is ascribed by that ascription. Thus, on their view, neither the fact that the name ‘Cicero’ refers to a given man, nor the fact that the imbedded sentence ‘Cicero is an orator’ expresses a given singular proposition, at all affects what is said or logically implied by the thought ascription (3). So the wide proposition expressed by the imbedded sentence in (3) plays no semantic role in determining what this thought ascription says; the ascription does not so much as logically imply that the wide proposition in question even exists. But given these claims, to go on to suggest that, nevertheless, this very proposition can be somehow “metaphysically” reconnected to whatever the ascription really *does* say (thus saving the day for externalism), is merely to indulge in nebulous handwaving.

PERHAPS MCLAUGHLIN AND TYE would respond by saying that I have given an unfair characterization of their view. After all, on their view there is a semantic contribution made by a proper name like ‘Cicero’ in a thought ascription like (3). They call this contribution ‘the concept of Cicero’, and they seem to think that an ascription like (3) characterizes Jones’s thought as involving this concept (p. 299). Similarly, they seem to think that in a thought ascription containing a small-scope natural kind term like

(8) Oscar is thinking that water is wet,

the word ‘water’ expresses a certain concept, so that as a result (8) characterizes Oscar’s thought in terms of this concept (p. 298).

But in fact, introduction of the “concepts” expressed by the relevant words does nothing to alleviate the obscurity of McLaughlin and Tye’s view. For how is the fact that the words ‘Cicero’ and ‘water’ express certain concepts supposed to have the consequence that the cognitive predicates containing these words express wide properties? Apparently, the concepts in question must be quite special: they must themselves be “wide” in some sense, and must correspond to the wide meanings of the words that express them. But what makes a concept wide?

For McLaughlin and Tye, the answer can't be that the existence of the concept logically or conceptually implies the existence of external objects of some sort. For they insist at several points in their paper that if one's thought involves the concept of Cicero, or the concept of water, then one can know *a priori* that one's thought involves that concept.

So they have to say that the concepts of Cicero and water are somehow "metaphysically" but not logically related to external objects. But now we have the same old obscurity back again. For what could it possibly mean to say that the concept of Cicero, for instance, metaphysically but not logically depends upon the existence of the man Cicero? This requires explanation, first, because in normal cases, the relation between a concept and what it is a concept of, is just a *semantic* relation, and is not any sort of necessary connection, either metaphysical or logical. Take the concept of God, for example. We might reasonably suppose that this is the concept, say, of an all-powerful, all-benevolent being who created the heavens and the earth by intelligent design. Perhaps in some possible worlds, there is such a being, so that in such worlds, the concept of God applies to that being. In other worlds, perhaps, there is no such being, and the concept has no application. So it seems just to be a matter of contingent fact whether or not the concept of God is a concept of any existing thing.

But the concept of Cicero is supposed to be different. It is supposed to be a concept that is necessarily connected to the actual man Cicero. How could this be? Here is one explanation: perhaps the concept of Cicero is a concept that can only be identified, individuated, or defined by reference to the man Cicero himself. Maybe we could say that the actual meaning of the name 'Cicero' is completely exhausted by the fact that the name refers to Cicero. Similarly, it might be said that the concept expressed by this word can be specified only by reference to that very man. Then we could say that to be the concept of Cicero just is (by definition) to be the concept of *that man*.

In the case of a natural kind term like 'water', there are various ways in which the meaning of the term, and the concept it expresses, might be wide. Perhaps, as in the above hypothesis about the name 'Cicero', the meaning of the word 'water' is exhausted by the fact that it refers to a certain natural kind W. Then necessarily, for a concept to be the concept of water is for it to be a concept of W. On the other hand, one might (as I do) prefer a view on which it is possible for the concept of water to exist, and hence possible for one to have the concept, even though there is no such stuff as water and hence no such kind as W. In this case, the meaning of the word 'water' could still be wide, for specification of the meaning might still require direct reference to some other contingent object or objects (besides the kind W). Perhaps, for instance, specification of the meaning requires direct reference to the planet Earth. "To be water," we might say, "is to be liquid that belongs to the same kind as the thirst-quenching stuff found in the streams and lakes of *this planet*." On this idea, notice, the

meaning of 'water' as used by our counterparts on Twin Earth would be exactly similar qualitatively to our meaning, but their meaning would still be different, since specification of their meaning would require direct reference to Twin Earth instead of Earth.

I have defended elsewhere a detailed proposal along these lines, a proposal that gives a precise explanation of the way in which the meanings of natural kind terms, and hence the concepts they express, are wide (see McKinsey, 1987, 1991b, and 1994). For our purposes here, though, it doesn't matter exactly which sorts of contingent objects are involved essentially in the meanings of such terms. The important point is that, on the sort of externalist account I've just sketched, the wideness of the meaning of, or the concept expressed by, a word like 'Cicero' or 'water' is a *logical* property of the meaning or concept. Thus on this account, to be the concept of Cicero, given the actual meaning of the word 'Cicero', a concept must be a concept of *a certain man*; to be the concept of water, given the actual meaning of the word 'water', a concept must be a concept involving a certain natural kind, or involving some other contingent object or objects. Notice that, on this externalist account, concepts like that of Cicero and water are of course related metaphysically to external objects; but they are so only because they are, in the first place, *logically* related to those objects. So it follows from this account that no one could possibly know *a priori* that he or she has the concept of Cicero or the concept of water. For to know these things *a priori*, one would have to know *a priori* that one has a concept of a certain man, and that one has a concept involving a certain natural kind, or other contingent object.

Of course, McLaughlin and Tye claim otherwise. They claim that if one has a thought involving the concept of Cicero, or a thought involving the concept of water, then one can easily know *a priori* that one has these concepts—no problem! But in making this claim, they seem to have forgotten that what we *say* when we say that someone has the concept of Cicero, or that someone has the concept of water, is determined by the actual meanings of the words 'Cicero' and 'water'. Once again, they seem to have forgotten the semantic basis of their own view. But as I have tried to make clear, it will follow from any intelligible externalist account of these meanings that the existence of the relevant concepts *logically* implies the existence of the external objects that individuate those concepts, and so it will also follow that one cannot know *a priori* that one has such a concept.

McLaughlin and Tye's glib insistence that one can easily know *a priori* that one has the concept of Cicero or the concept of water strongly suggests that they must be thinking of the properties of having such concepts as *narrow* properties, contrary to their own view. It is difficult to explain why they would be thinking this way. Perhaps they are confusing the wide property of having the concept of Cicero with the narrow property of having the concept expressed by the word 'Cicero', and the wide property of having the concept of water with

the narrow property of having the concept expressed by the word 'water'. In any case, by insisting that one can know *a priori* that one has the concepts in question, they are making it impossible for themselves to give any clear externalist account of what makes the property of having such a concept a wide property. For once again, they are forced to make the obscure claim that, in some inexplicable manner, the concepts in question are "metaphysically," though not logically, related to external objects. Once again, obscurity is the price they must pay for consistency.

WHY HAVE SO MANY externalists like McLaughlin and Tye uncritically endorsed the obscure idea that the concept or meaning expressed by a natural kind term like 'water' could be metaphysically but not logically dependent on external objects? Certainly there is no existing semantic account of such terms that implies or even remotely suggests this idea. Of course, concepts and words with given meanings can *refer* to things in the external world; but reference is a contingent semantic relation, not a metaphysical one. And of course some causal theories assert that certain concepts or meanings can only exist or be possessed when an agent bears a certain *causal* relation to given objects or kinds of object; but again, causation is certainly not a metaphysical relation. The only explanation for the idea's popularity that I can see is that many externalists have somehow managed to infer that the concept of water is metaphysically but not logically connected to H_2O , on the basis of the fact that it is metaphysically but not logically necessary that water is H_2O .⁷ But this is a grotesquely bad inference. The proposition that water is H_2O asserts an *a posteriori* metaphysical connection between *water* and H_2O , not between the *concept* of water and H_2O . Thus nothing whatever about the nature of the relation between the *concept* of water and H_2O follows from the modal status of the proposition that water is H_2O .

Moreover, the fact that it is not a logical truth that water is H_2O hardly shows that the connection between the concept of water and external objects is not that of logical necessity. As we have seen, the Twin Earth case makes it plausible to suppose that the concept or meaning expressed by 'water' can only be specified or individuated by direct reference to something external. Suppose that this something is water itself. Then as we have also seen, it would be a logical or conceptual truth that the concept of water exists only if water does. In this case, the connection between the concept of water and water itself—that is, H_2O —would be a conceptual or logical connection, even though it is not a logical truth either that water is H_2O , or that the concept of water exists only if H_2O does.

Notice that if the existence of the concept of water did logically or conceptually depend upon the existence of water, then since it is metaphysically but not logically necessary that water is H_2O , it would follow that it is metaphysically but not logically necessary that the concept of water exists only if H_2O

does. So it is possible to make sense of the idea that the concept of water is metaphysically but not logically connected to H_2O , provided that we make the prior assumption that the concept of water is logically connected with water in the first place.

But I do not think it is possible to make sense of the idea that a given concept or meaning could be metaphysically related to external objects, in the absence of any logical or conceptual relations that the concept or meaning bears to those objects. Kripke's (1972) suggestion that some propositions, such as the proposition that water is H_2O , are both necessarily true and *a posteriori*, makes sense because such propositions concern the essential *nature* of some type of substance, where such a nature can only be discovered by empirical investigation. But how could we make similar sense of the proposal that the concept or meaning expressed by a given word could bear such an *a posteriori* but necessary relation to things in the external world? Surely, the suggestion that some concepts and meanings could have hidden "natures" discoverable only by science (neurophysiology perhaps?), natures that would somehow necessarily connect these concepts and meanings to certain external objects, is quite unintelligible and should not be taken seriously. Surely, the truth is rather that if some concepts or meanings are indeed necessarily connected to external objects, they are so only because *we*, or the conventions of our language, have introduced or defined these concepts or meanings in part on the basis of reference to, or presupposition of, the external objects in question. Again, to make the existence of such necessary connections intelligible, we have to assume that these connections hold as a matter of logical or conceptual necessity.

SO FAR WE have seen that McLaughlin and Tye's claim that externalist theses are not about logical or conceptual dependency relations, results in a view that, while consistent with privileged access, is largely unintelligible. But it is important to realize that their endorsement of the principle of privileged access PAC results in a view that is also false regarding the semantic facts. In particular, all externalists who, like Burge (1988) and McLaughlin and Tye, defend compatibilism and explicitly endorse PAC, thereby commit themselves to the *denial* of semantic externalism SE, which I repeat:

Semantic Externalism (SE)

Many *de dicto*-structured predicates of the form 'is thinking that p' express properties that are wide, in the sense that possession of such a property by an agent logically implies the existence of contingent objects of a certain sort that are external to that agent.

For remember, as McLaughlin and Tye themselves agree, (PAC) is inconsistent with any externalist thesis on which having the property of thinking that p *log-*

ically implies the existence of external objects, and SE of course asserts that there are many such properties.

Now it is certainly strange to find contemporary defenders of externalism like Burge and McLaughlin and Tye, who are willing to explicitly commit themselves to the denial of semantic externalism. But what is most important is that this denial is just *false*. For SE is shown true by many simple cases of *de dicto*-structured predicates of the form ‘is thinking that p’ that contain small-scope proper names and indexicals. Many predicates of this sort do in fact typically express properties that are relational with respect to ordinary objects, and the possession of any such property of course logically implies the existence of the particular object in question.

Consider the case of Dave, a new graduate student who has not yet met Larry, the department’s Chair. Seeing a rumpled middle-aged man in old clothes busily cleaning the seminar room, Dave says ‘That janitor is really hardworking’. Hearing this, I turn to a colleague and say

(9) Dave is thinking that Larry is a janitor.

It seems intuitively obvious that in uttering (9) I would be using the name ‘Larry’ simply to refer to Larry, and to say that Dave is thinking that he is a janitor.

If the name ‘Larry’ had some sort of descriptive meaning in English, then perhaps it could be used to say something about the way Dave thinks of Larry. I have argued elsewhere (McKinsey 1999) that some ordinary proper names do in fact have descriptive meanings, and that as a result, uses of such names in cognitive contexts are nonrelational. However, I also believe, on the basis of Kripke’s (1972) famous Gödel-Schmidt case and others like it, that such names are very rare, and that most ordinary names have no descriptive meanings in public languages. In particular, it is clear that my use of ‘Larry’ in (9) has no descriptive meaning, and so it can function only to introduce its referent into what is said by (9). Hence (9) ascribes a property that is relational with respect to Larry: it says that Dave has an occurrent thought *about* Larry to the effect that he is a janitor.

It is even clearer that cognitive predicates containing small-scope demonstratives and other indexicals express relational properties. Consider

(10) Dave is thinking that *he* (or: *that man*) is a janitor.

Assuming that the occurrence of ‘he’ (or ‘that man’) in (10) refers to Larry, (10) like (9) says that Dave is having an occurrent thought about Larry to the effect that he is a janitor. Notice that both (9) and (10) ascribe relational cognitive properties, even though the occurrences of the relevant terms ‘Larry’ and ‘he’ (or

'that man') are assumed to fall both grammatically and logically in the scope of 'is thinking that'. Given this assumption, (9) and (10) are both *structurally* (or logically) *de dicto*, but due to the semantic character of the small-scope terms, both sentences turn out to be semantically relational, or *de re*.

Hence many *de dicto*-structured predicates that contain ordinary proper names and indexicals express properties that are logically wide, and thus such predicates show that semantic externalism (SE) is true, contrary to McLaughlin and Tye's view. Such predicates also provide straightforward counterexamples to PAC, also contrary to their view. For suppose that (9) is true. Then it immediately follows from PAC that Dave can know *a priori* that he is thinking that Larry is a janitor. But this consequence is false. Since what Dave allegedly knows *a priori* is relational with respect to Larry, Dave could just deduce from what he knows *a priori* that Larry exists, and hence he could also know *a priori* that Larry exists, which is clearly absurd.

SINCE PAC IS FALSE, it is an incorrect expression of the idea that we have a privileged way of knowing about our own thoughts. I have suggested elsewhere (McKinsey, 1994) that the correct principle would restrict the properties of a thought to which one has privileged access to those fundamental semantic properties that *individuate* the thought, in the following sense: A thought that a person x has in a possible world w is *individuated* by a property P just in case in any other possible world w' a person y would have the very same thought if and only if in w' y also has a thought that has P .

Then I would propose that the correct principle of privileged access is

Privileged Access to Individuating Properties (PAI)

It is necessarily true that if a person's thought is individuated by a given property Φ , then that person can in principle come to know *a priori* that he or she has a thought that has the property Φ .

Since, as we have seen, one never has privileged access to one's possession of logically wide properties, (PAI) implies that our thoughts are individuated only by logically *narrow* properties. I will call this principle

Metaphysical Internalism (MI)

It is necessarily true that if a person's thought is individuated by a given property Φ , then Φ is logically narrow.

I endorse both PAI and MI.

I have also said that I endorse semantic externalism (SE), the thesis that many *de dicto*-structured predicates express logically wide properties. SE is of course consistent with both PAI and MI. This is because, being solely a *seman-*

tic thesis, SE says merely that many cognitive predicates express logically wide properties; it says nothing about the metaphysical issue of whether or not these wide properties ever *individuate* our thoughts. Thus, the conjunction of SE and PAI provides a form of compatibilist view that is clearly consistent, and one that I think is true.

It is worth noting that at one point in their paper, McLaughlin and Tye assert that “the privileged access thesis is concerned with occurrent thoughts typed in as *fine-grained* a way as is necessary for the purposes of *any* rationalizing explanation” (1998, 293; their emphasis). I agree with this assertion, and it in fact provides one of the main motivations behind my (1994) proposal of PAI as a proper replacement for the false PAC. In my view, the properties that individuate our thoughts are precisely the properties that semantically characterize those thoughts in a maximally fine-grained way. This suggests that McLaughlin and Tye might themselves be willing to abandon PAC in favor of PAI, as I had proposed (1998, 295). However, they cannot consistently take this line. For as we saw above, they insist that a singular thought such as the thought that Cicero is an orator is individuated (in my sense) by the singular proposition that is the thought’s content (1998, 299). And, as they also correctly insist, no one can have privileged access to their thought’s having such a content. So McLaughlin and Tye are committed to the denial of PAI.

Those whose externalism, like mine, is restricted to a semantic thesis like SE are free to endorse the idea that we have privileged access to the properties that individuate our thoughts. But it seems to me that most externalists, like McLaughlin and Tye, have also wanted to endorse externalism as a *metaphysical* view about thoughts. These externalists claim that some thoughts are individuated by their wide contents, or by the wide property of having such a content. We might call this view

Metaphysical Externalism (ME)

In some cases, a person is thinking that *p*, the content that *p* is logically wide, and the person’s thought is individuated by the property of being a thought that has the content that *p*.

(By a “logically wide” content, I mean an abstract semantic entity, like a singular proposition, whose very existence logically implies the existence of contingent, typically concrete, objects.)

As we have just seen, however, anyone who endorses ME is precluded from endorsing what I take to be the most plausible principle of privileged access, namely PAI. This in fact is one of the main reasons why I believe that ME is false. There are other good reasons as well. Many philosophers, including myself, have pointed out that two persons’ thoughts that have the same singular proposition as content could nevertheless be different thoughts.⁸ This is because

a thought is not always individuated by its propositional content alone, but also by the agent's *way of thinking* of the content. I have also argued that two persons, in distinct possible worlds, could have the *same* thoughts, even though their thoughts have *different* singular propositions as contents (see McKinsey, 1994, 1999.) For these reasons, metaphysical externalism (ME) should be rejected in favor of a view that combines semantic externalism (SE) with metaphysical *internalism* (MI).

WHY HAVE SO MANY contemporary defenders of externalism come to ignore—and some to even deny—the semantic basis of their view? Perhaps one reason is that it has become common to state both externalism and the arguments for it in wholly nonsemantic terms. Consider the Twin Earth argument, for instance. When Putnam (1975) first described the Twin Earth example, he of course used it to make a semantic point, namely, that natural kind terms have wide meanings. But by now it is commonplace to state the moral of the example as if it had no semantic significance whatever. It is said that Oscar, a denizen of Earth, thinks that water is wet. But Toscar, who is Oscar's molecular duplicate on Twin Earth, has no thoughts about water at all. That is because all of the thoughts that Toscar expresses by use of the word 'water' are about a distinct kind of liquid composed of XYZ molecules instead of H₂O. Hence Oscar thinks that water is wet, but his molecular duplicate Toscar does not. Hence the property of having a thought with a given content—say, that water is wet—need not locally supervene upon the intrinsic physical properties of the person who has the thought.

Let us call a property that can thus fail to locally supervene upon a person's intrinsic physical properties, an "S-wide" property. Then the minimal consequence that follows from the above "short version" of the Twin Earth argument is a thesis that we might call

Weak Externalism (WE)

In some cases, the property of having a thought with the content that p is S-wide.

It has become commonplace for externalists to identify their view with this very weak thesis, and thus to identify their opponents (sometimes called "individualists") with those who (insanely) deny this weak thesis.⁹ This is a highly dubious practice. For one thing, I know of no philosopher who has ever actually *denied* WE, so it seems unlikely that the defender of WE really has any opponents at all.

Certainly, it is clear that no defender of internalism *should* deny WE, since there are paradigm cases of obviously narrow cognitive properties that do not supervene upon intrinsic physical properties. Consider *de se* properties, for in-

stance. The property P that we ascribe to Descartes when we say that Descartes is thinking that he himself is a thinking thing, is surely a *narrow* property if anything is. Yet P is S-wide. For P is the property of having a first-person thought about Descartes to the effect that he is a thinking thing. Notice that only Descartes could possibly have property P. In particular, Descartes' molecular duplicate on Twin Earth could not have P. Hence P is S-wide. Yet surely, P should count as *narrow*: Descartes' having P would logically imply the existence of no object external to him. Moreover, of course, anyone who has P (that is, anyone who is identical with Descartes) can know *a priori* that he has P, and so has privileged access to his possession of P, as Descartes himself first pointed out.¹⁰

There are many other types of logically narrow properties that are also S-wide. Consider, for instance, the *disjunction* of any logically narrow cognitive property N with a cognitive property W that is both logically wide and S-wide. The property N or W is logically narrow, since having it does not logically imply the existence of external objects (one can have N or W by having N). But this disjunctive property is also S-wide: for by assumption, W is S-wide, and so of two molecular duplicates x and y, x can have W while y does not; but suppose that neither x nor y has N; then x has N or W but y does not. So N or W is logically narrow but S-wide. Note also that in this case, x would not have privileged access to possession of N or W, since x has no privileged access to possession of W.

So S-wide properties are a heterogeneous lot. Some are logically narrow; some are logically wide. Some satisfy a principle of privileged access, and some do not. This means that versions of externalism like WE that make use of the notion of S-wideness are of very little philosophical interest. Again, no defender of any interesting form of internalism should want to deny WE. Moreover, WE is logically consistent with the explicit negation of semantic externalism (SE), so that one could consistently accept WE while insisting that absolutely every predicate of the form 'thinks that p' expresses only a logically narrow property. WE is also consistent with metaphysical internalism (MI) as well as with both principles of privileged access PAC and PAI. In fact, WE is such a weak thesis that, for every interesting principle concerning the externalism-internalism debate that I have so far identified in this paper, WE is consistent with both that principle and its negation. That is why I say that WE is philosophically uninteresting: it has no interesting logical implications regarding any traditionally held principles in the philosophy of mind.

But if the minimal consequence WE of the Twin Earth thought-experiments is so philosophically uninteresting, then why are these thought-experiments themselves so exciting? The answer is obvious. In the Twin Earth case, it is *not* the S-wideness of the property of thinking that water is wet that is philosophically interesting. Rather, what is interesting is the apparent explanation of *why* this particular property is S-wide. This explanation, as we saw earlier, would seem to imply that the property in question is S-wide because it is logically re-

lational with respect to contingent, external objects. This in turn implies the truth of semantic externalism (SE) and provides evidence in favor of, though it does not imply, the truth of metaphysical externalism (ME), both of which are philosophically significant principles.

But notice that to give this kind of explanation, one *must* do semantics. In the Twin Earth case, it is fairly clear how the explanation should go, since as Putnam originally described the case, it provides strong evidence that natural kind terms like ‘water’ have logically wide meanings. Once a clear account of these wide meanings is provided, we can understand the contribution made by a natural kind term to the meaning of a cognitive predicate containing the term such as ‘is thinking that water is wet’, and finally, we can obtain an understanding of what kinds of properties are expressed by such predicates and the precise sense in which these properties are “wide.”

So in order to have any sort of clear and philosophically interesting view, externalists *must* provide clear semantic explanations of the intuitions evoked by the thought-experiments that they use to motivate their stance. Yet as far as I know, no externalist has ever provided any such clear semantic explanation. And in fact, it seems that externalists have more recently stopped doing semantics entirely. Again, why has this happened?

This is only speculation, but it seems to me that most externalists have abandoned semantics out of an intense desire to avoid inconsistency with privileged access. Recall that PAC is the principle of privileged access endorsed by such externalists as Burge and McLaughlin and Tye. Also recall that PAC is straightforwardly inconsistent with semantic externalism SE. But if one abandons SE, then one is precluded from ever being able to give a clear semantic explanation of the wideness of any cognitive property. Having abandoned SE, externalists are then forced to make vague allusions to obscure and unexplained *a posteriori* metaphysical connections between cognitive properties and external objects, as a way of having a view as to what makes such a property “wide.” Or, as an alternative, externalists sometimes appeal to the notion of S-wideness as a way of stating their view. This has the advantage of providing a fairly clear and intelligible thesis in the form of WE, but unfortunately, the thesis is philosophically uninteresting in the absence of any further semantic explanation.

So externalists should give up their devotion to privileged access, at least in the form of PAC. After all, as we have seen, that principle is straightforwardly false anyway. Then having given up PAC, externalists can go back to doing good old-fashioned clear-headed semantics, and make some progress.

NOTES

1. See, for instance, Brueckner 1992.
2. As far as I know, the first to propose the metaphysical strategy was Burge 1988.
3. While I think that this is the most fundamental reason why the metaphysical strategy is

misconceived, I have elsewhere emphasized another serious problem, namely, that it results in a version of externalism that is trivial and uninteresting. See McKinsey 1991a, 1991b, 1994a, and (forthcoming).

4. Here and below I mean ‘logically implies’ in a broad sense that includes conceptual implication. See McKinsey 1991a, 14; 1991b, 152.
5. This *reductio* argument depends upon the following closure principle for *a priori*:

CLI Necessarily, for any person *x*, and any propositions *P* and *Q*, if *x* can know *a priori* that *P*, and *P* logically implies *Q*, then *x* can know *a priori* that *Q*.

The most common interpretations of my argument do not see it as relying on CLI, but rather on another very plausible closure principle, namely:

CAK Necessarily, for any person *x*, and any propositions *P* and *Q*, if *x* can know *a priori* that *P*, and *x* can know *a priori* that if *P* then *Q*, then *x* can know *a priori* that *Q*.

See, for instance, Brown 1995, Gallois and O’Leary-Hawthorne 1996, Boghossian 1997, McLaughlin and Tye 1998, Davies 1998, and Wright 2000. Although it is possible to give a version of my *reductio* that does appeal to CAK, the correct restatement would differ from any of the existing versions that use CAK. The difference is that all the existing versions that use CAK also rely on the (in my opinion false) assumption that externalist theses like (2) in the text are themselves knowable *a priori*. But neither my original (1991a) version that appeals to CLI, nor my restatement of it that appeals to CAK needs to rely on the (false) assumption that externalist theses such as (2) are knowable *a priori*. For details, see McKinsey (forthcoming).

6. For statements and defenses of the standard view, see, for instance, McKay 1981, Salmon 1986, and Soames 1987. I myself hold that the standard view is false (see McKinsey, 1986, 1994). However, I agree with defenders of the standard view that an ascription like (3) ascribes a property that is relational with respect to Cicero. See pp. 44–46 below, and also McKinsey 1999 and (forthcoming).
7. See for instance Brueckner 1992, 116. The possibility that the “metaphysical” externalists might be making this inference was suggested to me by Mark Huston.
8. See, for instance, Perry 1979, Fitch 1985, Loar 1985, Salmon 1986, and McKinsey 1986.
9. See, for instance, Brueckner 1995, 147, Gallois and O’Leary-Hawthorne 1996, 1, and McLaughlin and Tye 1998, 285.
10. See my discussion of the narrowness of *de se* properties in McKinsey 1991b, 146–57.

REFERENCES

- Boghossian, P. 1997. “What the Externalist Can Know *A priori*.” In C. MacDonald, B. Smith, and C. Wright (eds.), *Knowing Our Own Minds*. Oxford: Oxford University Press.
- Brown, J. 1995. “The Incompatibility of Anti-Individualism and Privileged Access.” *Analysis* 55: 149–56.
- Brueckner, A. 1992. “What an Anti-Individualist Knows *A priori*.” *Analysis* 52: 111–18.
- . 1995. “The Characteristic Thesis of Anti-Individualism.” *Analysis* 55: 146–48.
- Burge, T. 1988. “Individualism and Self-Knowledge.” *Journal of Philosophy* 85: 649–63.
- Davies, M. 1998. “Externalism, Architecturalism, and Epistemic Warrant.” In C. MacDonald, B. Smith, and C. Wright (eds.), *Knowing Our Own Minds*. Oxford: Oxford University Press.
- Fitch, G. 1985. “On the Logic of Belief.” *Nous* 19: 205–28.
- Gallois, A., and J. O’Leary-Hawthorne. 1996. “Externalism and Scepticism.” *Philosophical Studies* 81: 1–26.
- Kripke, S. 1972. “Naming and Necessity.” In D. Davidson and G. Harman (eds.), *Semantics of Natural Language*. Dordrecht: D. Reidel.
- Loar, B. 1985. “Social Content and Psychological Content.” Presented to the Twenty-fifth Annual Oberlin Colloquium in Philosophy; in R. Grimm and D. Merrill (eds.), *Contents of Thoughts*. Tucson: University of Arizona Press.
- McKay, T. 1981. “On Proper Names in Belief Ascriptions.” *Philosophical Studies* 39: 287–303.
- McKinsey, M. 1986. “Mental Anaphora.” *Synthese* 66: 159–75.

- . 1987. "Apriorism in the Philosophy of Language." *Philosophical Studies* 52: 1–32.
- . 1991a. "Anti-Individualism and Privileged Access." *Analysis* 51: 9–16. Reprinted in *The Twin Earth Chronicles*, edited by A. Pessin and S. Goldberg. Armonk, N.Y.: M. E. Sharpe, 1996, and in *Externalism and Self-Knowledge*, edited by P. Ludlow and N. Martin. Stanford, Calif.: CSLI Publications, 1998.
- . 1991b. "The Internal Basis of Meaning." *Pacific Philosophical Quarterly* 72: 143–69.
- . 1993. "Curing Folk Psychology of 'Arthritis'." *Philosophical Studies* 70: 49–63.
- . 1994. "Individuating Beliefs." *Philosophical Perspectives* 8 (*Logic and Language*): 303–30.
- . 1994a. "Accepting the Consequences of Anti-Individualism." *Analysis* 54: 124–28.
- . 1998. "The Grammar of Belief." In F. Orilia and W. Rapaport (eds.), *Thought, Language, and Ontology*.
- . 1999. "The Semantics of Belief Ascriptions." *Noûs* 33: 519–57.
- . (forthcoming). "Forms of Externalism and Privileged Access," presented to a conference on Anti-Individualism, Self-Knowledge, and Scepticism at the University of Bristol (February 1999), and to a conference on Internalism/Externalism at the University of Stockholm (April 1999).
- McLaughlin, B., and M. Tye 1998. "Externalism, Twin Earth, and Self-Knowledge." In C. MacDonald, B. Smith, and C. Wright (eds.), *Knowing Our Own Minds*. Oxford: Oxford University Press.
- Perry, J. 1979. "The Problem of the Essential Indexical." *Noûs* 13: 3–21.
- Putnam, H. 1975. "The Meaning of 'Meaning'." In K. Gunderson (ed.), *Minnesota Studies in the Philosophy of Science* 7. Minneapolis: University of Minnesota Press.
- Salmon, N. 1986. *Frege's Puzzle*. Cambridge, Mass.: MIT Press.
- Soames, S. 1987. "Direct Reference, Propositional Attitudes, and Semantic Content." In N. Salmon and S. Soames (eds.), *Propositions and Attitudes*. Oxford: Oxford University Press.
- Wright, C. 2000. "Cogency and Question-Begging: Some Reflections on McKinsey's Paradox and Putnam's Proof." *Philosophical Issues* 10: 140–163.

Acquaintanceless *De Re* Belief

Robin Jeshion

University of Southern California

THE *DE RE/DE DICTO* DISTINCTION AND ACQUAINTANCE

FAMINE HIT THE HEARTLAND and farmer Ralph's cows are dying. Ralph, who knows his cows, believes that Bessie is starving. His belief, which is based on his observations of Bessie's skinny state, is *de re*. It is manifest by his acceptance of the sentence

(1) Bessie is starving.

Doris, too, has been concerned about the famine. Muttering to herself, she says 'even the fattest cow is starving'. She does not determine that some particular cow is the fattest and that it is starving. She came to her belief by reasoning that the famine is so severe that even the fattest is starving. She has no interest in determining which cow is the fattest and thinking about that particular cow. Her belief is *de dicto*, and is manifest by her acceptance of the sentence

(2) The fattest cow is starving.

Ralph's belief is a canonical instance of *de re* belief. He stands in a direct perceptual relation to the object of belief, and he thinks of that object. Some theorists think of the content of these beliefs to be singular propositions—propositions whose constituents are individuals and properties. This is useful and, I think, largely right, and I will draw on it later. However, commitment to this metaphysics of the content of *de re* belief may ultimately be dispensable.¹ Doris's belief is a canonical instance of *de dicto* belief. Her belief is not directly about any particular object at all—not even in any extended sense. She is not even attempting to think about a particular object. The belief is, as it were, wholly conceptual. Some think of the content of these beliefs to be general propositions. But again, this is theory (a theory I think is largely correct) but may not be needed. What is important is that we have two distinct types of belief.

While it is appropriate to say that Ralph and Doris's beliefs differ in virtue of having different types of content—singular and general propositions—*by itself*, this point about the metaphysical/semantic features of the belief content sheds little light on the nature of the difference between *de re* and *de dicto* belief. It leaves unexplained precisely what needs explaining: What is it to believe, and what are the conditions on believing, a singular proposition? What is it to believe, and what are the conditions on believing, a general proposition? What is the *psychological* difference between these two types of beliefs? I will not here be attempting to offer a full theory that answers these questions. But what I say will, I hope, point us in the right direction.

Of course, characterizing the difference between *de re* and *de dicto* belief is notoriously difficult. In particular, it is difficult to know how to extend past the canonical cases for *de re* belief, cases in which agents stand in a direct perceptual relation to the concrete object of thought. If, twenty years down the road, Ralph thinks to himself that Bessie was a fine cow and my how she did suffer, yet he has completely lost all memory-images of Bessie, does he have a *de re* belief of Bessie? Many philosophers think the answer is yes, but it is not uncontroversial.² As long as Ralph was once directly perceptually acquainted with Bessie and his memory preserved information about her in the proper way, his belief can be *de re*, despite the absence of Bessie-memory-images. If Ralph tells Rhoda that Bessie is starving, yet Rhoda has never herself perceived Bessie, can Rhoda's belief, as manifest by her acceptance of (1), be *de re*? This is much more controversial. Still, here again, many think that the answer is yes—as long as Rhoda receives the news about Bessie from a causal communication (or information) chain that originates in someone who was directly perceptually acquainted with the cow,³ for these beliefs seem to have the same sort of structure as the canonical cases of *de re* belief. Ralph and Rhoda at least seem to be thinking of a particular object. Their way of taking Bessie does not seem to be conceptual. But now theorizing becomes exceptionally tricky: in what sense are these beliefs about or of the object? How can they be when the agents lack perceptual representations of the object, and, in the case of Rhoda, never even had such a representation? Problems multiply when we move away from concrete objects, and consider the possibility of having *de re* beliefs about mathematical entities, fictional characters and fictional works themselves, nonexistent objects of myth, false scientific theory, and hallucination, and any other nonconcrete entity.

Despite these difficulties, there is widespread agreement about one issue (at least for theorists who do not reduce *de re* to *de dicto*⁴): Acquaintance is a necessary condition on *de re* belief about concrete objects. Kaplan (1969) thinks we need to be “*en rapport*” with the object. Bach (1987), Boer and Lycan (1986), Burge (1977), and Recanati (1993) all maintain that there needs to be a “real relation” between the believer and the concrete object of thought. Ditto for Lewis (1983) and Evans (1982, 1985). Salmon (1987), Soames (1995), and Donnellan

(1979) embrace the necessity of acquaintance as well.⁵ It is rare to have such a meeting of (these) minds.⁶

Now, the notion of acquaintance here is, like the notion of *de re* belief, a term of art. And while I will not venture to explicate the notion, there are several things we can say about it. First, acquaintance is one thing, causal connection another. If Rhoda fertilizes her tomato plant with droppings from Ralph's farm, she is, to be sure, causally connected to Bessie; but she is not thereby acquainted with her. So causal connectedness is not sufficient for acquaintance. Is it necessary? Can you be acquainted with something yet fail to be causally connected with it? Perhaps. Perhaps we are acquainted with some mathematical objects yet are not causally connected with them. Or take Russell's (1956) favorite objects of acquaintance: sense data, universals, and (possibly) ourselves. It seems that only in a very extended sense can we say that we are causally connected with such objects. In any event, we need to keep the two notions separate.

Second, there are different ways of being acquainted with an object. One can be acquainted with it through a communication-chain, or through direct perceptual contact, as in our canonical cases. If we are acquainted with mathematical objects or universals, this must be a different way altogether.

Third, it seems that acquaintance comes in degrees. One can be more or less well acquainted with an object, depending on the amount or variety of contact one has had with it.

Fourth, being acquainted with an object is distinct from possessing knowledge-who or knowledge-which with respect to that object. Knowledge-who and knowledge-which are contextually sensitive.⁷ Within my reading group, I may know who Jeanette Winterson is: I have read many of her novels and know she is the author of them. But in the context in which I attend a party in her honor, I lack such knowledge, for I do not know her from Adam (or Eve). Acquaintance is not context sensitive in this way. I have acquaintance with Winterson through her novel and communication-chains and, no matter what the context, my inability to recognize her does not impinge upon that relation.

Within the literature, acquaintance is most often used as a catchall for a necessary condition on *de re* belief. So, those who countenance communication-based *de re* belief about concrete objects would maintain that we are, say, acquainted with Matisse even though we ourselves lack a direct perceptual relation (as in the canonical cases) with him. Restrictions on the range of objects of acquaintance result in restrictions on the class of *de re* beliefs. Russell's curious views are a case in point. Because he thought the *only* objects of acquaintance are sense data, universals, and (again, possibly) ourselves, they are, for him, the only possible candidates for objects of *de re* beliefs.

Like most everyone, I do not follow Russell in these restrictions on acquaintance. I shall be assuming a very liberal notion of acquaintance: that we are acquainted with some concrete objects, as in the canonical cases, and that we can

be acquainted with such objects in other ways, including via memory- and communication-chains. I shall be assuming (with many) that a minimal degree of contact with the object suffices for acquaintance. Most important, I shall also be assuming (with the masses) that if you and no one with whom you are memory- or communication-chain-related to ever stood in a canonical (direct perceptual) relation to a concrete object, you are *not* acquainted with that object.

Our question is: in this sense of acquaintance, is acquaintanceless *de re* belief about concrete objects possible?

The answer that I shall propose is, in short, yes. Since this goes against the dominant view, I have a lot of explaining to do. But let me hasten to add that my story of how we can have acquaintanceless *de re* belief accounts for the fundamental importance of acquaintance to *de re* belief. My view is that the canonical cases in which one is perceptually acquainted with the concrete object of one's belief are distinctive: one has a perceptually attained mental representation of the object. But what is essential to their being *de re* is not the acquaintance relation *per se*, or even the mental representation itself, but rather the role that beliefs of this kind play in cognition. Some acquaintanceless cases are *de re* precisely because they too play this role in cognition. The key idea is that the acquaintanceless cases are parasitic on the canonical acquaintance cases, and what ties them together is the function of proper names in thought.

THE IMPORTANCE OF OUR QUESTION

Before we get underway, it is worthwhile asking ourselves this: why should we care about the possibility of acquaintanceless *de re* belief about concrete objects? There are at least four in some ways overlapping reasons why we should be interested.

General Theory of De Re Belief

The first point is obvious. If we find that there is no reason to embrace a condition of acquaintance, we will be forced to reexamine our best analyses of what distinguishes *de re* and *de dicto* belief. Most theorists that embrace the necessity of acquaintance also, in effect, work it into their characterization of, or into the essential condition on having, *de re* belief.⁸ If my thesis is correct, such analyses will have to be abandoned, and we will have to confront the hard but fascinating task of (re)analyzing the notion.

Nonconcrete Objects

Second, if we can understand how we can have *de re* beliefs about concrete objects with which we are unacquainted, we may learn something about our capacity to have *de re* beliefs about certain nonconcrete entities, like fictional characters and mathematical objects. One might think that the possibility of *de re* beliefs about nonconcrete object hinges on their existing, and hence on actually

having *bona fide* ontological status as abstract entities. I will not go into the metaphysical issue here (see notes 3, 13, 14.) For now, my point is just that there is good reason for thinking that some of our beliefs about fictional characters and numbers are *de re*. There seems to be an intuitive distinction between *de re* and *de dicto* belief concerning fictional entities. An avid fan of the Holmes stories will accept the sentence

- (3) Sherlock Holmes smokes a pipe.

There is a very strong pull to say that this individual's belief, as manifest by accepting (3), is different in kind from the belief of an individual that is manifest by accepting the sentence

- (4) The most famous fictional detective smokes a pipe

where we suppose that this individual has never read or heard a Conan Doyle story, has never been in contact with anyone who has, and has no interest in thinking about the most famous detective in fiction. All the detectives that this person has had contact with smoke a pipe, and this leads her to generalize about detectives, real and fictional, and so form a basis for her (4) belief. The difference between a (3) belief and a (4) belief seems to closely parallel the difference between the (1)/(2) beliefs of Ralph and Doris.

The distinction also seems apparent in our beliefs about numbers. One might believe that

- (5) 17 is prime.

Alternatively, one might believe that

- (6) The seventh prime is prime

without in this case thinking of, or about, 17.⁹

The parallelism between these examples and our canonical examples is, I think, solid *prima facie*—but certainly defeasible—grounds for thinking that the *de re/de dicto* distinction applies quite readily to our fictional and mathematical beliefs. In my view, the analysis of the distinction ought to be general enough to apply to beliefs about both concrete and nonconcrete objects.¹⁰ An analysis of how it is possible to have *de re* beliefs about concrete objects with which we lack acquaintance would, then, be most welcome because it may shed light on how we can have *de re* beliefs about these nonconcrete objects. I will not be presenting a theory explaining how we have *de re* beliefs about mathematical and fictional entities, but I do think that what I say may at least suggest a fruitful line of investigation.

Nonexistent Entities

Third, our investigation may help in coming to grips with (the possibility of) *de re* beliefs about nonexistent entities. Consider the example of Vulcan. To account for irregularities in the orbit of Mercury, Babinet postulated the existence of, and named, a planet that circled the sun within Mercury's orbit. He called this planet 'Vulcan' and, along with other astronomers of the day, including Leverrier, attempted to discover it telescopically. We now know that no such planet exists, but scientists of the time were committed to its existence. Such astronomers' beliefs would be manifest by their acceptance of sentences containing 'Vulcan,' as in

(7) Vulcan circles the sun.

Cases like this aggravate theorizing about *de re* belief. Our intuitions pull in two directions. We have an intuition that the astronomers' Vulcan beliefs have the same form as other *de re* beliefs. The belief is not *de dicto* in that it is not fully conceptualized. Furthermore, the nonexistence of the planet seems inessential to the characterization of the subject's psychological state; it seems that the psychological state would be the same even if the planet existed. On the other hand, since *de re* belief is supposed to be belief that is directly *about* or *of* an object, it seems impossible to have a *de re* belief about something like Vulcan that does not exist. No object, no *de re* belief.¹¹ An explanation of how to have acquaintanceless *de re* belief about concrete objects may well assist us in demonstrating the *de re* character of (7) beliefs while retaining the *aboutness* that seems essential to *de re* belief.¹²

Descriptive Reference-Fixing

Fourth, acquaintanceless *de re* belief is fundamentally connected to a problem that was first raised by Kripke (1980)—the so-called problem of the contingent *a priori*. Recall the story of Leverrier and Neptune. No one in Leverrier's time perceived the planet we now call 'Neptune', but they did have scientific evidence that a certain planet was causing perturbations in Uranus's orbit. Armed with evidence of the existence of another planet, Leverrier introduced the name 'Neptune' into the language by stipulating that the term is to refer to the planet causing those perturbations. By hypothesis, Leverrier intends to fix the reference of the term, not necessarily to assign it a meaning. He is treating the introduced term like other proper names and, according to the theory, and to much semantic theory since Kripke, 'Neptune' is then a rigid designator and consequently, in all counterfactual situations, it denotes the same entity that it denotes in the actual world. Kripke claimed that Leverrier is now in quite a remarkable position. He can have *a priori* knowledge of the contingent proposition expressed by the sentence, 'If there is a planet causing the

perturbations in Uranus's orbit, then Neptune is'. But it has seemed to many that no one, stipulator included, can have *a priori* knowledge of a contingent proposition.

How is this problem of the contingent *a priori* related to our topic of acquaintanceless *de re* belief? Getting clear on this requires getting clear on the problem that Kripke set forth by examples. I have previously argued for a certain way of understanding the problem.¹³ I shall sketch these arguments just enough to properly set the stage.

The problem Kripke presented is not essentially about the modal status of the proposition allegedly known *a priori*. The basic reason for this is that there are cases that raise the same perplexing issues, yet the proposition supposedly known *a priori* by the stipulator is a necessary truth. Imagine that a scientist has some general evidence for thinking that there is an element with atomic number 121 and stipulates that the name 'Angelesium' is to refer to the (as yet undiscovered) element having atomic number 121. Then it seems that the scientist can know *a priori* that Angelesium is the element with atomic number 121, if anything has that atomic number. On the assumption that atomic numbers are essence-determining properties, the proposition known *a priori* (allegedly) is necessary, but the alleged *a priori* knowledge hardly seems kosher. The case raises exactly the same issue as the Neptune case.

It is natural at this stage to think that what distinguishes these cases is that the stipulator can have *a priori* knowledge of the proposition while nonstipulators cannot. This move locates the problem in the *a priori* itself. But *a priority*, in and of itself, is not the heart of the problem. There are cases in which stipulator and nonstipulator are on the same footing with respect to *a priority*, yet the problem remains. Imagine that a mathematician introduces a name 'N' to refer to the 69th prime. It seems that such an agent could know that if there is a 69th prime, then N is the 69th prime. Both stipulator and nonstipulator seem to have *a priori* knowledge. But we are still puzzled by the stipulator's alleged knowledge.

I think that these points suggest that the fundamental philosophical problem pertains to the possible epistemological consequences of the act of stipulative reference-fixing with a definite description. While Kripke introduced the problem in a way that may presuppose Millian semantics, in my view, the problem can and should be stated completely independent of any semantic theory. It is no more a problem for Millians than for neo-Fregeans. We have the following two theory-neutral claims about the cases:

Stipulative Descriptive Reference-Fixing (SDR): There appears to be a particular possible, and possibly actual, linguistic phenomenon—stipulative descriptive reference-fixing—in which an agent introduces a name 'N' into the language by stipulating: 'N' is to refer to the F.

Epistemic Privilege of Reference-Fixing (EPR): The act of descriptive reference-fixing appears to put the stipulator in a position to be non-inferentially *a priori* justified in believing the proposition expressed by the sentence ‘N is the F, if there is a unique F’.

The problem before us is a challenge to any full theory of mind and language to account for SDR and EPR. It is incumbent upon such theories to either countenance them as genuine and explain how they are possible; or to maintain they are not genuine and explain them away.

Now, the issue of acquaintanceless *de re* belief arises in connection to the Millian’s attempts to account for SDR and EPR. Millians maintain that the sole semantic content of a proper name is its referent. Names lack semantic descriptive content. For many Millians, the proposition expressed by a sentence ‘N is the F’ is a singular proposition $\langle O, P \rangle$, where O is the object denoted by ‘N’ and P is the property denoted by ‘the F’.

It is a further thesis—but one that is upheld by the most prominent Millians—that the content of an agent’s belief that is manifest by the agent’s acceptance of the sentence ‘N is the F’ is just the singular proposition: $\langle O, P \rangle$. Such beliefs will be *de re*—for the object O itself is a constituent of the belief content, such beliefs are directly about, or of, the individual O.¹⁴

Suppose S introduces ‘N’ into the language by fixing its reference with the definite description ‘the F’. And suppose that there is a unique F. Then S allegedly has *de re a priori* knowledge of a particular object O to the effect that O is the F, if anything is. Leverrier, for example, will have a *de re* belief about Neptune that it is the planet causing the orbital perturbations, if anything is, and will know this *a priori*.

This is thought to be impossible. How, after all, could anyone know that *a priori*, when there is no constitutive or conceptual relation between Neptune and the property of causing the orbital perturbations? And how, after all, could Leverrier even come to have a singular belief about Neptune when he has had no acquaintance with Neptune at all, and neither has anyone in his linguistic community?

The connection between acquaintanceless *de re* belief and this problem should be apparent. It is central to the problem raised by Kripke because the most widespread Millian solution is to deny that the stipulator ever attains a *de re* belief. That is, their solution to the problem is to countenance SDR as genuine, and to claim that no *de re* belief is ever attained—for the stipulator (and no one in his linguistic community) stands in a “real relation” to the object. Just by stipulating that ‘N’ is to refer to the F, it is not possible for the stipulator to thereby have a *de re* belief about the object O that it is the unique F. No acquaintance, no *de re* belief.¹⁵

The standard line to explain away EPR is to claim that only metalinguistic

knowledge is achieved. Donnellan famously gave an argument for this position. I think the argument does not work,¹⁶ but the position itself is, nevertheless, not implausible. But I am not convinced that it is right. There are roughly two reasons. At an intuitive level, the appeal to metalinguistic belief to explain away EPR is far from convincing. It certainly seems that the stipulator's belief is about nonlinguistic entities, but this point is highly defeasible. The second is more significant. The Millian who embraces the following theses runs into a theoretical difficulty:

Possibility of Stipulative Descriptive Reference-Fixing: It is possible to introduce a name 'N' into the public language by stipulating that its reference is to be fixed by the definite description 'the F'.

Purist Millianism about Proper Names: For all proper names, the sole semantic content of a name 'N' is its referent O.

Skepticism about Descriptive Reference-Fixing Generated De Re Belief: Just by stipulating that 'N' is to refer to the F, it is not possible for the stipulator to thereby have a *de re* belief about the object O.

The difficulty is that cases of descriptive reference-fixing appear to mark out counterexamples to a very natural thesis about linguistic understanding:

Accessibility of Content: For all expressions E in the language L, and all sentences S in L expressing some proposition P, if an agent A has semantical understanding of all the expressions E contained in S, then if A were apprised of all the relevant contextual information, A could have an attitude having P as its content.

The stipulator introduces the name 'N' into the language L, and hence, there are sentences containing 'N' in the language expressing propositions that no one, not even the stipulator, can grasp. Yet it seems that, given that names lack semantic descriptive content, the stipulator has all the understanding needed to grasp those propositions. The conditions for Accessibility of Content obtain in our cases. That is, it seems that the following thesis holds:

Understanding Millian Names: In cases of descriptive reference-fixing, the stipulator understands the sentence, 'N is the F, if anything is,' and is aware of any features of the context relevant to the determination of the content of that sentence.

These five theses are jointly incompatible. The Millian must give up at least one of them.¹⁷

Because of these problems, it may be in the Millian's best interest to con-

sider a nonskeptical stance about descriptive reference-fixing generated *de re* belief, even in cases in which there is no acquaintance relation between stipulator and the object satisfying the description.

HOW TO HAVE *DE RE* BELIEF WITHOUT ACQUAINTANCE

I want to argue that it is possible to have acquaintanceless *de re* belief and that such beliefs are, in a certain sense, generated from the descriptive reference-fixing act. This is, let me remind you, Kripke's line. He says:

[Imagine that] a rigid designator 'a' is introduced with the ceremony, 'Let "a" (rigidly) denote the unique object that actually has the property F, when talking about any situation, actual or counterfactual.' It seemed clear that if a speaker did introduce a designator into a language in this way, then *in virtue of his very linguistic act*, he would be in a position to say 'I know that Fa', . . .¹⁸

While Kripke seems to think there is something significant about the linguistic act, he does not offer us an account of why it contributes to making possible acquaintanceless *de re* belief. That is my aim here.

My argument is in two parts. First, I shall argue that there are fairly strict conditions on descriptive reference-fixing, and that only if these conditions are met is the stipulator even a candidate for having a reference-fixing generated *de re* belief about the object. This helps exclude certain cases where it is intuitively implausible that the stipulator has a *de re* belief about the relevant object. It locates the difficulty in a failure to actually introduce a name into the language, as opposed to a failure of the stipulator to have a *de re* belief. The analysis also reveals that satisfying the conditions for such a stipulative act puts one in a preferred position for having a *de re* belief about the relevant object. Second, I shall argue that if the conditions on descriptive reference-fixing are met, then the descriptive reference-fixing act itself can alter the way in which we think of the object. It can help shift thought from a *de dicto* belief to a *de re* belief.

Conditions on Stipulative Descriptive Reference-Fixing

Kripke (1980) isolated 'one meter', 'Neptune', and 'Jack the Ripper' as names in natural language whose references were fixed by descriptive reference-fixing. He also introduced the mathematical example of π tentatively suggesting that ' π ' is a name of a real number whose reference was fixed by the description 'the ratio of the circumference to the diameter of a circle'.

In his classical paper on the subject, "Reference and Contingency," Evans (1985) remarked: "Very few names which naturally occur in ordinary language can be regarded as descriptive names." He continues,

Nevertheless, no matter how rare examples may be, it would appear always to be open to create descriptive names by stipulation. For example, we might stipulate:

Let us use 'Julius' to refer to whoever invented the zip,
and, governed by such a stipulation, 'Julius' would appear to have the properties of a descriptive name.¹⁹

Evans, of course, took a Fregean line on the semantic content of descriptive names.²⁰ But the Julius case and cases like it involving “free” introduction of names into the language inspire skepticism among Millians. After all, the examples suggest that we can create *de re* beliefs at will, simply by stipulating: ‘N’ is the F. Evans’s oft-quoted remark—“We do not get ourselves into new belief states by ‘the stroke of a pen’” (in Grice’s [1969] phrase)—simply by introducing a name into the language—borders on being a platitude.²¹

If introducing a name into the language was executed simply by the stroke of a pen, I would surely agree. And I do agree that, as described, there is no *de re* belief generated in the Julius case. But my view on why no such belief was generated is that no name has been introduced into the language. There has been no act of descriptive reference-fixing. To be sure, we have the outward appearance of such reference-fixing, but there was no such act.²²

In my view, within this debate, philosophers have tended to forget that all acts of naming—ostensive and descriptive alike—are genuine performatives, oftentimes explicit performatives. Consequently there are conditions on successfully executing the act—what Austin (1962) aptly called felicity conditions.²³

There are numerous conditions that must be met for a speech act to constitute an act of naming. I will not go into them all (some are the relatively boring conditions that naming shares with other speech acts; others are more interesting—I am sure I do not have a full grip on these). But I want to draw your attention to some distinctive conditions on felicitous naming that are relevant to our concerns here, and that specifically concern the cognitive and communicative function of names.²⁴

In a case in which an agent S aims to introduce a name ‘N’ into her idiolect by fixing its referent, S succeeds in doing so only if the following *Sincerity* and *Psychological Neutrality* conditions obtain.²⁵ (I give formulations for both ostensive and descriptive reference-fixing; though the key idea is the same for both varieties of reference-fixing, the separate formulations are needed.)

Sincerity (Ostension): S intends for ‘N’ to name object O and to use ‘N’ as a name for O.

Sincerity (Description): S intends for ‘N’ to name the F, whatever object it is, and to use ‘N’ as a name for it.

Psychological Neutrality (Ostension): S introduces 'N' for object O because S aims to think about and speak about O by mentally tokening 'N', without necessarily thinking about O via any particular mode of presentation.

Psychological Neutrality (Description): S introduces 'N' for the F because S aims to think and speak about the object O that is the F by mentally tokening 'N', without necessarily thinking about O via any particular mode of presentation.

Notice that these are not two distinct conditions. *Sincerity* is weaker than *Psychological Neutrality*, but *Psychological Neutrality* entails *Sincerity*. Treating them separately helps highlight *Sincerity*. (When I do not specify either ostension or description, I intend both.)

Sincerity is, interestingly, a condition that Searle (1969) denies. He claimed that no special intention is needed for a felicitous act of naming. He lumps naming and greetings together as speech acts that lack sincerity conditions.²⁶ But I think that he was wrong on this score about naming. A parent who had no intention whatsoever to use 'N' as a name for her child yet who utters or thinks the words "I name you 'N'", would not, I think, be naming her child. The act was hollow. No name enters her idiolect.

There are complications on this condition if you try to extend it to the non-idiolect case of introducing a name into the public language. Imagine a context in which I am forced, perhaps by religious leaders or the state, to participate in a naming ceremony in which I am to publicly utter "I name this child Abraham (or Napoleon, or Adolf)." Yet I do not want to name him 'Abraham' and have no intention at all to use 'Abraham' as a name for him. Have I named my child 'Abraham' by making the utterance in the given context? This is not entirely clear to me. The act seems to be in some way infelicitous, but maybe the naming takes place in any event. It may be that in virtue of the relevant uptake—the fact that others in attendance regard my utterance as sincere—the naming occurs. But one thing seems clear: if I lack the intention *and* the rest of the community is in the know about this, no naming occurs.

With regard to *Psychological Neutrality*, my claim is that to introduce a name, agents must have a reason for doing so, one that accords with the function of names—as vehicles for thinking about objects in a way that requires no particular mode of presentation of the referent. I do not, however, hold that the namer must possess this intellectualized account of the function of names and their psychological neutrality as her reasons for naming the object. A parent's own (internal) reason why she names her child might just be that people name their children. But I think that nevertheless, this must be her reason for introducing the name, even if she never conceives it as such.

There are, I think, two other principles (not strict conditions) concerning

naming. One is general, applying to both ostensive and descriptive reference-fixing. The other governs descriptive reference-fixing only and concerns the purpose of descriptive reference-fixing and the relative advantage of ostensive over descriptive reference-fixing.

Single Tagging: Fix the reference of a term only if, so far as you know, the named object is not already named.

Primacy of Ostension: Fix the reference of a term descriptively just in case, so far as you know, you cannot do it ostensively.

Single Tagging is, no doubt, oversimplified, and there will be exceptions to the rule—think of pet names, names used just for business, and pen names. But I think *Single Tagging*, or some modification of it, must obtain in order to enable names to function as *common* ways of thinking of objects in a psychologically neutral way.²⁷

Primacy of Ostension expresses the fact that reference-fixing via ostension has a type of priority over reference-fixing via description. The psychological neutrality of naming is rooted in ostensive (demonstrative) reference-fixing. If you name your pet ferret ‘Willaby’, your demonstrative identification of your ferret is from a definite perspective (yours, at that time, in relation to your ferret). Yet you do not think about your ferret as that object that stands in such-and-such-perceptual relation to you. You think of him directly. The name ‘Willaby’ inherits that psychologically neutral means of thinking of Willaby. The psychological neutrality of names stems from ostensive reference-fixing. For this reason, ostension is a first choice when we introduce names. By contrast, naming via description’s capacity to generate psychological neutrality is parasitic on naming via ostension (more on this later). It is therefore a second choice, when psychological neutrality is desired, yet ostension is unavailable or will not suffice.²⁸

In any event, I think that these points indicate that Evans’s Julius example is highly suspect. With respect to *Sincerity*, I doubt that Evans himself had any intention to use ‘Julius’ to speak about whoever invented the zip and I doubt that any philosopher upon hearing the example took Evans as seriously intending to think of that individual. So ‘Julius’ never entered Evan’s idiolect or the less-restricted language of the community of philosophers. Furthermore, with respect to *Psychological Neutrality*, Evans had no legitimate reason for introducing ‘Julius’ into the language. This is shown by the fact that he had no real interest in the inventor of the zip, whoever it is—no interest in discovering him or communicating with others about him. He sought no communicative or cognitive advantage by introducing the name, and consequently he lacked a real reason for introducing a psychologically neutral way of thinking about the inventor of the zip. The act was merely artifice, and consequently never gets off the ground. These points are further reinforced by noticing that *Single Tagging* and

Primacy of Ostension are also violated. Given that Evans had good reason to think the zip inventor already had a name, if he really wanted to think about the zip inventor in a psychologically neutral way, he should have been moved to investigation (and consequently to discover his name), not to descriptive reference-fixing.²⁹

What I have done at this stage is to rule out some cases in which it is indeed intuitively implausible that a *de re* belief is generated. But I have done so by showing that in these cases, no reference-fixing occurs. We are not “free” to introduce names into the language in the sense that we are constrained by our intentions,³⁰ which are, in turn constrained by our cognitive and communicative goals.³¹ Things must be right with the agent (and context) for a name to secure a reference.

Generating De Re Belief from Descriptive Reference-Fixing

I have not, as yet, said how it is that descriptive reference-fixing enables the stipulator to in fact have *de re* beliefs about the object satisfying the description. This leads me to the second point in the argument: the descriptive reference-fixing act itself can foster a shift in thought from *de dicto* to *de re*, and thereby enable the stipulator to get into a different belief state.

My claim is that if the *Sincerity* and *Psychological Neutrality* conditions are met (as they must be in our cases), descriptive reference-fixing can contribute to altering the stipulator’s beliefs: if the stipulator’s beliefs initially have as their content general propositions, they can be transformed to singular propositions via the act of descriptive reference-fixing. The agent who intends to use ‘N’ as a name for the F and who does so because she aims to think of that object in a psychologically neutral way will have her psychology altered by descriptive reference-fixing.

The argument runs as follows: By *Sincerity* and *Psychological Neutrality*, our stipulator aims to think about the object that is the F by mentally tokening ‘N’. So, once the reference-fixing occurs, the stipulator’s subsequent uses of ‘N’ are mental tokens of the name. These mental tokens of ‘N’ function as *de re* modes of presentation of the object O. They do so because of the interplay between the way in which the name-type ‘N’’s reference has been fixed, the stipulator’s intentions, and the way in which our thought is tied to logical/semantic form and its symbolic representations: The reference of ‘N’ is determined not by its meaning (by hypothesis, it has none), but rather by the name’s having had its reference fixed. ‘N’ refers to that object that is the F. The stipulator knows both of these points and her thought is responsive to them.³² By mentally tokening ‘N’, the stipulator uses ‘N’ just as she would any name whose reference was fixed (by ostension), as standing for—as a symbolic *de re* representation for—its referent. And the mental tokens of the name in fact function in the same way as mental tokens of a name whose reference was fixed by ostension—as symbolic

de re representations for ‘N’'s referent. Their capacity to do so is parasitic on the capacity of ordinary names (whose reference is fixed by ostension) to do so. Though they will not be causally tied to any initial perceptual representations of the referent, as they usually do in ostensive reference-fixing, they nevertheless function as *de re* modes of presentation. Thenceforth, the stipulator does not, and need not, think of the object descriptively (satisfactionally) as the F. Her mental tokens of the name suffice for her to think of the object directly. Consequently, the stipulator's belief content is the singular proposition $\langle O, P \rangle$.³³

The key idea here is that using names instead of descriptions can alter one's psychology. It alters the form, or structure, of the stipulator's thoughts and beliefs about the object. The Grice-Evans point that we do not produce new beliefs simply by the stroke of a pen (i.e., by introducing names into the language) is rhetorically potent, but an overstatement. To be sure, descriptive reference-fixing generates no new *information* about the object, but this does not entail that the introduction of the name fails to bring about a psychological change. Imagine that agents A_1 and A_2 receive the same perceptual experiences and hear the same reports of information. Informationally, they are twins. Suppose that, initially, both have only *de dicto* beliefs about the object that is the unique F. Later, with respect to a certain name ‘N’ and the definite description ‘the F’, A_1 satisfies *Sincerity (Description)* and *Psychological Neutrality (Description)*, and that A_2 does not. Suppose also that A_2 does not intend to use any demonstrative or pronoun to refer to the F, does not intend to think of the F in a psychologically neutral way at all. I maintain that in virtue of A_1 's fixing the reference of ‘N’, A_1 's belief about the referent of ‘N’ is *de re* while A_2 's is merely *de dicto*. In this way, the possibility of having a *de re* thought is independent of one's informational state.

It is worthwhile reflecting on the metaphorical but still suggestive model of the mental file folder.³⁴ Initially, a stipulator has a single (or a series) of *de dicto* beliefs about the object, which are unorganized or ununited in cognition. By introducing the name into the language, the stipulator opens and labels a new mental file folder as a repository of information about the object. No new information is thereby deposited. But the creation of the file itself is nevertheless a significant change in the stipulator's cognitive architecture. For now her beliefs about the object have the same form or role in cognition as many of her other beliefs that are canonical instances of *de re* belief. What distinguishes *de re* thought is its structural or organizational role in thought; acquaintance, and any evidential or epistemic relation, is inessential.³⁵

I want to bolster these points by arguing that if we can have *de re* beliefs via a communication-chain, then we can have them in the more controversial acquaintanceless cases. It is specially directed to Millians (such as Salmon, Soames, and Donnellan) who want to allow the former while denying the possibility of

the latter. My argument employs Kent Bach's (1987) analysis of the mechanism for generating *de re* beliefs via communication chains.³⁶ Though Bach is not himself a Millian, his analysis of the mechanism should be congenial to Millians, and it is, in my view, the most developed account in the literature. Perhaps some of its details are incorrect. But I think that something very much like it must be correct if there is a tenable Millian account of how to have *de re* beliefs by communication-chains.

We start off with an individual who stands in a direct perceptual acquaintance relation to a particular concrete object and thereby comes to have a *de re* belief about it—like Ralph's belief about Bessie in our canonical case. Ralph has a perceptual *de re* mode of presentation of Bessie. How does Rhoda, who never perceives Bessie, come to have her *de re* belief that Bessie is starving? Ralph uses a name in communicating his thoughts about Bessie. When he thinks of Bessie, he does so via mentally tokening the name 'Bessie'. When he communicates to Rhoda about Bessie, he does so by uttering a physical token of the name. Rhoda in turn hears the name token and can consequently think thoughts about Bessie by herself mentally tokening the name. The pivotal idea as to why she is able to have *de re* attitudes toward Bessie is this: mental tokens of names are vehicles for *de re* thought and physical tokens of names are vehicles for transferring *de re* thought.

Notice that Rhoda's (1) *de re* belief is specifically about Bessie because she mentally tokens a name that she received from a communication-chain whose starting node is a reference-fixing of 'Bessie' to Bessie. To be sure, in this instance, 'Bessie' had its reference fixed ostensively. And, to be sure, Ralph had a *de re* mode of presentation of Bessie. But these facts about the ostension itself and Ralph's *de re* mode of presentation contributes nothing essential to the explanation why Rhoda's thought is *de re*. It would contribute something essential only if Ralph somehow passed along his *de re* perceptual mode of presentation to her. But, of course, he does not, and cannot. Perceptual modes of presentation are not transferable. Ralph cannot pass on to Rhoda his *de re* perceptual mode of presentation of Bessie by communicating about it. One can describe the content of one's perception of a cow, but one cannot pass it along. What the recipient receives is a description of a subjective perception, not the perception itself. Ralph also cannot pass on to Rhoda his perceptual mode of presentation of Bessie by using a name for Bessie. Names do not function as vehicles for passing along perceptual *de re* modes of presentation. Lacking semantic descriptive content and representational content, names cannot carry the information contained in perceptual *de re* modes of presentation. Rhoda has no perceptual representation of Bessie. What makes her thought *de re* and about Bessie in particular is simply that she thinks thoughts by mentally tokening the name 'Bessie.'

The upshot is that the account of communication-based *de re* belief makes essential use of only those facts that are needed for an individual to have de-

scriptive reference-fixing generated acquaintanceless *de re* belief. The individual needs to mentally token a name. What allows the individual to have *de re* thoughts about the named object is that mental tokens of the name serve in cognition as *de re* modes of presentation of the named object. So if one countenances the possibility of *de re* belief through communication chains (at least on a model that looks something like Bach's), then one ought to countenance the possibility of descriptive reference-fixing generated *de re* belief.

OBJECTIONS AND REPLIES

I will consider five objections to the thesis and arguments for it.

Information Objection

In order to have a *de re* belief about an object with which one is unacquainted, one must have a substantial amount of information about it. If one has only an extremely limited amount of information about the object, one will necessarily think of that object "satisfactionally," as whatever satisfies the conjunction of the descriptions that express one's set of information. Thought about that object cannot be structurally similar to canonical instances of *de re* belief.

While I see the force of this point, I disagree. True, when we reflect on canonical instances of *de re* belief, we think, correctly, that our perception of the object (or our memory of the perception) provides us with an exceptionally rich stock of information about the object. This is so even if we have but a fleeting glance at the object. The richness stems from the "pictorial" or "photographic" quality of our perceptual representation. Acquaintanceless *de re* belief generated from descriptive reference-fixing will be, by comparison, informationally emaciated. But, it is far from clear that what is essential to *de re* thought is the richness of information about the object. For one, not all of our *de re* belief is informationally rich in the way in which our canonical cases are rich. Cases involving acquaintance via communication-chains militate against this idea. If I hear you say, "My dog Elmer is the sweetest dog," and I have never met Elmer, all that I know about Elmer is that he's a dog and that you think he's the sweetest. My causal relation to Elmer does not help enrich my stock of information. Neither does the fact that you have a lot of information about him. My psychological state is not altered simply by the *existence* of yours plus my causal relation to you. If a host of information is demanded, much communication-based belief cannot be *de re*. Of course, the proponent of this objection may think that I cannot have a *de re* belief in the Elmer scenario.³⁷ But the example is not needed to make the point. The idea I have suggested is that acquaintanceless cases can be structurally similar to the canonical cases in virtue of playing the same role in cognition—roughly speaking, by the fact that the subject has a mental file folder for the object. The fact that the information on the object is extremely limited does not count against the possibility of opening a new

mental file by mentally tokening a name and so does not entail that the subject must think of the object “*satisfactionally*.” My point is that another mechanism makes the thought *de re*—mentally tokening a name—and its capacity to do so does not depend upon the extent of the subject’s information about the object.

Causal Connection Objection

One of the primary cases discussed—the Neptune case—does not seem to be an acquaintanceless case at all. After all, Leverrier has some causal contact with Neptune, and this is enough to secure a weak notion of acquaintance, and so enough for Leverrier to have a *de re* belief prior to his act of naming. One needs to establish a point for cases in which no causal relation obtains between subject and object of belief.³⁸

My reply is that there is nothing about causal contact *per se* that suffices for acquaintance or *de re* belief. It is not the case that all causal connections give rise to the perceptual representation that are supposed to capture “aboutness.” Although our canonical cases do, Leverrier did not obtain any perceptual representation or nondescriptive mode of presentation of Neptune as a result of his causal relation to Neptune. Furthermore, the fact that he recognized that there exists a causal relation between himself and the planet causing the perturbation in the orbits of Uranus also does not itself suffice. All that that offers is some causal description, which would itself do nothing more than provide a canonical *de dicto* way of thinking of the object. I do not doubt that Leverrier might already have had a *de re* belief about Neptune prior to his linguistic act. (See Irrelevance of Naming Objection on page 71). But, if so, his belief is obtained from his having consolidated information about the planet, i.e., he opened a new mental file folder for it. Yet his capacity to do this is not essentially tied to any acquaintance or causal contact with Neptune.

Basicness of Acquaintance Objection

Surely acquaintance is necessary for characterizing *de re* belief, and for marking out the *de re/de dicto* distinction. It is, after all, from the notion of acquaintance that we have the idea of a belief being directly about (relationally about) an object as opposed to being about it only indirectly (*satisfactionally*).

Although I have argued that acquaintance is not necessary for *de re* belief, I have not argued that acquaintance is not in some way significant to an understanding of *de re* belief. *De re* beliefs via acquaintance are developmentally primary. Also, I would hypothesize that acquaintanceless *de re* belief is impossible without *de re* belief with acquaintance. And, no doubt, it is (direct) acquaintance that suggests the idea of a belief being directly about an object. But acquaintance itself is sometimes not needed for *de re* belief. We can use proper names to generate beliefs that have the same structure and role in cognition as the canonical instances of *de re* belief. Though such beliefs are not accompanied

by perceptual representations, they are not thereby only indirectly about the objects of thought. The possibility of *de re* belief without acquaintance is parasitic on *de re* belief with acquaintance. Our analysis frees *de re* belief from the necessity of acquaintance while making sense of the fact that so many others have deemed it necessary.

Irrelevance of Naming Objection

Assuming one is right about the conditions on introducing a name into the language, and assuming one maintains that one can have *de re* belief about objects one has no name for, then there is no *de re*-belief generating role for descriptive reference-fixing to play. If one is in position to introduce a name for an object O, one will already have a *de re* belief about O, and hence, the act of naming will be otiose vis-à-vis generating *de re* belief. Alternatively put, descriptive reference-fixing is merely a symptom, not a cause, of the stipulator's *de re* belief.³⁹

Let me start off by addressing this important objection by stating that, yes, of course, I think that one can have a *de re* belief about objects for which one lacks a name. Also, I agree that there are many cases in which a stipulator introduces a name for an object he or she already has a *de re* belief about. So there *will* be cases in which descriptive reference-fixing will be a symptom, not a cause, of a *de re* belief. The introduction of 'Jack the Ripper' is probably a case of this type. Detectives that had reason to think that a single individual was responsible for a particular series of murders no doubt had *de re* beliefs about this individual prior to coining the name. This is, of course, consistent with the position I advocate. I maintain that mentally tokening a name for an individual is a sufficient, not a necessary, condition on having a *de re* thought.

Now, I will, in the end, attempt to turn back this objection. But I initially wish to assume it is right and explore the consequences. If this objection is cogent, then my discussion of descriptive reference-fixing as generating *de re* belief is wrong. However, since the objection assumes as correct my account of the conditions on introducing names into the language, it presupposes that agents who do so via descriptive reference-fixing already have *de re* beliefs about the object satisfying the description. Thus, if names are in fact introduced into the language or idiolect in cases in which the stipulator lacks an acquaintance relation with the object (e.g., 'Neptune'), we still need an explanation of how those acquaintanceless *de re* beliefs came about. This simply points us in the direction of further research.⁴⁰ I do not pretend to have exhausted or even explored all avenues for generating acquaintanceless *de re* beliefs. So if the objection is right, it does not establish that acquaintanceless *de re* belief is impossible. Indeed, it seems to put pressure on us to show how it is possible.

The objection is a concern that all cases of descriptive reference-fixing are, necessarily, merely indicative of the prior presence of a *de re* belief. The linguistic act contributes nothing. If this is so, my argument that the descriptive refer-

ence-fixing act can help generate *de re* belief is unsound. The line of thinking bolstering the objection runs as follows: if one could have different intentions—say, *de re* intentions about an object one does not intend to name—and these intentions are also enough to set up a mental file folder, then naming the object cannot play any role in generating *de re* belief. But surely this does not follow. The intentions involving names characterized in *Sincerity (Description)* and *Psychological Neutrality (Description)* can contribute to generating the *de re* belief. This invalid inference is tempting to those who think that satisfaction of *Sincerity (Description)* and *Psychological Neutrality (Description)* is one thing, and descriptive reference-fixing another. This is a mistake. The fact that an agent satisfies *Sincerity (Description)* and *Psychological Neutrality (Description)* is not normally something that is independent of the generation of *de re* belief via the descriptive reference-fixing act. By having the intentions in *Sincerity (Description)* and *Psychological Neutrality (Description)*, the agent has in effect done everything she needs to do to fix the reference of a name, short of actually using the name. It is not as if, in addition, she needs to say, “Let ‘N’ refer to the F,” and this *saying* is the reference-fixing. In fact, the satisfaction of *Sincerity (Description)* and *Psychological Neutrality (Description)* effectively sets up a mental file folder with the name as label.

But perhaps this objection stems only from a conviction that reference-fixing is, necessarily, psychologically insignificant. If so, I find it less than compelling. I described a case in another paper⁴¹ in which a contest is being held: the child with the best project for celebrating the millennium will be the winner. A girl decides it would be neat to discover, and then honor, the first person born in the twentieth century. To get her project underway, she engages in the following act. She fixes the reference of the term ‘Oldman 1’ as first person born in the twentieth century,⁴² and then sets out to identify this individual. She thinks about what she will say to Oldman 1 when first introduced, how she will honor Oldman 1, and so on. She initiates thought and plans involving this individual, actions that do not get underway at all until she names the individual. Of course, I cannot prove the point that the reference-fixing itself contributes to generating the *de re* belief, but I think that the case, as described, is possible, and now leave it as incumbent on the opposition to establish otherwise.

Existence Objection

Without an acquaintance condition, nothing precludes the possibility of having a *de re* belief about a nonexistent object. But there can be no *de re* belief if the object of belief does not exist. So we cannot give up acquaintance as a necessary condition for *de re* belief.

The issue about the (im)possibility of *de re* belief and its relationship to empty names needs extensive, separate treatment, but I will comment briefly on the supposed *reductio*. The second premise may be doubted. To be sure, it is pe-

cular to speak about having a *de re* belief about nothing. Nevertheless, I am inclined to think that we can still preserve the fundamental intuitive understanding of *de re* belief as belief about something in particular, while admitting the possibility of *de re* belief about nothing (paradoxical as it may sound!). The central idea is a variant on the one employed here: *de re* beliefs about nonexistent objects are initiated by a subject's opening a mental file; this essential structural feature is what makes such beliefs *de re*. In this view, *de re* beliefs about nonexistent objects are parasitic on the more basic *de re* belief about existent objects. This enables us to simultaneously deny and explain the common assumption that *de re* belief requires for its content an existent object.⁴³

NOTES

I first presented this paper in a colloquium at the University of Michigan in February 2000. A month later I presented a somewhat altered version at the Inland Northwest Philosophy Conference on Meaning and Truth, with Leora Weitzman commenting. Thanks to Michael O'Rourke for inviting me to that terrific conference, and for his very helpful comments. I also wish to thank Kent Bach, David Chalmers, Thomas Hofweber, David Hunter, Jim Pryor, Marga Reimer, Nathan Salmon, Jonathan Sutton, Ken Taylor, David Velleman and Gideon Yaffe for stimulating conversations and correspondence about these ideas. Two people really helped push this research forward—Michael Nelson, for his heels-in-the-ground resistance to the central thesis; and Leora Weitzman, for her thoughtful ideas on how to make the thesis work. I am grateful to them both.

1. To have a *de re* thought, it is not, I think, necessary for a singular proposition to be the content of the thought. I believe that there may be exceptions in instances in which an agent thinks with an empty name. If there is no object that is the object of the *de re* thought, then the agent cannot have a singular proposition as the thought content. (I am not sure whether "gappy" singular propositions will do.) I will not deal with this problem in this paper, but do so in "Farewell Acquaintance." There I attempt to argue that we ought to abandon not only acquaintance, but also the traditional existence condition on *de re* thought.
2. Bach (1987, 27–31) explicitly considers cases of this type and offers a theory as to why such beliefs are *de re*. This position might be embraced by Boer and Lycan (1986), Burge (1977), and Perry (1980), as it is consistent with the principles on *de re* belief they advocate, but they do not spell out an account of why such a belief would be *de re*. Evans (1982) and Recanati (1993) would probably object on the grounds that the subject can no longer individuate the object.
3. Bach (1987, 31–39) and Boer and Lycan (1986, 128–30) examine cases of this kind and claim they are *de re*.
4. See Sosa 1970 for contemporary expression of the view that all belief is ultimately *de dicto*. The view has roots in Frege 1952.
5. Although Lewis (1979) reduces all *de re* belief to *de se*, he does uphold a *de re/de dicto* distinction. He maintains that acquaintance relations ground and distinguish, and hence are necessary for, *de re* beliefs about concrete objects. Boer and Lycan (1986) regard *de re* belief as a special case of *de dicto* belief, yet uphold a distinction that is rooted in grades of acquaintance relations.
6. Kaplan (1989a) abandons an acquaintance relation.
7. Boer and Lycan (1986, 133) present a highly developed theory of knowledge-who. They also present convincing cases for thinking one could have a *de re* belief without knowing who.
8. Consider, for example, Bach's influential (1987, 12) where he characterizes *de re* belief as, essentially, belief whose object is determined relationally. For him, this comes down to some variety of acquaintance. For Bach, *de dicto* (or descriptive) belief is belief whose object is determined satisfactorially. While I think that Bach's characterization is, in the

end, wrong, I think it is a clear and useful heuristic for getting at the distinction. Recanati also embraces this basic distinction: “a *de re* mode of presentation involves a certain *relation to the reference*” (1993, 103). Burge (1977, 346) holds a similar thesis though his characterization is problematic: “A *de re* belief is a belief whose correct ascription places the believer in an appropriate nonconceptual, contextual relation to objects the belief is about.” This analysis unfortunately glosses together *de re* belief with *de re* belief ascriptions. Also, appeal to a contextual relation makes it hard to fathom how to apply it to *de re* mathematical beliefs.

9. Charles Parsons (1979, 1993) has developed the view that many of our beliefs about numbers are *de re*, and are attained by a kind of Kantian intuition. I find many aspects of his account quite persuasive.
10. There is a pervasive tendency in the literature to restrict discussion to *de re* belief about concrete objects, laying aside or ignoring belief about nonconcrete objects, cf. Bach 1987, 11; Burge 1977, 339; Recanati 1993, 116. I think this failure to confront *de re* belief about nonconcrete entities contributes to the prevalence of an acquaintance condition.
11. The trend in the literature has been to deny the possibility of *de re* belief of nonexistent objects, while attempting to explain away our intuitions about Vulcan-like cases. Some theorists maintain that Vulcan is, in fact, existent; it is a mythical object that exists as an abstract entity. Our astronomers may have *de re* belief about Vulcan, but the possibility is secured at the expense of a pleasing metaphysics. Salmon (1998, 2000) favors this approach. The main problem as I see it is not with the abstract entities *per se*, but rather with the fact that mind-dependent abstract entities are brought into existence without any individual’s intention to do so. An alternative route is to deny that our astronomers have any *de re* belief. Without an existent object, the subject has no thought, or has, at best, a *de dicto* belief. Cf. McDowell 1984, 1986 for the no-thought approach, Evans 1982, Burge 1977, Alston 2000 for the *de dicto* approach. The problem with these lines is that they regard our intuition about psychological states as trumped by a theory of the *de re*, without offering a convincing explanation why this should be so. It is at least intuitively implausible that astronomers of the nineteenth century had no thought at all, or just *de dicto* thought, corresponding to their acceptance of (7).
12. In “Farewell Acquaintance,” I attempt to offer a solution to the problem of belief about nonexistent object. I suggest that such beliefs can be *de re*, and that the metaphysical standing of such objects (whether one takes them to be nonexistent, to be abstract entities, or to have some kind of Meinongian standing) is of little significance to this issue.
13. Ideally, this chapter should be read as a companion to Jeshion 2000, 2001a. In “Ways of Taking a Meter,” I focus on the epistemological issue of how the stipulator attains *a priori* justification for her belief. In “Donnellan on Neptune,” I criticize Donnellan’s argument for the metalinguistic position about what the stipulator knows, and I attempt to set up a dilemma for one who adheres to the metalinguistic view.
14. Cf. Salmon 1989, Soames 1995.
15. The *locus classicus* is Donnellan 1979. Cf., also Salmon 1987.
16. The full argument against Donnellan’s argument is in §1 and §2 of Jeshion 2001a. Here’s a thumbnail sketch. Donnellan attempted to establish that no *de re* belief was attained by Leverrier, but do so without advancing necessary and sufficient conditions for having a *de re* belief. Instead he advanced a loose principle governing *de re* thought. Call it the *De Re Principle*: If one has a name ‘N’ for an individual, and there is a belief that one would express by saying “N is the F”, then if one subsequently meets the individual it will be true to say to that individual “I believed that you are the F” (or, if the individual is not a person, pointing, “I believed that that is the F”). He claimed that the Leverrier case appears to strain this principle. At an intuitive level, it seems that Leverrier would speak falsely if he said, pointing to Neptune, “I believed that that is the cause of the orbital perturbations.” Donnellan claimed that in the absence of an alternative explanation, this is a sign that the subject never had a *de re* belief in the first place. He claimed that there was no alternative. I argued that there is an alternative explanation. The strain on the *De Re Principle* may be

- due solely to the Frege's Puzzle structure of the case, i.e., to the fact that Leverrier has two different ways of taking Neptune, one involving acquaintance, the other generated from stipulative descriptive reference-fixing.
17. In Jeshion 2001a, I discuss reasons for thinking that the Millian should not give up the theses of Possibility of Stipulative Descriptive Reference-Fixing, Purist Millianism, Accessibility of Content, and Understanding Millian Names. Salmon (1987) suggests that Understanding Millian Names is false. Soames (1995, *fn* 5) seems to suggest that Accessibility of Content must be qualified. Kim (1977) denies the Possibility of Stipulative Descriptive Reference-Fixing, as does Russell (1956).
 18. Kripke 1980, 14, 56, my emphasis. It is not clear to me whether Kripke intends for the linguistic act to be capable of generating both a *de re* belief and its justification or just the justification of a prior *de re* belief. My view here is like his only if he intended the former. Kaplan (1989a, especially 536, 560) advances a similar line, though Kaplan seems to reject the idea that there are strict conditions on descriptive reference-fixing.
 19. Evans 1985, 181.
 20. That is, he held that 'Julius' is synonymous with the definite description with which its reference was fixed. To secure the contingency of the sentence, "N is the F, if anything is", the definite description must be the rigidified description 'the actual inventor of the zip'. A Millian who adopts this position would be rejecting our thesis of Purist Millianism.
 21. Evans 1985, 202; Grice 1969; Recanati 1993, 108; Sutton 2000.
 22. In addition to Evans, Plantinga (1969), Kaplan (1989a), Salmon (1987, 1998), and Donnellan (1979) all seem to think that one can always introduce a name into the language *simply* by uttering "Let 'N' denote the F". Like Evans's Julius, Kaplan's Newman1, Salmon's Nappy and Curly-o examples fail to satisfy our conditions and hence no such names are introduced into the language.
 23. Austin 1962. On naming as a performative, see also Searle 1969 and Bach and Harnish 1979.
 24. There are some related ideas on the utility of names in Strawson 1974, 42–48. Thanks to Gil Harman for drawing my attention to Strawson's discussion.
 25. The notion of sincerity conditions goes back to Austin. The idea that names have an essential "psychological neutrality" runs through the writings of many Millians and some contemporary neo-Fregeans. I borrow the term from Recanati 1993, though I must not be understood as adopting his view on the meaning of or conditions for thinking with names.
 26. Searle 1969, 61. By contrast, Alston (2000) classifies names and greetings in different categories, as what he calls exercitives, and expressives, respectively.
 27. Reflection on exceptions actually helps bolster *Single Tagging*. Many pet names (e.g., 'Emmie' for 'Emily') are, not insignificantly, phonologically similar to the nonpet name. This mitigates the extent to which they count against the role of names as devices for interpersonal communication and psychologically neutral thought about individuals. Pen names are normally introduced to conceal the identity of the writer; their capacity to do so depends essentially upon the presence of a normative principle like *Single Tagging*. Notice that there is something not just semantically nonstandard, but semantically deviant about multiple aliases.
 28. Acquaintanceless cases (Neptune, π) are typical instances in which we desire psychological neutrality, yet ostension is unavailable. Even if acquaintance is present, sometimes descriptive reference-fixing is needed. Take the meter stick scenario in which the stipulator is standing before a particular stick and so is acquainted with its length. Ostension will not suffice. If one says: "*that* is one meter," pointing to the stick, one will ordinarily be taken to be naming the stick, not its length. Description here is needed to supplement ostension.
 29. If you are not already convinced that there are—and must be—conditions on naming, read this remarkable passage by Plantinga (1969, 253): "Is there really any reason why I can't name all the real numbers, or, indeed, everything whatsoever in one vast, all-embracing baptism ceremony? I can't see any such reason, and I hereby name everything 'Charley'."
 30. This is not to deny another sort of 'freedom': any object is, arguably, a candidate for a

- name, for any (or just about any) object is a possible object of our interest. Hence, provided we have the right intentions, intentions that are responsive to rules on naming, we are free to name any object.
31. So I reject a wholesale voluntarism about belief; we do not have that sort of control on our cognitive lives. I simply assume this point.
 32. In saying that the stipulator knows that the reference of ‘N’ is determined by reference-, not meaning-fixing, I do not presuppose any capacity to articulate or perhaps even fully conceptualize the distinction. I only assume that such a stipulator has a well-grounded awareness of the difference between giving a name a meaning and tagging a name with a reference, an awareness that will be possessed by rational agents that engage in the sophisticated practice of reference-fixing via definite description.
 33. The skeleton of this idea is, I think, found in Harman (1977), who claims that for an individual to have a thought about object O, it is enough if that individual uses a “mental name” as a name for that individual, and such mental names can be introduced descriptively. Though he does not claim that such thought will be *de re*, much in his discussion suggests that it should be.
 34. The notion of a mental file folder or a mental dossier is popular in some of the current literature on *de re* thought. For a nice development of the idea, cf. especially Perry (1980). It is not clear to me whether Perry would be sympathetic to the line being pushed here. Recanati (1993) uses a similar notion of a mental file folder but claims that *de re* thought is impossible without non-descriptive identifying information. The notion of the mental dossier goes back to Grice (1969).
 35. I am inclined to think that what is essential to *de re* thought, and what distinguishes it from *de dicto* thought, is something like the presence of a mental file folder that classifies information about the object. Psychological neutrality is probably a concomitant of mental dossiers. The relation between *de re* thought, names, mental files, and psychological neutrality deserve much more attention than I can give them here, and the remarks in this paragraph are intended only to give a sketch of my view; they are not intended as an argument. Thanks to Leora Weitzman for her probing questions about these matters.
 36. Bach 1987, 31–39.
 37. For example, anyone who buys an individuation requirement (like Evans and Recanati) will think that I have only a *de dicto* belief about Elmer.
 38. Nathan Salmon advanced this point (in conversation).
 39. Kent Bach and Leora Weitzman pressed this point.
 40. Maybe the act of introducing a name is separable from introducing a mental file folder, and the latter is primary. I am dubious about separating the two, but this is an issue that requires much more attention than I can give it here.
 41. The Oldman 1 case is in Jeshion 2001a. In that case, the stipulator is in fact acquainted with the object satisfying the description used to fix the reference of the introduced name, yet she is not aware that this is so. She thinks she is unacquainted with him. But these matters do not concern us here. The intuitive pull of the case for our purposes does not depend on whether the stipulator is acquainted (under a different guise) with the first-born of the twentieth century.
 42. It is true that this act violates Single Tagging. But I think that the violation, in this instance, is one of those exceptions that makes the rule. Our subject introduces the name so that she has a psychologically neutral way of thinking of the individual while engaged in her project to identify him.
 43. I would conjecture that the standard existence assumption stems not only from the intuitive characterization of *de re* belief, but also from the substitutivity properties (existential generalization in particular) ascribed to *de re* readings of belief sentences, cf., Quine 1956. In short, there is a carryover of a key assumption from the tradition that understood the *de re/de dicto* distinction by exploring the semantics of belief sentences. We need to more fully liberate our understanding of *de re* belief from this Quinean tradition.

REFERENCES

- Alston, W. 2000. *Illocutionary Acts and Sentence Meaning*. Ithaca, N.Y.: Cornell University Press.
- Austin, J. L. 1962. *How To Do Things with Words*. Cambridge, Mass.: Harvard University Press.
- Bach, K. 1987. *Thought and Reference*. Oxford: Clarendon Press.
- Bach, K. and Harnish, M. 1979. *Linguistic Communication and Speech Acts*. Cambridge, Mass.: MIT Press.
- Boer, S and Lycan, W. 1986. *Knowing Who*. Cambridge, Mass.: MIT Press.
- Burge, T. 1977. "Belief *De Re*." *Journal of Philosophy* 74: 338–62.
- Davies, M. 1981. *Meaning, Quantification, Necessity*. London: Routledge and Kegan Paul.
- Donnellan, K. 1979. "The Contingent *A Priori* and Rigid Designators." In P. French, T. Uehling, and H. Wettstein (eds.), *Contemporary Perspectives in the Philosophy of Language*. Minneapolis: University of Minnesota Press.
- Dretske, F. 1988. *Explaining Behavior: Reasons in a World of Causes*. Cambridge, Mass.: MIT Press.
- Evans, G. 1982. *The Varieties of Reference*. Oxford: Clarendon Press.
- . 1985. *Collected Papers*. Oxford: Clarendon Press.
- Frege, G. 1952. "On Sense and Reference." In P. Geach and M. Black (eds.), *Translations from the Philosophical Writings of Gottlob Frege*. Oxford: Basil Blackwell.
- Grice, P. 1969. "Vacuous Names." In D. Davidson and J. Hintikka (eds.), *Words and Objections*. Dordrecht: Reidel.
- Harman, G. 1977. "How to Use Propositions." *American Philosophical Quarterly* 14: 173–76.
- Jeshion, R. 2000. "Ways of Taking a Meter." *Philosophical Studies* 99: 297–318.
- . 2001a. "Donnellan on Neptune." *Philosophy and Phenomenological Research*, forthcoming.
- . 2001b. "Farewell Acquaintance." Ms.
- Kaplan, D. 1969. "Quantifying In." In D. Davidson and J. Hintikka (eds.), *Words and Objections*. Dordrecht: Reidel.
- . 1989a. "Demonstratives." In J. Almog, H. Wettstein, and J. Perry (eds.), *Themes from Kaplan*. New York: Oxford University Press.
- . 1989b. "Afterthoughts." In J. Almog, H. Wettstein, and J. Perry (eds.), *Themes from Kaplan*. New York: Oxford University Press.
- Kim, J. 1977. "Perception and Reference Without Causality." *Journal of Philosophy* 74: 606–20.
- Kneale, W. 1962. "Modality, *De Dicto* and *De Re*." In E. Nagel, P. Suppes, and A. Tarski (eds.), *Logic, Methodology and Philosophy of Science*. Stanford, Calif.: Stanford University Press.
- Kripke, S. 1980. *Naming and Necessity*. Oxford: Basil Blackwell.
- Lewis, D. 1979. "Attitudes *De Dicto* and *De Se*." *Philosophical Review* 88: 513–43.
- . 1983. "Individuation by Acquaintance and by Stipulation." *Philosophical Review* 92: 3–32.
- McDowell, J. 1984. "*De Re* Senses." In C. Wright (ed.), *Frege: Tradition & Influence*. Oxford: Basil Blackwell.
- . 1986. "Singular Thought and the Extent of Inner Space." In P. Petit and J. McDowell (eds.), *Subject, Thought and Context*. Oxford: Clarendon Press.
- Parsons, C. 1979. "Mathematical Intuition." *Proceedings of the Aristotelian Society* 80: 145–68.
- . 1993. "On Some Difficulties Concerning Intuition and Intuitive Knowledge." *Mind* 102: 233–46.
- Perry, J. 1980. "A Problem About Continued Belief." *Pacific Philosophical Quarterly* 62: 317–22.
- Plantinga, A. 1969. "*De Re et De Dicto*." *Nous* 3: 235–58.
- Quine, W. V. O. 1956. "Quantifiers and Propositional Attitudes." *Journal of Philosophy* 53: 177–87.
- Recanati, F. 1993. *Direct Reference: From Language to Thought*. Oxford: Basil Blackwell.
- Russell, B. 1918. "Knowledge by Acquaintance and Knowledge by Description." In his *Mysticism and Logic and Other Essays*. London: Longmans, Green and Co.
- . 1956. In R.C. Marsh (ed.), *Logic and Knowledge: Essays 1901–1950*. London: George Allen & Unwin.
- Salmon, N. 1987. "How to Measure the Standard Meter." *Proceedings of the Aristotelian Society* 88: 193–217.

- . 1989. "How to Become a Millian Heir." *Nous* 23: 211–20.
- . 1998. "Nonexistence." *Nous* 32: 277–319.
- . 2002. "Mythical Objects." In J. Campbell, M. O'Rourke, and D. Shier (eds.), *Meaning and Truth: Investigations in Philosophical Semantics*. New York: Seven Bridges Press.
- Schiffer, S. 1977. "Naming and Knowing." *Midwest Studies in Philosophy* 2: 28–41.
- Searle, J. 1969. *Speech Acts*. Cambridge: Cambridge University Press.
- Soames, S. 1995. "Beyond Singular Propositions." *Canadian Journal of Philosophy* 25: 515–50.
- . 1998. "The Modal Argument: Wide Scope and Rigidified Descriptions." *Nous* 32: 1–22.
- Sosa, E. 1970. "Propositional Attitudes *De Dicto* and *De Re*." *Journal of Philosophy* 67: 883–96.
- Strawson, P. 1974. *Subject and Predicate in Logic and Grammar*. London: Methuen & Co Ltd.
- Sutton, J. 2000. "The Contingent *A Priori* and Implicit Knowledge." *Philosophy and Phenomenological Research*, forthcoming.

CHAPTER 5

Meanings

Stephen Schiffer
New York University

DO THERE EXIST such entities as *the things we mean*? If such things do exist, they are also the things we believe and assert, and thereby the things in terms of which we must understand what our words and sentences mean. This question about the existence of things we mean and believe is the first question in the theory of linguistic and mental content, for its answer sets the agenda for the theory of content. If, for example, the answer is yes, then the next question is: What is the nature of those things that are the things we mean and believe?

In this chapter I assume, as a working hypothesis, that there are such things as the things we mean and believe; my goal is an account of their nature. The things we mean and believe are typically ascribed as contents via that-clauses, such as ‘that Australia is next to Germany’ in ‘Harold believes that Australia is next to Germany’, and my focus will be on belief reports of the form

(1) *A* believes that *S*,

since this is the most familiar exemplar in discussions of content, and what applies to them applies, *mutatis mutandis*, to every other that-clause-containing ascription of content.

To make things even easier for myself, I also assume, as a second working hypothesis, (a) that ‘believes’ in (1) stands for a two-place relation—the belief relation—that holds between believers and the things they believe,¹ and (b) that the that-clause that results from a substituent for ‘*S*’ refers to that which *A* is being said to believe.²

BELIEF CONTENTS AS PROPOSITIONS

Consider

(2) Ralph believes that lobsters are fish.

If, as I am assuming, there exist such things as things believed, then one of those things is the reference of the that-clause in (2), ‘that lobsters are fish’, and our question is about the nature of that thing. It happens we can say a fair amount about it right off the bat. ‘That lobsters are fish’ plainly refers to *that lobsters are fish*, and the following seems straightforwardly true of this thing, *that lobsters are fish*, which is the reference of the that-clause:

- *That lobsters are fish* is *abstract*: it has no spatial location or any other physical properties.
- It is *mind- and language-independent* in two senses. First, its existence is independent of the existence of thinkers or speakers. *That lobsters are fish*, which can be believed by different people speaking different languages, didn’t spring into existence the first time, or each time, someone believed or asserted it; it wasn’t brought into existence by anything anyone said or thought. Second, *that lobsters are fish* can be expressed by a sentence of just about any natural language but itself belongs to no language.
- It has a *truth condition*: it’s true iff lobsters are fish.
- It has its truth condition *essentially*; it is a *necessary truth* that *that lobsters are fish* is true iff lobsters are fish. The contrast here is with sentences. The *sentence* ‘Lobsters are fish’ is also true iff lobsters are fish, but that is a *contingent* truth that would have been otherwise had English speakers used ‘lobster’ the way they now use ‘flounder’.
- It has its truth condition *absolutely*, i.e., without relativization to anything. The contrast is again with sentences. The sentence ‘Lobsters are fish’ has its truth condition only *in English* or *among us*. There might be another language or population of speakers in which it means that camels snore; but *that lobsters are fish* has its truth condition everywhere and everywhen.

From all this we may conclude, by an obvious generalization, that things believed are what philosophers call *propositions*: abstract, mind- and language-independent entities that have truth conditions, and have their truth conditions both essentially and absolutely.

SOME DIVIDING ISSUES

This still leaves plenty of room for philosophers to disagree about the further nature of the propositions we believe. The first big issue is whether these propositions are *structured* or *unstructured*. ‘Structured’ here is a term of art whose use may be explained in the following way.

Propositions are *structured* if (nearly enough for now) they stand in a one-to-one correspondence with ordered *n*-tuples whose basic components are things

that are not themselves propositions, otherwise *unstructured*. Structured propositions are said to be composed of the constituents of the n -tuples to which they stand in one-to-one correspondence, and the basic components of those constituents are called “propositional building blocks,” the entities from which propositions are built up via certain recursive rules. Parties to the debate about whether that-clauses refer to structured or to unstructured propositions generally share a certain presupposition that frames the terms of their debate. This shared presupposition, which I will call the *Compositionality Hypothesis* (CH) and which is very widely held, states that the reference of a that-clause is a function of the references of its primary expressions.³ If CH is correct, it is both possible and useful to construe the arguments of the proposition-yielding function to be ordered pairs of the form

$$(3) \langle \langle x_1, \dots, x_n \rangle, X^n \rangle$$

where the values of the variables are the references the primary expressions in that-clauses have in those that-clauses. The theorist of structured propositions claims that the function is bijective, i.e., that it determines a one-to-one correspondence, so that no two instances of (3) can determine the same proposition. Consequently, for the structuralist, two that-clauses can refer to the same propositions only if their primary expressions have the same references. The theorist of unstructured propositions who accepts CH agrees that the propositions to which that-clauses refer are a function of instances of (3), but he denies that this function yields a one-to-one correspondence, and thus holds that two that-clauses may refer to the same proposition even though their corresponding primary expressions do not share the same references.

There is reason to doubt that both CH is true and that-clauses refer to unstructured propositions. For it would seem that the only way of individuating unstructured propositions given CH is via the possible worlds in which they are true, which entails that there is only one necessarily true proposition and that, consequently, Livonia knows every mathematical truth just by virtue of knowing that every dog is a dog. Robert Stalnaker, an ingenious defender of unstructured propositions, has gone to some lengths to explain away the counterintuitiveness of such consequences (see, e.g., Stalnaker 1984). There is some question whether he succeeds. It must be emphasized, however, that Stalnaker’s counterintuitive result requires holding *both* that that-clauses refer to unstructured propositions *and* that CH is true. There are two ways to endorse unstructured propositions: relative to CH, and relative to the denial of CH. It is only the theorist of unstructured propositions *who accepts CH* [and must therefore hold that the propositions to which that-clauses refer are a function of ordered pairs of form [3)] who must individuate propositions by their possible-worlds truth conditions. A theorist who denies that the reference of a that-clause

is a function of the references of its primary expressions can individuate propositions as finely as anyone, if not more finely. There will be more on this later.

Most propositionalists accept CH and reject unstructured propositions. Among theorists of structured propositions the big debate is about the references of expressions in that-clauses, and therewith about the building blocks of structured propositions. It is useful to begin with a distinction that yields an exclusive but not exhaustive partitioning of logical space, a distinction between what I will call the *strong Russellian* position and the *strong Fregean* position.⁴ Both positions presuppose CH.

The strong Russellian holds that the references expressions have in that-clauses are the objects and properties our beliefs are about, and that therefore the references of that-clauses are always *Russellian propositions*—propositions whose basic constituents are the objects and properties our beliefs are about. According to this theorist, the references of the expressions in the that-clause in

- (4) Ralph believes that Bill loves Chelsea.

determine the ordered pair

- (5) $\langle\langle\text{Bill, Chelsea}\rangle, \text{the love relation}\rangle$,

which in turn determines, by a one-to-one correspondence, the Russellian proposition to which the that-clause refers. The basic components of the Russellian proposition are the same as the ones in the ordered pair that determines the proposition. In fact, Russellians typically *identify* their propositions with ordered pairs like (5), although a more cautious theorist may want to say that although structured propositions stand in one-to-one correspondence with ordered pairs, the propositions themselves are *sui generis* abstract entities not identifiable with any set-theoretic constructions. Still, there can be no harm in *representing* Russellian propositions by the ordered pairs that determine them.⁵ The Russellian proposition (5) represents will be true in a possible world w iff $\langle\text{Bill, Chelsea}\rangle$ instantiates the love relation in w , false in w otherwise.

The strong Fregean holds that the references expressions have in that-clauses are never the objects and properties our beliefs are about but are rather things she calls *concepts*, or *modes of presentation* (*guises, ways of thinking*), of the objects and properties our beliefs purport to be about. (Henceforth, I drop ‘mode of presentation’, which was Frege’s own metaphor, and use ‘concept’, even though this use of ‘concept’ differs from Frege’s technical use.) Fregean propositions, then, are structured entities whose basic building blocks are concepts of the objects and properties our beliefs purport to be about. This theorist will want to say that there are concepts c_b , c_c , and C_L of Bill, Chelsea, and the love relation, respectively, such that in (4) ‘Bill’ refers to c_b , ‘Chelsea’ to c_c , and ‘loves’ to C_L ,

and that, therefore, the Fregean proposition to which the that-clause refers may be represented by

$$(6) \langle \langle c_b, c_c \rangle, C_L \rangle^6$$

which is true just in case

$$\exists xyR[(x \text{ falls under } c_b \ \& \ y \text{ falls under } c_c \ \& \ R \text{ falls under } C_L) \ \& \ (\langle x, y \rangle \text{ instantiates } R)];$$

false if $\langle x, y \rangle$ does not instantiate R ; and where it is left as a matter of possible disagreement among Fregeans whether to say that (6) is false or neither true nor false if nothing falls under one of its three concepts.

There is, we should notice at this point, an important asymmetry between the Russellian position and the generic Fregean position just described—namely, that the Russellian conception is fully specified in a way the Fregean conception is not. The Russellian conception is fully specified because we know what the components of Russellian propositions are, and thus know what those propositions are. Not so for the Fregean conception. There is the *illusion* that we know what the components of Fregean propositions are because the Fregean has borrowed familiar terms—‘concept’, ‘mode of presentation’, ‘way of thinking’, and so on—to stand for those components; but it is an illusion because in the context of her theory these terms are technical terms meaning little more than *the basic components of propositions, assuming those components are not the objects and properties our beliefs are about*. To be sure, the Fregean chooses terms like ‘concept’ and ‘mode of presentation’ because of the suggestiveness of their pretheoretic meanings, but no one of those many pretheoretic meanings does all the technical work the Fregean requires. And if we do not know what Fregean concepts are, then we do not yet know what Fregean propositions are. The generic Fregean theory, in other words, does not say what concepts, the components of Fregean propositions, are, and this is an issue on which Fregeans may differ among themselves. There will be more on this presently.

Although the distinction between strong Russellianism and strong Fregeanism is not an exhaustive classification of theories of structured propositions, it does yield an exhaustive and exclusive distinction concerning the reference of a given expression in a given that-clause: trivially, given the generic Fregean’s use of ‘concept’, this reference will either be an object or property or else a concept of one. This in turn yields the following partitioning of theories of structured propositions, which is exhaustive as regards structured propositions:

- The reference of a that-clause is always/sometimes a Russellian proposition (strong/weak Russellianism).

- The reference of a that-clause is always/sometimes a Fregean proposition (strong/weak Fregeanism).
- The reference of a that-clause is sometimes a Fressellian proposition (if I may), such as the proposition represented by $\langle\langle \text{Bill}, c_c, C_L \rangle\rangle$, which could result from the reference of 'Bill' in (4) being Bill, while the other terms have Fregean references (the Fressellian position).⁷

Each of these positions has problems.

One famous problem with strong Russellianism is the problem of "empty names." Intuitively,

- (7) Reggie believes that George Eliot was a woman, but in fact there was no such person; a committee wrote all the novels

might well be true if there were no such person as George Eliot. But, on the face of it, the strong Russellian must hold that (7) would have no truth-value if George Eliot did not exist, as in that case its that-clause would fail to refer.⁸

Another famous problem comes to light with an example such as

- (8) Ralph believes that George Eliot adored groundhogs but does not believe that Mary Ann Evans adored woodchucks.

It seems obvious that (8) might well be true, notwithstanding that George Eliot = Mary Ann Evans and the property of being a groundhog = the property of being a woodchuck. But if the two that-clauses in (8) refer to Russellian propositions, then they refer to the same Russellian proposition and—given the working hypothesis that belief reports of form (i) ('A believes that S') say that A stands in the two-place belief relation to the reference of 'that S'—it therefore follows that (8) is not only false, but necessarily false.⁹ It is true that Nathan Salmon and others have shown that the strong Russellian is not without sophisticated resources for trying to explain away the intuition to which I just appealed, but my own view, which I cannot now defend, is that it is problematic whether those resources can do the job (see, e.g., Salmon 1986 and 1989, and Schiffer 1987b).

A third problem is actually just a variant of the second, except that it involves utterances of the same that-clause. Thus, Lois Lane may say

- (9) I believe that he flies, but I don't believe that *he* flies

where the first utterance of 'he' is accompanied by her pointing to a photo of a man in a caped spandex outfit and the second utterance of it is accompanied by

her pointing to a photo of a bespectacled man wearing a suit and tie. Intuitively, her utterance of (9) may be true even when both utterances of 'he' refer to the same person. Or, to take an example made famous by Saul Kripke (Kripke 1979), two simultaneous utterances of

Herbert believes that Paderewski was musical.

may have different truth-values even though both occurrences of 'Paderewski' refer to the same person. This could happen when Herbert mistakenly believes there were two famous Poles named 'Paderewski', one a statesman, the other a pianist, and when what matters in the conversational context of the first utterance is Herbert's willingness to say 'I've no idea whether Paderewski was the slightest bit musical', while what matters in the other conversational context is Herbert's willingness to say 'Paderewski was astoundingly musical'. The strong Russellian is committed to denying the intuitive data, since for her both utterances of the belief sentence must have the same truth-value, given the working hypothesis about the logical form of belief reports (together, to be sure, with her views about the references of names).

There is another possible problem I'll discuss later, since if it is a problem for the strong Russellian, then it is equally a problem for every theorist of structured propositions.

Weak Russellianism, according to which only some that-clauses have Russellian propositions as their references, is also problematic. Consider the problem raised for the strong Russellian by Lois Lane's utterance (9). If the strong Russellian sticks to her guns (unreasonably, I should think) and denies that (9) can be true, then there is no reason for her not to take the same hard line, *mutatis mutandis*, regarding the other *prima facie* counterexamples to her theory. But if she allows that (9) may be true, and that therefore one of its two occurrences of 'that he flies' does not have a Russellian proposition as its reference, then where will she find a that-clause whose reference is a Russellian proposition? The problem to which I am alluding may be elaborated in the following way. When t appears to enjoy a Russellian reference in the true utterance

A believes that $S(t)$ ¹⁰

it will always be possible that there should be another utterance

A believes that $S(t')$

where t' may or may not be the same as t , such that, first, the two utterances have different truth-values, and, second, t' also appears to enjoy the same

Russellian reference as *t*. For example, after an accident that caused her to suffer amnesia, Claudia Schiffer (no relation) may say both

I don't believe that I am German

and

I believe that Claudia Schiffer is German

where, intuitively, both reports are true. A theorist who claims that one but not the other of the two that-clauses has a Russellian proposition as its reference will find it impossible to say which one enjoys that reference; whatever she says about the reference of the one that-clause, she will be equally constrained to say about the other. One might protest that this creates an impossible situation for me: it is trivially analytic that the reference of a term in a that-clause is either Russellian or Fregean; yet, the protest continues, the reference surely cannot be Fregean, since in the examples under discussion the terms in question refer to ordinary objects—the superhero with two names, Claudia Schiffer—rather than to concepts of them. Presently we shall see that the Fregean may have a way of denying that the reference “surely can't be Fregean.”

If the weak Russellian position fails for the reason just indicated, then the Fressellian position fails for the same reason. The Fressellian is, as it were, a theorist who begins as a strong Fregean but feels the need to revise her position because terms in that-clauses often refer to things our beliefs are about. But recourse to Fressellian propositions is a well-motivated response only if the weak Russellian account of the reference of the expressions in question is adequate, and I have just suggested that it is not.

I realize, of course, that there is more to be said for and against the foregoing three positions, but a definitive discussion is not the purpose of my brief critical survey. Nevertheless, I would like to register my belief that neither strong Russellianism, weak Russellianism, nor Fressellianism can survive the full treatment. If that-clauses refer to structured propositions, those propositions are Fregean; that is to say, given CH, strong Fregeanism is correct.

At the same time, strong Fregeanism is not without its problems. One problem is that, as noted, all we know from the generic Fregean about concepts is hardly more than that they are the references expressions have in that-clauses, assuming those references are not Russellian references. So the generic Fregean has yet to tell us what the building blocks of his propositions are, and so has yet to tell us what Fregean propositions are. I call this a problem, as opposed to a mere further task, because it is arguable that no Fregean has so far succeeded in giving an adequate account of Fregean concepts.

A second problem is one to which I have already alluded. The strong

Fregean holds that the reference of a term in a true belief report's that-clause is a concept of an object or property the belief purports to be about, yet it is obvious that terms in such that-clauses often refer to things that are not concepts but are the very things the belief is about. When your husband's brother says to you

I believe that I'm falling in love with you,

isn't it obvious that both utterances of 'I' refer to him and that his utterance of 'you' refers to you? To this problem the Fregean may have a reply that I believe was first clearly articulated by Gareth Evans (1982), but it is a reply that makes a solution to the first problem even more urgent. Inspired by Evans, the Fregean may say that the problematic referring expressions have as their references *object-dependent* concepts, concepts that are individuated partly in terms of the objects of which they are concepts, and would not exist if those objects did not exist. The Fregean can then distinguish between (i) x 's being *the reference* of t and (ii) t 's merely *referring to* x . The idea would be that t refers to x if x is the reference of t , but not necessarily conversely. An object may be so transparently contained (so to say) in an object-dependent concept of it that one cannot refer to the concept without thereby indirectly referring to the object. Moreover, object-dependent propositions will have the objects contained in object-dependent concepts entering rigidly into the possible-worlds truth conditions of those propositions. In this way, the Fregean may say, the that-clauses in Claudia's two utterances may refer to distinct object-dependent propositions about her that have the same possible-worlds truth condition: each is true in a possible world w just in case Claudia herself exists and is German in w . But the Fregean who takes this line still owes an account of the nature of object-dependent (not to mention object-independent) concepts.

A third problem is that it is apt to seem that, whatever concepts turn out to be, there are cases where it is implausible to think anything worth calling reference to a concept is going on. An example of such a case is the belief report

Just about everyone who visits New York City believes that it's noisy,

which is both true and easily understood, even though, it would appear, nothing worth calling a concept would be understood to be the reference of 'it'. Similarly, you may believe what I tell you when I say 'Hilda believes that that guy is on his way there from Paris', but would you thereby know the concepts under which Hilda, who is not party to the utterance, is thinking about Jacques Derrida, the *x-is-on-the-way-to-y-from-z* relation, the University of Idaho, and Paris?

Finally, there is an argument—due to Adam Pautz, an NYU graduate stu-

dent—against the Fregean position in either its strong or weak form that I think is worth stating. Frankly, the argument surprises me. On the one hand, I am not aware of having encountered it before I heard it from Pautz in a seminar I gave in the fall of 1999, while, on the other hand, no one has succeeded in telling me what exactly is wrong with it. The argument may be put thus:

- (1) If the Fregean theory is true, then [*] the reference of ‘Fido’ in ‘Ralph believes that Fido is a dog’ is a concept of Fido.¹¹
- (2) If [*], then the following inference is valid when ‘ x ’ is taken to range over concepts:

Ralph believes that Fido is a dog
 $\therefore \exists x$ (Ralph believes that x is a dog)
- (3) But the inference is not valid in the intended sense: a concept can be the value of ‘ x ’ that makes the conclusion true only in the unlikely event that Ralph mistakes it for a dog.
- (4) \therefore The Fregean theory isn’t true.

Pretty clearly, if this argument is not sound, it is because (2) is false. The problem is that (2) is based on an evidently well-established logico-semantical principal: if o is the reference of t in the true sentence $S(t)$ [and, thinking of ‘Giorgione was so-called because of his size’, t makes no other contribution to the truth-value of $S(t)$], then o makes true the existential generalization $\exists xS(x)$. And please don’t remind me that the Fregean theory precludes substitution *salva veritate* of ‘the concept of Fido’ for ‘Fido’ in the that-clause, since when ensconced in the that-clause, ‘the concept of Fido’ would refer not to the concept of Fido but to the concept of the concept of Fido. The argument proceeds in full awareness of that aspect of Frege’s theory and is not intended to challenge it; the force of the argument turns only on the fact that, if the Fregean theory is right, a concept is *the reference* of ‘Fido’ in the that-clause. In this connection, it may be helpful to keep in mind that for the Fregean, the position of the that-clause is entirely referential and transparent, so that if the that-clause refers to the Fregean proposition $\langle c_p, C_d \rangle$, then from ‘Ralph believes that Fido is a dog’ we can get, *salva veritate*, ‘Ralph believes $\langle c_p, C_d \rangle$ ’ [which unproblematically entails ‘ $\exists x$ (Ralph believes $\langle x, C_d \rangle$)’]. This is why inferences such as

Louise believes that existence precedes essence
 That existence precedes essence = Jean-Paul’s theory
 \therefore Louise believes Jean-Paul’s theory

are valid. I’m not saying that the substitutivity business is not relevant to the falsity of (2), if (2) is false; I am just saying that an explanation would be needed to show that, and how, it is relevant. To say that it is merely a quirk of English

(and, presumably, of every other language!) that exportation fails for expressions in that-clauses, even though they occur there as singular referring expressions, is no explanation at all. Now, one might feel that there *must* be an explanation of why the inference example in (2) is not valid in the way in question, since if it were valid one would then have to choose between two unpalatable alternatives: that the reference of ‘Fido’ in the that-clause was Fido, or that the that-clause did not refer to something Ralph is being said to believe. But, as we are about to see, there is another option: deny the Compositionality Hypothesis (CH), according to which the reference of a that-clause is a function of the references its primary expressions have in the that-clause. Then we could take the that-clause to refer to a finely individuated *unstructured* proposition, and we would not have to say the problematic inference was valid in the sense in question. To this one might reply that CH is required by the hypothesis that every natural language has a compositional truth theory, which is itself required to explain certain features of natural languages. But, as I have tried to show elsewhere (see especially Schiffer 1987a, ch. 7, and Schiffer 1994), the features in need of explanation can be explained without recourse to either a compositional truth theory or a compositional meaning theory (given the relevant technical meanings of those labels).

There is also an inference argument whose soundness would refute every theory of structured propositions:

- (1) If any theory of structured propositions is true, then [#] ‘is a dog’¹² functions in ‘Ralph believes that Fido is a dog’ as a singular term whose reference is either the property of being a dog or else a concept of that property.
- (2) If [#], then the following inference is valid:
 Ralph believes that Fido is a dog
 ∴ $\exists x$ (Ralph believes that Fido x).
- (3) But not only isn’t this valid, it’s not even coherent.
- (4) ∴ No theory of structured propositions is true.

Curiously, when confronted with this argument, philosophers immediately challenge (1) and do not even mention (2). There is much to justify (1); ‘is a dog’ on theories of structured propositions certainly is not functioning as any kind of predicate; its function is on all fours with ‘doghood’/‘the concept of doghood’ in ‘< . . . , doghood/the concept of doghood>’, viz., to introduce the property or concept into the structured proposition to which the singular referring expression containing it (‘that Fido is a dog’, ‘<Fido/the concept of Fido, doghood/the concept of doghood>’) refers. I will postpone further discussion for another occasion. I present the objection here merely as food for thought.

So where do we go from here? To a better understanding of the nature of propositions, for which I now set the stage.

PLEONASTIC ENTITIES

This is my label for entities whose existence is typically secured by what I call a *something-from-nothing transformation*. We have a something-from-nothing transformation when from a statement involving no reference to an F we can deduce a statement that does refer to an F . The property of being a dog is a pleonastic entity. From the statement

(10) Lassie is a dog,

whose only singular term is ‘Lassie’, we can validly infer its pleonastic equivalent

(11) Lassie has the property of being a dog,

which contains the new singular term ‘the property of being dog’, whose reference is the property of being a dog.

There are many other kinds of pleonastic entities, including propositions, but before turning to the objects of our immediate concern, let me continue to set the stage for them by staying a little with the example of properties.

How are pleonastic properties possible? That is, what explains the something-from-nothing transformations by which they are introduced into our conceptual scheme? How *can* (10) entail (11), thereby entailing that the property of being a dog exists? We can thicken the plot by first taking on a seemingly unrelated question: How are we able to have knowledge about properties, mind- and language-independent abstract entities that are wholly incapable of causally interacting with us? This question is made more vivid in the following way. Suppose there is a possible world exactly like ours except that our counterparts in that world do not have any property-hypostatizing linguistic or conceptual practices, and hence have no concept of a property. These people can think that Lassie is a dog, but they cannot infer from this that Lassie has the property of being a dog, even though in that world, as in every world in which Lassie exists, it is necessarily true that if Lassie is a dog, then Lassie has the property of being a dog. Lacking the concept of a property, these people are entirely ignorant of properties, even though they live in a world as rich in properties as the actual world. What would it take to bring these people up to epistemological snuff with us?

What it would take, and all that it would take, would be for them to engage in a certain manner of speaking, a certain language game—namely, our property-hypostatizing practices, in particular our property-yielding something-

from-nothing transformations. But how can this be? They certainly could not discover the existence of *volcanoes* by engaging in any language game. How can merely engaging in a linguistic, or conceptual, practice give one knowledge of things that exist independently of that practice? The answer to this question also answers our question about the something-from-nothing transformation, about how (I0) can entail (II).

The fact that Lassie is a dog entails that Lassie has the property of being a dog. Jones knows that Lassie is a dog but does not know that Lassie has the property of being a dog. This is because he lacks that concept of a property that comes (nearly enough) with the practice of making the something-from-nothing inference from the fact that Lassie is a dog to the fact that Lassie has the property of being dog. *Once Jones has the concept of a property, it will be a **conceptual truth** for him that every dog has the property of being a dog.*

Maybe you feel like reading Kant to me. Kant, in response to the ontological argument for the existence of God, famously held that “existence isn’t a predicate,” where by this he meant that no mere concept, however defined, can secure that there exist things that fall under the concept. Harry Field, endorsing Kant’s point, has succinctly restated it thus:

An investigation of conceptual linkages can reveal conditions that things must satisfy if they are to fall under our concepts; but it can’t yield that there are things that satisfy those concepts (as Kant pointed out in his critique of the ontological argument for the existence of God). (Field 1989, 5)

Consider the concept of a *wishdate*:

x is a *wishdate* =_{df} x is a person whose existence supervenes on someone’s wishing for a date, every such wish bringing into existence a person to date.

The point that Kant and Field are making implies that while this is a perfectly kosher definition, it does not result in its being true that there are any wishdates, no matter who wishes for a date. All that follows from the stipulative definition of a wishdate is that *if (per impossibile) wishdates exist, then their existence supervenes on the mere wish for a date.*

There is, however, a relevant difference between the concept of a wishdate and the concept of a property, and this difference allows us to see that, and how, properties are an exception to the Kantian dictum. The intuitive idea that will need precisification is that the existence of properties, but not the existence of wishdates, would make no causal difference. Here is one way we can try to precisify this intuitive idea. As a preliminary, let me remind you that a theory T' is

a *conservative extension of a theory T with respect to statements of kind K* provided that T' contains T and whatever K statements are entailed by T' are also entailed by T . We may then explain the relevant difference between the concept of a property and that of a wishdate via the following claim:

The concept of an F can secure the existence of F s **only if** when we add the statement ‘If T , then F s exist’ to any true theory T that doesn’t itself employ the notion of an F , we get a new theory that’s a conservative extension of T with respect to statements (i) that don’t formally entail the existence of F s and (ii) whose truth would have causal consequences.¹³

The concept of a wishdate fails this test, but the concept of a property passes it, and this accounts for the crucial difference between the two concepts. Kant was right about the concept of a wishdate, wrong about the concept of a property. The displayed criterion is stated as a necessary condition, but when it is elevated into a sufficient condition as well, we may take the notion of a pleonastic entity to be defined by its satisfying the criterion. I am being cautious about sufficiency because presumably something would have to be added to rule out concepts whose instantiations were physically impossible independently of the way they contain the notion of existence, such as the way the concept of an *impotent wishdate*, a wishdate with no causal powers, is physically impossible just by virtue of requiring there to be a person who has no causal powers. Maybe other qualifications would be needed; I will not now try to state a sufficient condition. In the meantime, I submit that we are in a position to appreciate that, *pace* Kant, it is a *conceptual truth* that if Lassie is a dog, then Lassie has the property of being a dog.

We can also see (nearly enough) how engaging in the relevant property-hypostatizing linguistic or conceptual practices is necessary and sufficient for having knowledge of the existence of properties.¹⁴ First, if you engage in the practice, you have the concept of a property, and to have any concept is to engage in certain conceptual (and thereby, under certain assumptions, linguistic) practices. Second, if you have the concept of a property, then you know *a priori* the conceptual truth that, e.g., every dog has the property of being a dog, and thereby know that the property exists. We also see how we can explain the something-from-nothing transformation that takes one from (I0) (‘Lassie is a dog’) to (I1) (‘Lassie has the property of being a dog’). That is simply a direct and obvious consequence of its being a conceptual truth that if Lassie is a dog, then Lassie has the property of being a dog.

PLEONASTIC PROPOSITIONS

Propositions, too, are introduced into our conceptual scheme via something-from-nothing transformations. From (I0) (‘Lassie is a dog’) we can validly infer another of its pleonastic equivalents, viz.,

That Lassie is a dog is true

or, more colloquially,

It's true that Lassie is a dog,

which contains the new singular term 'that Lassie is a dog', whose reference is the proposition that Lassie is a dog. Of course, the language game we play with that-clauses is not limited to these something-from-nothing transformations, or even to what is indirectly implied by them. Crucially, there is also the use of that-clauses in ascriptions of meaning, speech acts, and psychological states like belief. These practices presuppose the validity (subject to certain qualifications¹⁵) of the something-from-nothing transformation that yields the truth schema for propositions,

The proposition that S is true iff S

but the role of that-clauses in propositional-attitude discourse is not deducible from the something-from-nothing practice and is essential to completing the account of pleonastic propositions. *The relation between that-clauses and the propositions to which they refer is quite unlike the usual relation between singular terms and their references.*

Typically, if t is a referential singular term in an utterance of the sentence $S(t)$, then in order to evaluate the statement made by the utterance we must first identify the reference of t and then determine whether that reference has the property expressed by $S(x)$. The reference of t , once itself determined, partially determines the criteria for evaluating the statement made by the utterance of $S(t)$: we fix the reference of t , and thereby fix, or partially fix, the criteria for truth-evaluating the utterance of $S(t)$. A consequence of this is that the reference of t is guaranteed to have an identity and individuation that is entirely independent of the criteria of evaluation which that reference helps to determine. These points are illustrated by the pair of statements

(12a) Bill loves Chelsea.

(12b) Bill loves Hillary.

In order to evaluate the statement made in (12a), we must first identify the references of 'Bill' and 'Chelsea'; likewise, *mutatis mutandis*, for (12b). It would be outrageous to suppose that we first fix the criteria of evaluation and then use those criteria to determine the references. We evaluate the statements made in (12a,b) by glomming onto the references of 'Bill', 'Chelsea', and 'Hillary' and then determining whether the first stands in the love relation to the other two. Accordingly, the identity and individuation of Bill, Chelsea, and Hillary owe

nothing at all to the criteria of evaluation they help to determine. It is because Chelsea and Hillary each have an identity and individuation that is entirely independent of the criteria for evaluating the statements about them that we immediately see the absurdity of supposing that we know that Chelsea \neq Hillary because we know that the statements made in (12a) and (12b) may differ in truth-value. On the contrary, of course: we know that the two statements may differ in truth-value *because we know that Chelsea \neq Hillary*.

Matters are just the opposite when we turn to that-clauses and their references. In a belief report, we *first* have contextually determined criteria of evaluation and *then* implicitly use those criteria to determine *both* the proposition to which the that-clause refers *and* the individuation of that proposition. These criteria of evaluation are in part determined by contextual factors pertaining to the communicative interests of speakers and their audiences, even after disambiguation and obvious reference-fixing has taken place.¹⁶ Two literal utterances of

(13) Ralph believes that George Eliot was a man

may have different truth-values, because in one conversational context but not the other the truth of the utterance requires thinking of George Eliot as a famous author. Given our on-going assumption about logical form—that ‘*A* believes that *S*’ says that *A* stands in the two-place belief relation to the reference of ‘that *S*’—it follows that the two utterances of ‘that George Eliot was a man’ refer to different propositions, albeit, no doubt, to propositions with the same possible-worlds truth condition: each is true in a possible world *w* just in case George Eliot—the woman who *actually* bears that name—exists and is a man in *w*. The same applies to two utterances of

(14) Ralph believes that she wrote *Ivanhoe*

when both occurrences of ‘she’ refer to George Eliot. If we were evaluating an utterance of ‘Ralph admires her’ we would first determine the reference of ‘her’ and that would in turn complete the determination of the criteria for evaluating the statement. In evaluating the statement made in the utterance of (14), however, we first implicitly fix the criteria for evaluating the statement, and that is what fixes the reference of the that-clause. This is not to deny that the semantic properties of expressions in a that-clause are not crucial to the determination of the that-clause’s reference; in order to fix the criteria of evaluation for the utterances of (13) and (14) we must first know to whom the utterances of ‘George Eliot’ and ‘she’ refer. My point is that these semantic properties on their own do not determine the references of that-clauses; rather, those semantic properties help to determine the criteria of evaluation for belief reports, which criteria in turn fix the references of that-clauses.

At the same time, the contextually determined criteria of evaluation may themselves determine what semantic properties of expressions in that-clauses are relevant. We noticed above (p. 84) that a problem for the Russellian was that an utterance of, e.g.,

Mary believes that George Eliot wrote *Middlemarch*

may possibly be true even if it turns out that George Eliot never existed, the novels credited to a bearer of that name having been written by a committee. The contextual factors determining criteria of evaluation, and thus determining the proposition to which the utterance of the that-clause refers, may depend on how much the belief reporter and her audience know. If, for example, they believe that George Eliot existed, then the proposition to which the that-clause refers will most likely be an *object-dependent* proposition, a proposition that would not exist if George Eliot had not existed and is partly individuated in terms of her. But if they knew, or even believed, that George Eliot did not exist, then the contextually determined criteria of evaluation would determine an object-independent proposition as the reference of the that-clause.¹⁷

Contextually determined criteria of evaluation for a belief report not only determine *which* proposition its that-clause refers to; *they also enter into the very individuation of that proposition*. If, for example, a particular utterance of (13) is true only if Ralph thinks of George Eliot as a famous author, then the proposition to which that utterance of ‘that George Eliot was a man’ refers is such that believing it entails thinking of George Eliot as a famous author. A different occurrence of the that-clause will refer to a different proposition—albeit one with the same truth condition—if the truth of the belief report containing it does not entail thinking of George Eliot as a famous author. Although a proposition may exist in possible worlds in which there are neither thinkers nor speakers, its identity and individuation is owed in part to criteria that must be satisfied in order to count as believing it. It belongs to the essence of the proposition that it, and it alone, can satisfy the criteria of evaluation that determine it.

The striking differences between the way ordinary singular terms operate and the way that-clauses do is nicely illustrated when we compare the pair (12a,b) with the following two pairs:

- (15a) Nobody doubts that whoever believes that all ophthalmologists are ophthalmologists believes that all ophthalmologists are ophthalmologists.
- (15b) Nobody doubts that whoever believes that all ophthalmologists are ophthalmologists believes that all ophthalmologists are eye doctors.¹⁸
- (16a) Lois believes that Superman flies.
- (16b) Lois believes that Clark Kent flies.

We know that the two members of all three pairs may well differ in truth-value, but there is a very important difference between (12a,b), on the one hand, and, on the other hand, (15a,b) and (16a,b). As already noted, it is absurd to suppose we know that Chelsea \neq Hillary because we know that (12a) and (12b) may differ in truth-value; rather, we know that (12a) and (12b) may differ in truth-value *because* we know that Chelsea \neq Hillary. But just the opposite obtains as regards the other two pairs. We do not know that (15a) and (15b) may differ in truth-value because we know that the proposition that whoever believes that all ophthalmologists are ophthalmologists believes that all ophthalmologists are ophthalmologists \neq the proposition that whoever believes that all ophthalmologists are ophthalmologists believes that all ophthalmologists are eye doctors; rather, we know that the proposition that whoever believes that all ophthalmologists are ophthalmologists believes that all ophthalmologists are ophthalmologists \neq the proposition that whoever believes that all ophthalmologists are ophthalmologists believes that all ophthalmologists are eye doctors *because we know (15a) and (15b) may differ in truth-value*. Likewise for (16a) and (16b): we first know that they may differ in truth-value, and on this basis we know that the proposition that Superman flies \neq the proposition that Clark Kent flies.

ARE PLEONASTIC PROPOSITIONS STRUCTURED?

Suppose they are structured propositions. Then they are Fregean propositions (albeit *pleonastic* Fregean propositions) composed of pleonastic concepts of the objects and properties our beliefs purport to be about, and we would perforce take those concepts to be the references expressions have in that-clauses. What more could we say about these pleonastic concepts?

(i) The proposition to which the that-clause in a belief report refers is determined by, and owes its identity and individuation to, contextually determined criteria for truth-evaluating the belief report that are as if the proposition to which the that-clause refers did not even exist. Consequently, whatever concepts “compose” that proposition must also be determined by, and owe their identity and individuation to, the same criteria of evaluation. But these criteria directly determine the proposition; they do not first determine propositional components, which only then enable us to determine a proposition. Pleonastic concepts composing pleonastic propositions would have to be abstractions from the propositions containing them, with, so to say, no life of their own apart from those propositions. They would perforce be individuated in terms of the propositions containing them and would be tantamount to equivalence classes of propositions: the concept to which an expression in a that-clause refers would, for all intents and purposes, be the class of propositions equivalent in such-and-such respect to the proposition to which the that-clause refers. For example, the that-clause in a particular utterance of ‘Ralph believes that George Eliot was a man’ may refer to a proposition that, intuitively speaking, requires thinking of

George Eliot in the same way as other propositions to which that-clauses containing ‘George Eliot’ refer, and we may trivially think of those tokens of ‘George Eliot’ as associated with a certain equivalence class of propositions: the class of propositions equivalent to the one in question with respect to how they require thinking of George Eliot. The same may be said, *mutatis mutandis*, for any expression.

(ii) We may, however, speak without difficulty of pleonastic concepts being concepts *of* the objects and properties our beliefs are about. Every belief report

A believes that *S*

will entail the canonical truth condition

A's belief that *S* is true iff *S*.

It is because of this that all the somewhat different propositions to which occurrences of ‘that George Eliot wrote *Ivanhoe*’ may refer in distinct utterances of ‘Ralph believes that George Eliot wrote *Ivanhoe*’ will all be true iff George Eliot wrote *Ivanhoe*. We can therefore say that, roughly speaking, a pleonastic concept is a concept *of* *x* just in case *x* enters into the canonical truth condition of every proposition containing the concept. There would be no further need, (see Peacocke 1992), for a substantial theory of reference for concepts.

(iii) If pleonastic propositions are structured entities composed of pleonastic concepts, then we at least have a version of Fregeanism which answers the first three problems for the theory mentioned earlier (pp. 86–87).

First, of course, we would have a version of Fregeanism—*Pleonastic Fregeanism*—which says what concepts, *qua* references of expressions in that-clauses and propositional building blocks, are.

Second, Pleonastic Fregeanism would be unthreatened by the fact that words in that-clauses often refer to the things the belief is about. Earlier we noticed the intuitiveness of holding that Claudia spoke truly when she said

I believe that Claudia Schiffer is German but not that I am,

even though ‘Claudia Schiffer’ and ‘I’ there referred to the same person. What the Pleonastic Fregean can say is that the references of the name and pronoun are distinct concepts of the demonstrated woman, albeit *object-dependent* concepts of her. Since the two distinct concepts are dependent in the way of object-dependent concepts on the woman of which they are concepts, there is a sense in which they transparently contain her, so that in referring to the concepts one is also referring to the woman. The woman, then, is not *the reference* of the tokens of ‘Claudia Schiffer’ and ‘I’; but because the concepts that are the

references are object-dependent concepts of the woman, the Pleonastic Fregean can say that Claudia is also referring to herself in uttering ‘Claudia Schiffer’ and ‘I’. As already noted (p. 87), the Fregean can say that the fact that an expression refers to a thing does not entail that that thing is *the reference* of the expression, for the reference may itself transparently require the existence of the thing in question. This diagnosis gets support from the observation, confirmed by the Claudia example, that an expression in a that-clause often plainly refers to a thing even though, just as plainly, substitutivity *salva veritate* is not assured (were Claudia to replace her utterance of ‘I believe that Claudia Schiffer is German but not that I am’ with an utterance of ‘I believe that I am German but not that I am’, she would be replacing a truth with a falsehood). As before, the Pleonastic Fregean would be able to explain this by saying that the tokens of ‘Claudia Schiffer’ and ‘I’ refer to distinct concepts of the same person, but, since those concepts are object-dependent concepts of her, she, too, is referred to in the very reference to the concepts.¹⁹ (To be sure, any Fregean who agrees with the data must hold that the singular terms in question have as their references object-dependent concepts. The sense in which the Pleonastic Fregean has an explanatory advantage is that she has an account of object-dependent concepts.)

Finally, Pleonastic Fregeanism can accommodate the fact that we typically make correct belief reports without being in a position to say very much about how the believer is thinking about the things the belief is about. The theory holds that the concepts to which the expressions in a that-clause refer are determined by implicit contextual factors that determine what will count as verifying and falsifying the belief ascription, and therewith individuating the proposition to which the that-clause refers. The “concepts” thus determined would reflect contextually relevant beliefs about the things the concepts were of, because such beliefs would be among the things determining the ascriptions’ truth-values, and therewith determining the propositions to which their that-clauses refer. To this extent, pleonastic concepts reflect how believers are thinking about the things their beliefs are about, but the way in which concepts are determined in no way requires belief ascribers to know much at all about how believers are thinking about the things their beliefs are about. This allows the concept to which an expression in a that-clause refers to be quite thin. In an extreme case, one I think hardly, if ever, occurs, the “concept” determined may be so thin as to be tantamount to a mere reference to the thing itself. Perhaps something close to this is going on in the statement ‘Just about everyone who’s been to New York City believes that it’s noisy’ (thus, some propositions would in the limit, so to say, be in effect Russellian as well as Fregean).

But it is not clear that the proponent of pleonastic propositions should accept Pleonastic Fregeanism. It is not clear that he should hold that pleonastic concepts are the references expressions have in that-clauses, so it is not clear that

he should hold that pleonastic propositions are structured propositions composed of pleonastic concepts.

There is no pressure for a proponent of pleonastic propositions to accept Pleonastic Fregeanism unless he also accepts the Compositionality Hypothesis (the reference of a that-clause is a function of the references of its component expressions), and there are the following reasons for doubting that he should accept it.

(a) For many semantically complex singular terms, it is intuitively correct that the reference of the complex whole is a function of the references of its parts. But, to repeat (p. 81), the justification for extending this to all complex singular referring expressions would be that it is required by the hypothesis that natural languages have compositional semantics and that that hypothesis is needed to explain various features of language; yet by my lights, once again (Schiffer 1987a, 1994), compositional semantics is not needed to explain those features.

(b) Pleonastic Fregeanism has *per se* no reply to Pautz's objection (p. 88), the last of the Fregean's problems I mentioned.

(c) When the reference of a singular term is compositionally determined by the references of its component expressions, we expect that in order to know the reference of the complex singular term, one must first know the references of its parts. For example, one cannot know the reference of the singular term '<Lassie, doghood>' except by first fixing the references of 'Lassie' and 'doghood'. But if 'Lassie' and 'is a dog' (or 'dog') in

(17) Fiona believes that Lassie is a dog

refer to concepts composing the pleonastic proposition to which 'that Lassie is a dog' there refers, then those concepts play no role at all in one's determination of the reference of the that-clause. Since the contextually determined criteria for truth-evaluating (17) determine the proposition to which the that-clause refers, as well as the conditions for individuating that proposition, those criteria *ipso facto* also determine the identity and individuation of the concepts to which, we are supposing, the words in the that-clause refer. The concepts to which the words would refer would be determined, as it were, by abstraction from the already-determined proposition to which the that-clause refers and would play no role at all in our determination of that reference.

(d) There is independent reason to be skeptical of pleonastic concepts as propositional building blocks. To see the problem, suppose it is suggested—as I in fact suggested in (Schiffer 1994)—that that-clauses refer to *unstructured* pleonastic propositions. Would this suggestion then be refuted by the correct observation that each pleonastic proposition can easily be construed as a func-

tion of equivalence classes of pleonastic propositions? I do not see that it really would be. Theorists who speak of propositional building blocks invariably have in mind entities that can be identified and individuated apart from the propositions they build, and this would not be true if the alleged building blocks were equivalence classes of the propositions they were supposed to build. Now, if that shows that equivalence classes of pleonastic propositions cannot serve as pleonastic concepts, i.e., as the basic components of structured pleonastic propositions, then it also shows there can be no pleonastic concepts. For if there are pleonastic concepts, then their conditions of individuation make them tantamount to equivalence classes of pleonastic propositions, if it is even possible for them to be conceived as anything other than such classes.

Should we therefore conclude—putting aside Pautz’s objection—that Pleonastic Fregeanism is untenable? I do not know if that question has a determinate answer. I am, however, confident that there is no substantial reason to care about the answer. What on earth can turn on the difference between the view that that-clauses refer to unstructured pleonastic propositions and the view that they refer to structured pleonastic propositions whose basic components are equivalence classes of pleonastic propositions, or things tantamount to them?

CONCLUSION

After assuming, as a working hypothesis, that believing is a two-place relation between believers and things to which that-clauses refer, I argued that those things are *pleonastic propositions*. Let us say that *Heavy-Duty Fregeanism* is the doctrine that the references of expressions in that-clauses, and therewith the building blocks of propositions, have an identity and individuation distinct from that of the propositions they build. Heavy-Duty Fregeanism is incompatible with belief’s being a relation to pleonastic propositions. This leaves open the question whether pleonastic propositions are structured entities whose basic components are pleonastic concepts construed, for all intents and purposes, as equivalence classes of the pleonastic propositions they compose, or whether pleonastic propositions are unstructured. I said we should not worry whether this question even has a determinate answer, since nothing of interest turns on it. The complete story on all these matters will be discussed elsewhere (Schiffer forthcoming).

NOTES

This chapter is a revision of what I gave as the University of Idaho’s annual Seaman lecture and the keynote address at the third annual INPC. My revision benefited from the excellent comments made during the discussion of my talk. I am also indebted to the comments on an earlier draft by Ray Buchanan, Cian Dorri, Paul Horwich, Michael O’Rourke, Chris Peacocke and Josh Schecter. Versions of this paper were also given as talks at the CUNY Graduate Center, CREA in Paris, and the GAP conference in Bielefeld, Germany, all in the fall of 2000. Comments made in those discussions were also very helpful.

1. For simplicity of exposition I am ignoring considerations pertaining to tense.

2. I argued against this working hypothesis in Schiffer 1987a, but the deflationary aspect of what I shall be arguing for in this article makes my present position more of a sequel to Schiffer 1987a than an apostasy of it.
3. The primary expressions in a that-clause are those whose references contribute directly to the truth-value of the proposition to which the that-clause refers; secondary expressions contribute their references only to determine the references of the primary expressions containing them. For example, if, rather implausibly, the that-clause in 'Louise believes that that guy talking to Maria is a violinist' refers to a "singular proposition" containing the guy talking to Maria and the property of being a violinist, then 'the guy talking to Maria' is a primary expression, and 'Maria' is a secondary expression.
4. 'Russellian' and 'Fregean', of course, after Bertrand Russell and Gottlob Frege.
5. For the Russellian, every proposition may be taken to be an ordered pair of the form $\langle\langle x_1, \dots, x_n \rangle, R^n\rangle$, where $\langle x_1, \dots, x_n \rangle$ is an n -ary sequence of things of any ontological category, R^n is an n -ary relation (properties are one-place relations), and where $\langle\langle x_1, \dots, x_n \rangle, R^n\rangle$ is true iff $\langle x_1, \dots, x_n \rangle$ instantiates R^n , false otherwise. For example, the proposition that roses are red and violets are blue becomes $\langle\langle$ the proposition that roses are red, the proposition that violets are blue \rangle , CONJ \rangle , and the proposition that there are tigers becomes $\langle\langle$ the property of being a tiger \rangle , SOME \rangle (*vide* Soames 1988).
6. As the Russellian may take $\langle\langle x_1, \dots, x_n \rangle, R^n\rangle$ to represent the form of every proposition, so the Fregean may take it to be represented by $\langle\langle c_1, \dots, c_n \rangle, C^n\rangle$.
7. Cf. Horwich 1998, 122.
8. The Russellian Nathan Salmon seeks to resolve the problem of empty names by arguing in his contribution to this volume (Salmon 2001) that even when a speaker intends her use of a name in a that-clause to refer to a bearer of the name, the occurrence of the name refers to a certain sort of *mythical entity* should it turn out that the intended bearer of the name does not exist.
9. Strong Russellians like Salmon and Soames accept the working hypothesis, but it is possible to hold what I have elsewhere called a *hidden-indexical theory* of belief reports according to which (i) believing is a three-place relation among a believer, the Russellian proposition the believer believes, and the "mode of presentation" under which the believer believes the Russellian proposition and (ii) an utterance of a belief sentence of form (i) (' A believes that S ') requires an implicit contextually determined reference to a mode of presentation or to a type of mode of presentation (see Schiffer 1977, where I first presented and argued for HIT, and Crimmins and Perry 1989, which independently argued for it; and see Schiffer 1992 and 1996a where I argue against HIT). I believe the working hypothesis to be true and eminently defensible. I'm assuming it here so I can manageably focus on the nature of propositions.
10. Read ' $S(t)$ ' as *the sentence S containing the singular term t* .
11. In order to state the argument in a way that brings it directly to bear on the weak Fregean position, 'is' in [*] should be changed to 'may be'.
12. Or 'dog'—it makes no difference, although some rewording would be required.
13. I am grateful to Paul Horwich for pointing out to me the need to qualify an earlier formulation.
14. Owing to Grelling-like paradoxes, I do not mean to imply that every instance of the inference form ' a is F ; so, a has the property of being F ' is valid. The claim that every instance was valid would land us with the property of being a property that does not instantiate itself and thus with the contradiction apparently entailed by 'The property of being a property that doesn't instantiate itself is a property that doesn't instantiate itself', which is false if true and true if false, a lose-lose situation if ever there was one. In Schiffer 1996b I argued that such exceptions to the validity of the something-from-nothing schema in question actually supports the account of pleonastic entities.
15. These again pertain to the semantic paradoxes and also to doubts about the status of the principles of bivalence and excluded middle (see Schiffer 2000). In Schiffer 1996b and forthcoming I argue that the complications introduced by these qualifications further

- bolster the pleonastic aspect of the ontological and conceptual status of pleonastic propositions.
16. The point of the “obvious” qualification will soon be apparent.
 17. I was pleased to see essentially this point made by Richard Holton (2000) when he conceded that the thought that, e.g., Ronald Reagan is retired is essentially about Reagan, but then pointed out that “that does not preclude . . . that, were he to turn out not to exist . . . we would still have a mistaken but contentful thought. . . . It is just that this thought would not have the same content as it has given that he does exist” (13).
 18. The examples, but not the use to which they are put, are borrowed from Mates (1952).
 19. But does not the fact that ‘Claudia believes that she isn’t German’ entail $\exists x$ —viz., Claudia—(Claudia believes that x isn’t German) show that Claudia is the reference of ‘she’? No; if it were we could substitute ‘Claudia Schiffer’ for ‘she’, *salva veritate*, and we have just seen that we cannot. There is no reason to take certain canonical representations in elementary logic as giving a fully accurate representation of ordinary language quantification.

REFERENCES

- Crimmins, M., and J. Perry 1989. “The Prince and the Phone Booth: Reporting Puzzling Beliefs.” *Journal of Philosophy* 86: 685–711.
- Evans, G. 1982. *The Varieties of Reference*. J. McDowell (ed.). Oxford: Oxford University Press.
- Field, H. 1989. *Realism, Mathematics and Modality*. Oxford: Blackwell.
- Holton, R. 2000. “Review of John McDowell.” In *Meaning, Knowledge, Reality. Times Literary Supplement*, June 23.
- Horwich, P. 1998. *Meaning*. Oxford: Oxford University Press.
- Kripke, S. 1979. “A Puzzle about Belief.” In A. Margalit (ed.), *Meaning and Use*. Dordrecht: Reidel.
- Mates, B. 1952. “Synonymity.” In L. Linsky (ed.), *Semantics and the Philosophy of Language*. Urbana: University of Illinois Press.
- Peacocke, C. 1992. *A Study of Concepts*. Cambridge, Mass.: MIT Press.
- Salmon, N. 1986. *Frege’s Puzzle*. Cambridge, Mass.: MIT Press.
- . 1989. “Illogical Belief.” In J. Tomberlin (ed.), *Philosophical Perspectives* 3. Atascadero, Calif.: Ridgeview.
- . 2002. “Mythical Objects.” In J. Keim Campbell, M. O’Rourke, D. Shier (eds.), *Meaning and Truth: Investigations in Philosophical Semantics*. New York: Seven Bridges Press.
- Schiffer, S. 1977. “Naming and Knowing.” In P. French, T. Uehling, and H. Wettstein (eds.), *Midwest Studies in Philosophy* 2. Minneapolis: University of Minnesota Press.
- . 1987a. *Remnants of Meaning*. Cambridge, Mass.: MIT Press.
- . 1987b. “The ‘Fido’-Fido Theory of Belief.” In J. Tomberlin (ed.), *Philosophical Perspectives* 1. Atascadero, Calif.: Ridgeview.
- . 1992. “Belief Ascription.” *Journal of Philosophy* 89: 499–521.
- . 1994. “A Paradox of Meaning.” *Noûs* 28: 279–324.
- . 1996a. “The Hidden-Indexical Theory’s Logical-Form Problem: A Rejoinder.” *Analysis* 56: 92–97.
- . 1996b. “Language-Created Language-Independent Entities.” *Philosophical Topics* 24: 149–67.
- . 2000. “Vagueness and Partial Belief.” *Philosophical Issues* 10: 220–57.
- . (forthcoming). *The Things We Mean*. Oxford: Oxford University Press.
- Soames, S. 1988. “Semantics and Semantic Competence.” In S. Schiffer and S. Steele (eds.), *Cognition and Representation*. Boulder, Colo.: Westview Press.
- Stalnaker, R. 1984. *Inquiry*. Cambridge, Mass.: MIT Press.

PART II
Linguistic Meaning and the World

CHAPTER 6

Mythical Objects

Nathan Salmon

University of California, Santa Barbara

NOTIONAL AND RELATIONAL

IT IS WIDELY recognized that a sentence like

(1) Ralph wants a sloop

is subject to a nonlexical ambiguity not duplicated in, for example,

(2) Ralph owns a sloop.

(1) may indeed be read analogously with (2): There is a sloop that Ralph wants/owns. This is what W. V. Quine (1956) calls the *relational* reading. On this reading, (1) is like (2) in logically entailing the existence of at least one sloop, and the author of (1), like that of (2), is thus ontologically committed to sloops. But (1) may be read instead as indicating an aimless desire on Ralph's part for the very state of affairs described by (2): Ralph's relief from slooplessness. Quine calls this the *notional* reading. Here (1) asserts not that a relation obtains between Ralph and a sloop, but that one obtains between Ralph and the generalized, nonspecific concept of *some sloop or other* (or some counterpart of this concept, like the property of being a class that includes some sloop or other among its elements). No sloop in particular need be the object of Ralph's desire; for that matter, all sloops everywhere may be destroyed. Ralph can still notionally want one. There is no analogous reading for (2). Sloop ownership is as commonplace as sloops. Whatever it would be to stand in the ownership relation to a concept, it is clear that (2) does not attribute such a state to Ralph.

The same asymmetry arises in connection with the following pair:

- (3) Ralph believes a spy has stolen his documents.
- (4) A spy has stolen Ralph's documents.

On its relational reading (3) asserts that there is a spy whom Ralph suspects of having stolen his documents—just as (4) asserts that there is a spy who has indeed taken the missing documents. This is the so-called *de re* reading of (3), what Russell (1905) calls the *primary occurrence* reading.¹ On this reading, some spy is under suspicion, and the speaker is logically committed to there being at least that one spy, in just the same way that the author of (4) is committed to the existence of at least one spy. On its notional reading, (3) reports Ralph's more generalized belief of the very proposition contained in (4): that some spy or other has made off with the documents. No one in particular need be under suspicion. There need not even be any spies anywhere, as long as Ralph believes otherwise. This is the *de dicto* reading, what Russell calls the *secondary occurrence* reading. It asserts a relation not between Ralph and a spy, but one between Ralph and the concepts of *some spy or other* and *stealing documents*. There is no analogous reading for (4). Concepts are not thieves, nor does (4) make any accusation against any concept. Underlying the relational/notional dichotomy in (1) and (3) is the pertinent fact that wanting and believing are psychological states that may be directed equally toward concepts or objects (or concepts that involve objects, or propositions that involve objects, etc.). Ownership and theft are not states of this sort.

Care must be taken not to confuse the notional/relational distinction with various alternative distinctions. One such alternative concerns different uses that a speaker might make of an indefinite descriptive phrase. Though 'a sloop' expresses the indefinite concept *some sloop or other*, there is no bar against using the phrase with reference to a particular sloop (as, for example, in "I was in a sloop yesterday. Was it yours?"). Such a use flies in the face of the indefinite character of the concept semantically expressed by the phrase. We say something nonspecific and mean something specific; in effect, we say "some sloop" but mean "that sloop." And yet life goes on relatively unperturbed. Keith Donnellan (1966) famously pointed out (as did some others independently) that definite descriptions are likewise used sometimes with a particular object in mind ("referential use"), sometimes not ("attributive"). Let us call a use of a definite or indefinite description in uttering a sentence *directed* when there is a particular object to which the use is relevantly connected (e.g., the speaker intends a specific object or person) and the speaker may be regarded as thereby asserting (or asking) something specific directly about that object, and let us call a use of a description *undirected* when the speaker instead merely intends something general to the effect that whatever (whoever) is the only such-and-such/at least one such-and-such or other.²

The distinction between directed and undirected uses is clearly genuine; of that there can be no legitimate doubt. What is subject to serious dispute is

whether the distinction has a direct bearing on the semantics of descriptive phrases. In particular, is (2) literally true, even if only by dumb luck, when ‘a sloop’ is used directedly for a sloop Ralph does not in fact own, if Ralph nevertheless owns a sloop? Intuition strongly favors an affirmative response. Russell recognized the point, and urged it in favor of his theory (now generally taken for granted) that indefinite descriptions function univocally as existential quantifiers:

What do I really assert when I assert “I met a man”? Let us assume, for the moment, that my assertion is true, and that in fact I met Jones. It is clear that what I assert is *not* “I met Jones.” I may say “I met a man, but it was not Jones”; in that case, though I lie, I do not contradict myself, as I should do if when I say I met a man I really mean that I met Jones. It is clear also that the person to whom I am speaking can understand what I say, even if he is a foreigner and has never heard of Jones.

But we may go further: not only Jones, but no actual man, enters into my statement. This becomes obvious when the statement is false, since then there is no more reason why Jones should be supposed to enter into the proposition than why anyone else should. . . . Thus it is only what we may call the concept that enters into the proposition.

More systematic considerations can also be brought to bear, discrediting the thesis that the directed/undirected distinction is relevantly relevant.³ Still some remain unconvinced. Joseph Almog (1998, 77–81) has claimed that the distinction (or one like it in all relevant respects) is not only semantically significant, but indeed provides the basis for the notional/relational distinction.⁴ The relational reading of (1), Almog contends, is generated by a directed use of the relevant indefinite description, the notional reading by an undirected use (1998, 79–81; Almog speaks of “readings” rather than “uses”). The account extends the notional/relational distinction to (2), portraying the undirected use of the indefinite as generating a notional reading. In fact, Almog explains the notional reading of (1) as the exact analog of the reading generated by an undirected use of (2).

The fact that the directed/undirected distinction applies to sentences like (2) and (4), not just (1) and (3), is in itself reason for suspicion of the proposal. Almog’s account gets things exactly reversed with the facts. It is the relational reading of (1), not the notional, that arises by reading it on the model of (2): There is some sloop or other that Ralph owns/wants. A genuinely notional reading of (2) should depict Ralph as somehow standing in the ownership relation to a nonspecific concept! Likewise, it is the relational reading of (3), not the notional, that arises by reading it on the model of (4): There is some spy or other who has stolen Ralph’s documents—or whom Ralph believes has stolen them.⁵

The explanation for the collapse of the notional/relational distinction on

Almog's account is straightforward. Consider the relational reading of (1): A sloop is such that Ralph specifically wants it. Whereas 'a sloop' may be used directedly, there is nothing to prevent the speaker from instead using the indefinite phrase undirectedly, and to mean by (1), understood relationally, that Ralph's desire is focused on some sloop or other: There is a very particular sloop—which sloop is not here specified—that Ralph has his heart set on. (I maintain that this accords with the literal meaning of (1), read relationally, regardless of whether the indefinite is used directedly or undirectedly, whereas the specific thought that Ralph wants *that sloop I have in mind* provides more information than is semantically encoded into the relational reading.) Exactly similarly for (3): There is a very particular spy—which spy is not here specified—to whom Ralph's finger of blame is pointed in a most *de re*, accusatory way. In neither case does an undirected use preclude the relational reading; read relationally, the indefinite may be used either directedly or undirectedly.

Ironically, an undirected use in fact evidently precludes the notional reading. If (1) is read notionally, the description 'a sloop' functions not to *express* the generalized concept of some sloop or other, but to *refer to* it, in order for (1) to express that Ralph stands to this very concept in the specified relation.⁶ Analogously, on the notional reading of (3), the complement clause functions not to express the proposition that some spy or other has stolen Ralph's documents but to refer to the proposition, enabling the sentence to express that Ralph believes it. As Frege noted, in such cases the indefinite phrase does not have its customary content or reference, i.e., its customary *Sinn* or *Bedeutung*. Instead it is in *ungerade* ("oblique") mode. Insofar as the phrase is used to refer to a generalized concept, it is naturally used directedly for that very concept. The notional reading is thus generally accompanied by a directed use by the speaker (albeit an *ungerade* use), not an undirected one. Here again, Almog's account has matters exactly reversed with the facts.⁷

Taking (2) as a model for the notional reading of (1) inevitably yields exactly the wrong results. In effect, Almog attempts to capture the relational/notional distinction by contrasting directed and undirected uses of the relational reading, missing the notional reading altogether. The failure of the directed/undirected distinction as an analysis of the notional/relational is confirmed by Russell's insight that the latter distinction replicates itself in increasingly complex constructions. This is Russell's notion of scope. Thus the sentence

Quine doubts that Ralph wants a sloop

yields not merely two, but three distinct readings: There is a sloop that Quine specifically doubts Ralph wants (*wide scope*); Quine doubts that there is any sloop that Ralph wants (*intermediate*); Quine doubts that Ralph seeks relief from slooplessness (*narrow*). The intermediate-scope reading is notional with respect to Quine and relational with respect to Ralph; the narrow-scope reading is dou-

bly notional. The intermediate- and narrow-scope readings report Quine's doubt of the relational and notional readings, respectively, of (1). The wide-scope reading is the next generation of readings. Prefixing further operators introduces successive generations ("You understand that Salmon reports that Quine doubts . . ."). By contrast, the directed/undirected distinction does not reproduce with operators. The distinction naturally arises in the wide-scope reading, which is neutral between a directed and an undirected use of 'a sloop'. Each is permissible. ("A sloop [*that sloop I have in mind* vs. *some sloop or other*] is such that Quine specifically doubts that Ralph wants *it*.") In both the intermediate- and narrow-scope readings, 'a sloop' is in *ungerade* mode, and hence, insofar as it is used directedly or undirectedly, is presumably directed.⁸

GEACH'S PUZZLE

THE NOTIONAL/RELATIONAL DISTINCTION may be tested by anaphoric links to a descriptive phrase. Consider:

Ralph wants a sloop, but it is a lemon
 Ralph believes a female spy has stolen his documents; she also tampered
 with the computer.

These sentences strongly favor a relational reading. Appropriately understood, each evidently entails the relational reading of its first conjunct, even if the first conjunct itself is (somewhat perversely) read notionally. If, as alleged, it is a lemon, then there must be an *it* that is a lemon, and that *it* must be a sloop that Ralph wants. Similarly, if she tampered with the computer, then there must be a *she* who is a spy and whom Ralph suspects of the theft.

The notional/relational distinction comes under severe strain, however, when confronted with Peter T. Geach's (1967) ingenious Hob/Nob sentence:

- (5) Hob thinks a witch has blighted Bob's mare, and Nob wonders whether she (the same witch) killed Cob's sow.⁹

This puzzling sentence seems to resist both a relational and a notional reading. If there is a *she* whom Nob wonders about, then that *she*, it would appear, must be a witch whom Hob suspects of mare blighting. But the sincere utterer of (5) intuitively does not seem committed in this way to the reality of witches. Barring the existence of witches, though (5) may be true, there is no actual witch about whom Hob suspects and Nob wonders. Any account of the notional/relational that depicts (5) as requiring the existence of a witch is *ipso facto* wrong. There is a natural reading of (5) that carries an ontological commitment to witches, viz., the straightforward relational reading. The point is that the intended reading does not.

A tempting response construes (5) as fully notional, along the lines of

- (5_n) (i) Hob thinks: a witch has blighted Bob's mare; and (ii) Nob wonders whether: the witch that (Hob thinks) blighted Bob's mare also killed Cob's sow.

Yet this will not do; (5) may be neutral concerning whether Nob has a true belief about, let alone shares, Hob's suspicion. Nob's wondering need not take the form "Did the same witch that (Hob thinks) blighted Bob's mare also kill Cob's sow?" It may be that Hob's thought takes the form "Maggoty Meg has blighted Bob's mare" while Nob's takes the form "Did Maggoty Meg kill Cob's sow?" If so, (5) would be true, but no fully notional reading forthcoming.

Worse, Hob's and Nob's thoughts need not involve the same manner of specification. It may be that Hob's thought takes the form "Maggoty Meg has blighted Bob's mare" while Nob's wondering takes the form "Did the Wicked Witch of the West kill Cob's sow?" This appears to preclude a neo-Fregean analysis along the lines of the following:

- (F) $(\exists\alpha)[\alpha$ **corepresents** for both Hob and Nob & Hob **notionally-thinks** $\lceil\alpha$ is a witch who has blighted Bob's mare \rceil & Nob **notionally-thinks** $\lceil\alpha$ is a witch \rceil and Nob **notionally-wonders** \lceil Did α kill Cob's sow? $\rceil]$.¹⁰

Geach himself argues (1967, pp. 148–149) that since (5) does not commit its author to the existence of witches, it must have some purely notional reading or other. He suggests an alternative neo-Fregean analysis, evidently along the lines of the following:

- (G) $(\exists\alpha)(\exists\beta)[\alpha$ is a witch-representation & β is a witch-representation & α and β **corepresent** for both Hob and Nob & Hob **notionally-thinks** $\lceil\alpha$ has blighted Bob's mare \rceil & Nob **notionally-wonders** \lceil Did β kill Cob's sow? $\rceil]$.¹¹

This proposal faces certain serious difficulties, some of which are also problems for (F): The relevant notion of a *witch-representation* must be explained in such a way as to allow that an individual representation α (e.g., an individual concept) may be a witch-representation without representing any actual witch, and for that matter, without representing anything at all. More important, the relevant notion of *corepresentation* needs to be explained so as to allow the following: that a pair of individual representations α and β may co-represent for two thinkers without representing anything at all for either thinker. Geach does not explicitly employ the notion of corepresentation. I have included it on his behalf because it, or something like it, is crucial to the proposed analysis. Any analysis, if it is correct, must capture the idea that Hob's and Nob's thoughts

have a common focus. Though there is no witch, Hob and Nob are, in some sense, thinking about the *same* witch. It is on this point that notional analyses generally fail. Even something as strong as (ζ_n)—already too strong—misses this essential feature of (ζ). On the other hand, however the notion of vacuously corepresenting witch-representations is ultimately explained, by contrast with (G), (ζ) evidently commits its author no more to corepresenting witch-representations than to witches. More generally, any analysis along the lines of (F) or (G) cannot forever avoid facing the well-known difficulties with neo-Fregean, notional analyses of relational constructions generally (e.g., the Twin Earth considerations).¹²

An alternative approach accepts the imposingly apparent relational character of (ζ) at face value, and construes it along the lines of the following:

- (6) There is someone whom: (*i*) Hob thinks a witch that has blighted Bob's mare; (*ii*) Nob also thinks is a witch; and (*iii*) Nob wonders whether she killed Cob's sow.

This happily avoids commitment to witches. But it does not provide a solution. Hob's and Nob's thoughts need not concern any real person. Maggoty Meg is not a real person, and there may be no one whom either Hob or Nob believe to be the wicked strega herself.

Some proposed solutions to Geach's puzzle make the unpalatable claim that Hob's and Nob's musings concern a Meinongian Object—a particular witch who is both indeterminate and nonexistent.¹³ Many proposed solutions instead reinterpret relational attributions of attitude so that they are not really relational, i.e., they do not make genuine reference to the individuals apparently mentioned therein by name or pronoun. These responses inevitably make equally unpalatable claims involving relational constructions—for example, that Nob's wondering literally concerns the very same witch/person as Hob's belief yet neither concerns anyone (or anything) whatsoever, or that relational constructions mention or generalize over speech-act tokens and/or connections among speech-act tokens.¹⁴ It would be more sensible to deny that (ζ) can be literally true on the relevant reading, given that there are no actual witches.¹⁵ The problem with this denial is that its proponent is clearly in denial. As intended, (ζ) can clearly be true (assuming Hob and Nob are real) even in the absence of witches. Numerous postmodern solutions jump through technical hoops to allow a pronoun ('she') to be a variable bound by a quantifier within a belief context ('a witch') despite standing outside the belief context, hence also outside the quantifier's scope, and despite standing within an entirely separate belief context. These "solutions" do not satisfy the inquiring mind as much as boggle it. It is one thing to construct an elaborate system on which (ζ) may be deemed true without 'There is a witch'. It is quite another to provide a satisfying explanation of the content of Nob's attitude, one for which the constructed system is

appropriate. How can Nob wonder about a witch, and a particular witch at that—the very one Hob suspects—when there is no witch and, therefore, no particular witch about whom he is wondering? This is the puzzle in a nutshell. It combines elements of intensionality puzzles with puzzles concerning non-existence and puzzles concerning identity, and has been deemed likely intractable.¹⁶

MYTHS

THE SOLUTION I SHALL URGE takes (5) at face value, and takes seriously the idea that false theories that have been mistakenly believed—what I call *myths*—give rise to fabricated but genuine entities.¹⁷ These entities include such oddities as: Vulcan, the hypothetical planet proposed by Babinet and which Le Verrier believed caused perturbations in Mercury's solar orbit; the ether, once thought to be the physical medium through which light waves propagate; phlogiston, once thought to be the element (material substance) that causes combustion; the Loch Ness Monster; Santa Claus; and Meinong's Golden Mountain. Such *mythical objects* are real things, though they are neither material objects nor mental objects ("ideas"). They come into being with the belief in the myth. Indeed, they are created by the mistaken theory's inventor, albeit without the theorist's knowledge. But they do not exist in physical space, and are, in that sense, abstract entities. They are an unavoidable by-product of human fallibility.

Vulcan is a mythical planet. This is not to say, as one might be tempted to take it, that Vulcan is a planet but one of a rather funny sort, e.g., a Meinongian Object that exists in myth but not in reality.¹⁸ On the contrary, Vulcan exists in reality, just as robustly as you the reader. But a mythical planet is no more a planet than a toy duck is a duck or a magician is someone who performs feats of magic. A mythical object is an imposter, a pretender, a stage prop. Vulcan is not a real planet, though it is a very real object—not concrete, not in physical space, but real. One might say that the planet Mercury is also a "mythical object," in that it too figures in the Vulcan myth, wrongly depicted as being gravitationally influenced by Vulcan. If we choose to speak this way, then it must be said that some "mythical planets" are real planets, though not really as depicted in the myth. Vulcan, by contrast with the "mythical" Mercury, is a *wholly mythical* object, not a real planet but an abstract entity inadvertently fabricated by the inventor of the myth. I shall continue to use the simple word 'mythical' as a shorthand for the notion of something wholly mythical.¹⁹

The existence of fictional objects, in something close to this sense, has been persuasively urged by Peter van Inwagen (1977) and Saul Kripke (1973) as an ontological commitment of our ordinary discourse about fiction.²⁰ Their account, however, is significantly different from the one I propose. Kripke contends that

a mythical-object name like ‘Vulcan’ is ambiguous between two uses, one of which is parasitic on the other. It would be less deceptive to replace the ambiguous name with two univocal names, ‘Vulcan₁’ and ‘Vulcan₂’. The name on its primary use, ‘Vulcan₁’, was introduced into the language, *sans* subscript, by Babinet as a name for an intra-Mercurial planet. Le Verrier used the name in this way in theorizing about Mercury’s perihelion. In this use, the name names nothing; ‘Vulcan₁’ is entirely vacuous. Giving the name this use, we may say such things as that Le Verrier believed that Vulcan₁ affected Mercury’s perihelion. Le Verrier’s theory is a myth concerning Vulcan₁. The name on its secondary use, ‘Vulcan₂’, is introduced into the language (again *sans* subscript) at a later stage, when the myth has finally been exposed, as a name for the mythical planet erroneously postulated, and thereby inadvertently created, by Babinet. Perhaps it would be better to say that a new *use* of the name ‘Vulcan’ is introduced into the language. ‘Vulcan₂’ is fully referential. Using the name in this way, we say such things as that Vulcan₂ was a mythical intra-Mercurial planet hypothesized by Babinet. The difference between Vulcan₁ and Vulcan₂ could not be more stark. The mistaken astronomical theory believed by Babinet and Le Verrier concerns Vulcan₁, which does not exist. Vulcan₂, which does exist, arises from the mistaken theory itself. Vulcan₂ is recognized through reflection not on events in the far-off astronomical heavens but on the more local story of man’s intellectual triumphs and defeats, particularly on the history of science.

Kripke’s account is vulnerable to a familiar family of thorny problems: the classical problem of true negative existentials and the more general problem of the content and truth value of sentences involving vacuous names. Vulcan₁ does not exist. This sentence is true, and seems to say about something (*viz.*, Vulcan₁) that it fails to exist. Yet the sentence entails that there is nothing for it to attribute nonexistence to. Furthermore, on Kripke’s account, Le Verrier believed that Vulcan₁ has an impact on Mercury’s perihelion. What can the content of Le Verrier’s belief be if there is no such thing as Vulcan₁? Furthermore, is the belief content simply false? If so, then it may be said that Vulcan₁ has no impact on Mercury’s perihelion. Yet this claim too seems to attribute something to Vulcan₁, and thus seems equally wrong, and for exactly the same reason, with the claim that Vulcan₁ does have such an impact. Kripke is aware of these problems but offers no viable solution.

I submit that Kripke’s alleged primary use of a mythical-object name is itself a myth. To be sure, Babinet believed himself to be naming a real planet in introducing a use of ‘Vulcan’ into the language, and other users like Le Verrier believed themselves to be referring to a real planet. But this linguistic theory of the name ‘Vulcan’ is mistaken, and is in this respect exactly like the astronomical theory that Vulcan is a real planet. The two theories complement each other, and fall together hand in hand. The situation should be viewed instead as fol-

lows: Babinet invented the theory—erroneous, as it turns out—that there is an intra-Mercurial planet. In doing this, he inadvertently created Vulcan. Indeed, Babinet even introduced a name for this mythical planet. The name was intended for a real planet, and Babinet believed the name thus referred to a real planet (notionally, not relationally!). But here again, he was simply mistaken. Other astronomers, most notably Le Verrier, became convinced of Babinet’s theory, both as it concerns Vulcan (that it is a very real intra-Mercurial planet) and as it concerns ‘Vulcan’ (that it names the intra-Mercurial planet). Babinet and Le Verrier both believed, correctly, that the name ‘Vulcan’, on the relevant use, refers to Vulcan. But they also both believed, mistakenly, that Vulcan is a real planet. They might have expressed the latter belief by means of the French version of the English sentence ‘Vulcan is a planet’, or other shared beliefs by means of sentences like ‘Vulcan’s orbit lies closer to the Sun than Mercury’s’. These beliefs are mistakes, and the sentences (whether English or French) are false.

Importantly, there is no relevant use of the name ‘Vulcan’ by Babinet and Le Verrier that is vacuous. So used, the name refers to Vulcan, the mythical planet. Le Verrier did *not* believe that Vulcan₁ is an intra-Mercurial planet—or, to put the point less misleadingly, there is no real use marked by the subscript on ‘Vulcan’ on which the string of words ‘Vulcan₁ is an intra-Mercurial planet’ expresses anything for Le Verrier to have believed, disbelieved, or suspended judgment about. To put the matter in terms of Kripke’s account, what Le Verrier believed was that Vulcan₂ is a real intra-Mercurial planet. Le Verrier’s belief concerns the mythical planet, a very real object that had been inadvertently created, then named ‘Vulcan’, by Babinet. Their theory about Vulcan was completely wrong. Vulcan is in fact an abstract object, one that is depicted in myth as a massive physical object.

A common reaction is to charge my proposal with miscasting mythical objects as the objects with which myths are concerned. On the contrary, it is objected, if they exist at all, mythical objects enter the intellectual landscape only at a later stage, not in the myth itself but in the subsequent historical account of the myth. A robust sense of reality demands that the myth itself be not about these abstract objects but about *nothing*, or at most about representations of nothing. No one expresses this sentiment more forcefully than Russell (1919):

[Many] logicians have been driven to the conclusion that there are unreal objects . . . In such theories, it seems to me, there is a failure of that feeling for reality which ought to be preserved even in the most abstract studies. Logic, I should maintain, must no more admit a unicorn than zoology can; for logic is concerned with the real world just as truly as zoology, though with its more abstract and general features. To say that unicorns have an existence in heraldry, or in literature, or in imagination, is a most pitiful and paltry evasion. What exists in heraldry is not an animal, made of flesh and blood, moving and breathing of its own initiative. What exists is a picture, or a description in words. . . . A

robust sense of reality is very necessary in framing a correct analysis of propositions about unicorns . . . and other such pseudo-objects.²¹

I heartily applaud Russell's eloquent plea for philosophical sobriety. But his attitude toward "unreal" objects is fundamentally confused. To repeat, a mythical planet is not a massive physical object but an abstract entity, the product of creative astronomizing. Likewise, a mythical unicorn or a mythical winged horse is not a living creature but a fabricated entity, the likely product of blurred or fuzzy vision, just as mermaids are the likely product of a deprived and overactive imagination under the influence of liquor—creatures not really made of flesh and blood and fur or scales, not really moving and breathing of their own initiative, but depicted as such in myth, legend, hallucination, or drunken stupor.

It is frequently objected even by those who countenance mythical objects that the Vulcan theory, for example, is merely the theory that there is an intra-Mercurial planet, not the bizarre hypothesis that the relevant abstract entity is that planet. Babinet and Le Verrier, it is observed, did not believe that an abstract entity is a massive heavenly object. Quite right, but only if meant notionally. Understood relationally—as the claim that, even if there is such an abstract entity as the mythical object that is Vulcan, Babinet and Le Verrier did not believe it to be an intra-Mercurial planet—it turns mythical objects into a philosophical black box. What role are these abstract entities supposed to play, and how exactly are their myth-believers supposed to be related to them in virtue of believing the myth? In fact, this issue provides yet another reason to prefer my account over Kripke's. On my account, in sharp contrast, the role of mythical objects is straightforward: They are the things depicted as such-and-such in myth, the fabrications erroneously believed by wayward believers to be planets or the medium of light-wave propagation or ghosts, the objects the mistaken theory is about when the theory is not about any real planet or any real medium or any real ghost. It is not merely that being depicted as such-and-such is an essential property of a mythical object, a feature the object could not exist without. Rather, being so depicted is the metaphysical function of the mythical object; that is *what* it is, its *raison d'être*. To countenance the existence of Vulcan as a mythical planet while at the same time denying that Babinet and Le Verrier had beliefs about this mythical object, is in a very real sense to miss the point of recognizing Vulcan's existence. It is precisely the astronomers' false beliefs about the mythical planet that makes it a mythical planet; if no one had believed it to be a planet, it would not *be* a mythical planet. Come to that, it would not even exist.²²

Another important point: I am not *postulating* mythical objects. For example, I am not postulating Vulcan. Even if I wanted to, Babinet beat me to it—though he postulated Vulcan as a real planet, not a mythical one.²³ Mythical objects would exist even if I and everyone else had never countenanced or recognized them, or admitted them into our ontology. Rather, I see myself as uncovering some evidence for their independent and continued existence, in

something like the manner of the paleontologist who infers dinosaurs from their fossil remains, rather than the theoretical physicist who postulates a new category of physical entity in order to make better sense of things (even if what I am actually doing is in important respects more like the latter).²⁴

Perhaps the most important evidence in favor of this theory of mythical objects is its logical entailment by our thoughts and beliefs concerning myths. We are sometimes led to say and think such things as, “An intra-Mercurial planet, Vulcan, was hypothesized by Babinet and believed by Le Verrier to affect Mercury’s perihelion, but there has never been a hypothetical planet whose orbit was supposed to lie between Mercury and Venus” and “Some hypothetical species have been hypothesized as linking the evolution of birds from dinosaurs, but no hypothetical species have been postulated to link the evolution of mammals from birds.” The distinctions drawn cannot be made without a commitment to mythical objects, i.e., without attributing existence, in some manner, to mythical objects. No less significant, beliefs are imputed about the mentioned mythical objects, to the effect that they are not mythical. Being wrongly believed not to be mythical is just what it is to be mythical. Furthermore, beliefs are imputed to distinct believers concerning the very same mythical object.²⁵

Further evidence—in fact, evidence of precisely the same sort—is provided by the Hob/Nob sentence. The puzzle is solved by construing (5) on its principal reading, or at least in one of its principal readings, as fully relational, not in the manner of (6) but along the lines of:

- (7) There is a mythical witch such that (i) Hob thinks: she has blighted Bob’s mare; and (ii) Nob wonders whether: she killed Cob’s sow.²⁶

This has the distinct advantage over (6) that it does not require that both Hob and Nob believe someone to be the witch in question. In fact, it allows that there be no one in particular whom either Hob or Nob believes to be a witch. It does require something not unrelated to this, but no more than is actually required by (5): that there be something that both Hob and Nob believe to be a witch—*something*, not *someone*, not a witch or a person, certainly not an indeterminate Meinongian Object, but a very real entity that Nob thinks a real witch who has blighted Bob’s mare. Nob also believes this same mythical witch to be a real witch and wonders about “her” (really: about *it*) whether she killed Cob’s sow. In effect, the proposal substitutes ontological commitment to mythical witches for the ontological commitment to real witches intrinsic to the straightforward relational reading of (5) (obtained from (7) by deleting the word ‘mythical’). There are other witch-free readings for (5), but I submit that any intended reading is a variant of (7) that equally commits the author to the existence of a (real or) mythical witch, such as:

- (i) Hob thinks: some witch or other has blighted Bob’s mare; and (ii) the (same) mythical witch that Hob thinks has blighted Bob’s mare is such that Nob wonders whether: she killed Cob’s sow.²⁷

Significantly, one who accepts Kripke's account may not avail him/herself of this solution to Geach's puzzle. On Kripke's account it may be observed that

- (i) Hob thinks: Meg₁ has blighted Bob's mare; and (ii) Nob wonders whether: Meg₁ killed Cob's sow.

The Hob/Nob sentence (5) is not obtainable by existential generalization on 'Meg₁', since by Kripke's lights, this name is supposed to be vacuous and to occur in nonextensional ("referentially opaque," *ungerade*) position. Nor on Kripke's account can 'Meg₂' be correctly substituted for 'Meg₁'; Hob's and Nob's theories are supposed to concern the nonexistent witch Meg₁ and not the mythical witch Meg₂. Kripke might instead accept the following, as a later-stage observation about the Meg₁ theory:

Meg₂ is the mythical witch corresponding to Meg₁.

Here the relevant notion of *correspondence* places 'Meg₂' in extensional position. While 'Meg₂' is thus open to existential generalization, 'Meg₁' supposedly remains in a nonextensional position where it is not subject to quantification. It is impossible to deduce (5) from any of this. Geach's puzzle does not support Kripke's account. On the contrary, the puzzle poses a serious threat to that account, with its denial that Hob's and Nob's thoughts are, respectively, a suspicion and a wondering regarding Meg₂.

On my alternative account, we may instead observe that

Maggoty Meg is a mythical witch. Hob thinks she has blighted Bob's mare.
Nob wonders whether she killed Cob's sow.

We may then conjoin and EG to obtain (7). In the end, what makes (7) a plausible analysis is that it (or some variant) spells out in more precise language what (5) literally says to begin with. Babinet and Le Verrier provide a real-life case in which the thoughts of different thinkers converge on a single mythical object: Babinet thought he had seen an intra-Mercurial planet, and Le Verrier believed that it (the same "planet") impacted Mercury's perihelion. The primary lesson of Geach's puzzle is that when theoretical mistakes are made mythical creatures are conceived, and in acknowledging that misbelievers are sometimes related as Nob to Hob, or as Le Verrier to Babinet, we commit ourselves to their illegitimate progeny.²⁸

NOTES

This paper was presented at various venues before and after the turn of the millenium. I am grateful to Mark Fiocco, Steven Humphry, Genoveva Marti, Michael McGlone, and Teresa Robertson for discussion, as well as my audiences and the participants in my UCSB seminar during Spring 2000.

1. Russell would extend his primary/secondary occurrence distinction to (1) by rewriting it in sentential-operator form, for example, as ¹Ralph desires that (2)¹.
2. Donnellan's referential/attributive distinction for definite descriptions is a special case of the directed/undirected distinction, which also covers indefinite descriptions. A use of 'some atheist' in uttering 'Some atheist is a spy' may be undirected even if the speaker is regarded as thereby designating a higher-order entity relevantly connected to that same use (for example, the function from functions-from-individuals-to-truth-values that assigns *truth* to any function assigning *truth* to at least one atheist and otherwise assigns *falsity*).
3. See Kripke 1977.
4. Others who also maintain that the directed/undirected distinction is semantically relevant include Barbara Partee (1972, Almog follows Partee in confusing relational/notional with directed/undirected), Jon Barwise and John Perry (1983), and Howard Wettstein (1981, 1983). I challenge Wettstein's account in Salmon (1991).
5. Almog explains the notional reading of "Madonna seeks a man" (misidentified with its undirected use) by saying that it is true if and only if Madonna seeks at least one instance of the kind *Man* (1998, 57, 80). This is at best a tortured expression of Madonna's objective ("Mankind, schmankind. I'm just looking for a man."). Worse, the formulation leaves the notional/relational ambiguity unresolved. In seeking at least one instance of mankind, is there anyone in particular who is the object of Madonna's desire, or is she merely seeking relief from her unbearable loneliness? Almog disambiguates in exactly the wrong direction, saying: (i) "Madonna met a man" . . . is true on this parsing [its undirected use] . . . iff Madonna met at least one man"; furthermore, (ii) "no special treatment accrues intensional verbs. Thus to get the truth conditions of the [notional] reading of 'Madonna seeks a man', simply substitute 'seek' [in (i)]" (1998, 80). Substitution of *seeking* for *having met* in Madonna's having met at least one man (or in Madonna's standing to mankind in the relative product, *x* met at least one instance of *y*) directly results in a targeted search by the diva.

- Almog denies (1998, 53–54) that (2) logically entails (2') 'There is a sloop that Ralph owns', on the grounds that (1), which has the same logical form as (2), can be true without (1') 'There is a sloop that Ralph wants'—while conceding that it is nevertheless necessary and knowable *a priori* that if Ralph owns a sloop then there is a sloop that he owns. This argument carries no conviction. Logic can no more tolerate a divergence in truth value between 'Ralph owns at least one sloop' and 'At least one sloop is such that Ralph owns it' than it can between 'The number of planets is such as to be not even' and 'It is not the case that: the number of planets is even'. The second pair are equivalent despite the fact that substitution of 'possibly' for 'not' yields a falsehood and a truth, respectively. There is a reading of (1) on which it evidently entails (1')—viz., the relational reading. In any event, on this reading (1) yields (1') with the same sort of modality as between (2) and (2')—whether the connection is deemed logical or only necessary, *a priori*, intuitive, conceptual, true by virtue of meaning, and whatever else (knowable by reason alone?). The relational/notional distinction may even be defined or characterized by contrasting the reading of (1) on which it yields (1') *via* the same sort of modality as between (2) and (2'), with that on which it instead attributes a desire for slooplessness relief compatible with (1')'s denial. Owning and finding provide a template for wanting and seeking, but only for wanting and seeking in the relational senses. The desire for mere relief from slooplessness provides a new paradigm (familiarity of grammatical form notwithstanding).
6. This is not to say that Ralph wants to own a concept. There is no sloop or concept that Ralph specifically wants in virtue of wanting relief from slooplessness. Rather, Ralph stands in a certain relation to the generalized concept, *some sloop or other*. The relation is expressed in some English constructions by 'wants'. To say that Ralph notionally wants a concept is to assert that this same relation obtains between Ralph and a concept of a concept. Cf. Alonzo Church (1956, 8n20).
 7. Almog depicts (2) on an undirected use as expressing (or at least as true exactly on the condition) that Ralph stands to the kind *Sloop* in the relative product, *x* owns at least one instance of *y*. This would suggest that, in such a use, the word 'sloop' refers to, and is directed toward, the kind *Sloop* while the words 'owns a' express the relative product (1998,

79). Similarly for the analogous use of (1), yielding its relational reading (directly contrary to Almog's stated purpose; see note 5 above). The phrase 'a sloop' (as opposed to the word 'sloop' occurring therein) on such a use would refer neither to the kind nor to the relative product, nor to anything else. In effect, it is contextually defined away. (Alternatively, it might be taken as referring to a higher-order entity, e.g., $(\lambda F)[(\exists z)(z \text{ is an instance of the kind } Sloop \ \& \ Fz)]$; cf. note 2 above. But Almog eschews such entities in his semantic analysis.) By contrast, 'a sloop' on the notional reading of (1) refers to, and its use is directed toward, the concept, *some sloop or other* (or if one prefers, *at least one instance of Sloop*).

8. The various considerations demonstrating the failure of the directed/undirected analysis of relational/notional are well known in connection with definite descriptions. Cf. Kripke (1977, 9–10). Analogous considerations are at least as forceful with regard to indefinite descriptions. In responding to Kripke's arguments against the alleged semantic significance of the directed/undirected distinction, Almog (1998, 91–98) barely acknowledges these more decisive—and more fundamental—considerations against his proposal. Almog's defense of the semantic-significance thesis suffers furthermore from the confusions limned above, including, for example, the false premise that the notional reading of (1) asserts that one sloop or other has the property of being wanted by Ralph (something in fact entailed by the relational reading). Michael McGlone has pointed out (in conversation) that Almog might restrict his directed/undirected account of relational/notional to constructions like (1), not extending it to (3). (Cf. Almog 1998, 104n20.) Such a restriction would be both *ad hoc* and irrelevant. (The scope considerations apply equally to 'Diogenes wants to seek an honest man'.) The account fails for both sorts of cases, and for the same basic reason: The analogue for (1)/(3) of an undirected use of (2)/(4) is a straightforwardly relational reading, and hence fails as an analysis of the notional reading.

Perhaps Almog will recant and concede that verbs like 'want' and 'seek' do after all require special treatment to capture the elusive notional readings. On its notional reading, (1) is true iff Ralph is related to the kind *Sloop* by notionally wanting at least one instance of the latter, as opposed to relationally wanting one, as entailed by the discredited account. (See notes 5 and 7 above.) This of itself leaves the former condition unexplained. In particular, appealing to an alleged undirected use of 'a sloop' by the reporter yields the wrong reading. But Almog also explicitly rejects the Frege-inspired analysis (which I believe is essentially correct): that certain expressions including 'seek' and 'want' (not including 'find' and 'own') are *ungerade* operators, which induce 'a sloop' to refer to rather than to express the concept *some sloop or other*, eliciting a directed use by the speaker. (The relational reading of (1) is explicable on this analysis as a matter of wide scope/primary occurrence.)

A case can be made that the relational reading of (1) goes hand in hand with a directed use of 'a sloop', or a propensity toward a directed use, *on the part of Ralph* rather than the speaker, and the notional reading correspondingly with an undirected use, or a propensity thereto, by Ralph. A logico-semantic account of relational/notional along these lines, although not as conspicuously flawed as Almog's, is also significantly wide of the mark. (Suppose Ralph speaks no English. Consider also the Church-Langford translation test.) Almog anyway explicitly rejects the idea (1998, 56).

9. Though the puzzle has generated a considerable literature, its general importance to the philosophy of logic and language remains insufficiently appreciated. (As will emerge, I believe Geach's moniker for the puzzle as one of "intentional identity" is a likely misnomer.)
10. Cf. David Kaplan (1969, 225–31). Contrary to Daniel C. Dennett (1968), the intelligibility (indeed the fact) of Hob's and Nob's thoughts having a common focus, somehow on the same unreal witch, does not require that they agree on every possible issue regarding the witch in question—which would in any case entail their agreeing on every possible issue.
11. Geach 1976, 314–18.
12. Stephen Neale (1990, 221), proposes analyzing the relevant reading of (5) along the lines of: (i) Hob thinks: a witch has blighted Bob's mare; and (ii) Nob wonders whether: the *such-and-such* witch killed Cob's sow, where 'the *such-and-such* witch' is fleshed out by the

context, e.g., as ‘the local witch’. But (5) evidently does not attribute to Nob the particular thought ‘Did *the local witch* kill Cob’s sow?’ nor any similarly descriptive thought. Worse, Neale’s proposal fails to capture the crucial feature of (5) that Nob’s wondering allegedly regards the very witch that Hob suspects. Michael McKinsey (1986) argues that the only readings of (5) that do not commit its author to the existence of a witch (or to there being some real person whom Hob and Nob relationally suspect of witchcraft) are given by (ζ_n) (which he regards as ambiguous). Dennett (1968) apparently holds that the only such readings of (5) are either those given by (ζ_n) or else something similar to the less specific (F). *Pace* Geach, Dennett, McKinsey, and Neale, (5) is evidently relational yet free of commitment to witches (or to anyone who is a suspect). (Contrary to Dennett, the speaker’s basis or justification for uttering (5) is mostly irrelevant.)

13. Cf. Esa Saarinen (1978). A variant of this approach imputes thoughts to Hob and Nob concerning a particular possible and fully determinate but nonexistent witch. This proposal cannot be summarily dismissed on the ground of an alleged ontological commitment to merely possibles. The proposed analysis may be understood instead as follows: There *might have existed* (even if there does not exist) a witch such that *actually*: (i) Hob thinks she has blighted Bob’s mare; and (ii) Nob wonders whether she killed Cob’s sow. Whereas this is in some sense committed to merely possible witches, it avoids commitment to their actual existence. The more serious difficulty is that neither Hob nor Nob (assuming they are real) is connected to any particular possible witch, to the exclusion of other possible witches, in such a manner as to have relational thoughts about her. How could they be? Witches do not exist. Cf. Kripke (1972, 158) “. . . one cannot say of any possible person that he *would have been* Sherlock Holmes, had he existed. Several distinct possible people, and even actual ones such as Darwin or Jack the Ripper, might have performed the exploits of Holmes, but there is none of whom we can say that he would have *been* Holmes had he performed these exploits. For if so, which one?”
14. The Hob/Nob sentence (5) is logically consistent with neither Hob nor Nob articulating his musings, explicitly or implicitly. Tyler Burge’s (1983, 94–98) analysis seems to be roughly the following:

Hob **believes** $(\exists x)(x \text{ is a witch who has blighted Bob's mare})$ & \therefore Hob **believes**
 $(\text{the}_{13} x)(x \text{ is a witch who has blighted Bob's mare exists})$ & Nob **wonders** $\{y_{13} \text{ killed}$
 Cob's sow $\}$

Burge stipulates that the recurring subscript “marks the anaphoric or quasi-anaphoric connection between the terms” (1983, 97), where “a more explicit way of capturing the point of the subscripts” would explicitly generalize over communication chains, including both Hob’s application of ‘the₁₃’ and Nob’s application of ‘y₁₃’ (1983, 98).

Burge’s apparatus is not explained sufficiently for this to qualify as a proposed solution to the problem. Aside from questions raised by the connective adjoining the first two conjuncts (how does a single statement contain an argument?), the analysis is inadequate on its most natural interpretations. An immediate problem is that (5), as intended, does not entail that Hob notionally thinks only one witch has blighted Bob’s mare; the argument of the first two conjuncts is invalid. More problematic, if the special quotation marks indicate ordinary quotation (as seems to conform with Burge’s intended interpretation), the analysis miscasts relational constructions as reporting dispositions toward sentences (e.g., purported utterances or implicit utterances) rather than the content of the attitudes thereby expressed and their relation to objects. Assuming instead (apparently contrary to Burge’s intent) that the occurrence of ‘y₁₃’ is in bindable position, the variable remains free even assuming that the definite-descriptions operator ‘the₁₃’ is variable binding. Burge’s stipulation suggests the variable is to have a value assigned to it *via* Hob’s alleged description ‘the witch who has blighted Bob’s mare’, thus recasting the third conjunct into ‘Nob wonders whether she—the witch who has blighted Bob’s mare—killed Cob’s sow’. (Otherwise, the ‘y₁₃’ evidently remains both free and value-less, leaving (5) without propositional content, hence untrue.) This, however, is evidently ambiguous between a reading on which the value-fixing is affected on the part of the author of (5)—call it *primary occurrence*—and a *secondary-occurrence* reading on which the value-fixing is allegedly affected on the part of Nob. (The terminology is intended to recall Russell’s

distinction. The ambiguity corresponds even more closely to two competing interpretations of David Kaplan's rigidifying operator '*dthat*'.) On the secondary-occurrence reading, the value-fixing description plays a representational role on Nob's behalf. On the primary-occurrence reading, the value-fixing is shielded from the shift-from-customary-mode function of the quotation marks, leaving the pronoun to carry the weight of representing for Nob. The analysis on the secondary-occurrence reading, like (5_n), commits not only Hob but also Nob to the existence of a witch who has blighted Bob's mare. Worse, the more likely primary-occurrence reading commits (5)'s author to the existence of such a witch. Neither is correct.

A further problem with the proposal is that the truth of (5) does not require that Nob make any pronominal application that is anaphoric on an application by Hob. The two might never communicate. Burge therefore offers something like the following as an alternative analysis (1983, 96):

The community **believes** $(\exists x)(x \text{ is a witch wreaking havoc})$ & ∴ the community **believes** $(\text{the}_{13} x)(x \text{ is a witch who is wreaking havoc exists})$ & Hob **thinks** $(y_{13} \text{ has blighted Bob's mare})$ & Nob **wonders** $(z_{13} \text{ killed Cob's sow})$

This is subject to some of the same difficulties as the previous analysis and more besides, including some of the same defects as Neale's proposal (see note 12)—as well as some of the defects of the Fregean analyses that Burge eschews. By contrast, for example, (5) makes no claim regarding community-held beliefs, let alone regarding a specific alleged community belief that there is only one witch wreaking havoc.

15. The account in Almog (1998, 68, 75–76, and *passim*), extended to propositional-attitude attributions, apparently depicts (5) as modally equivalent on its intended reading to 'Hob thinks Maggoty Meg has blighted Bob's mare, and Nob wonders whether she killed Cob's sow', and depicts the latter as expressing a necessary falsehood in virtue of the failure of 'Maggoty Meg' to refer.
16. Michael Clark (1975, 124).
17. Cf. Salmon 1998, 304–5; especially 317n50.
18. Geach (1967b) misconstrues the claim in just this way.
19. Sachin Pai asks whether there is in addition to Mercury a wholly mythical planet that astronomers like Le Verrier wrongly believed to be Mercury. I leave this as a topic requiring further investigation.
20. Kripke does not himself officially either accept or reject an ontology of mythical objects. My interpretation is based partly on notes I took at Kripke's seminars on the topic of reference and fiction at Princeton University during March–April 1981 and on recordings of his seminars at the University of California, Riverside, in January 1983. Kripke's account of fictional and mythical objects is explicated and criticized, and my alternative theory defended, in Salmon 1998, 293–305.
21. Bertrand Russell, *Introduction to Mathematical Philosophy*, chap. 16, 169–70.
22. Mythical objects are of the same metaphysical/ontological category as fictional characters, and it is an essential property of any such entity that it be of this category. Perhaps a mythical object might instead have been a fictional character, or vice versa, but no mythical or fictional object could have been, say, an even integer. Some philosophers who accept the reality of fictional characters nevertheless reject mythical objects. The usual motivation is the feeling that whereas Sherlock Holmes is a real object, a character created by Sir Arthur Conan Doyle, the Vulcan theory was wrong precisely because Vulcan simply does not exist. This ignores the nearly perfect similarity between fiction and myth. Whatever good reason there is for acknowledging the real existence of Holmes extends to Vulcan. The Vulcan theory is wrong not because there is no such thing as Vulcan, but because there is no such *planet* as Vulcan as it is depicted. Or better put, Vulcan is no such planet. (Likewise, there was no such detective as Holmes, who is a fictional detective and not a real one.) Myths and fictions are both made up. The principal difference between mythical and fictional objects is that the myth is believed while the fiction is only make-believe. This difference does nothing to obliterate the reality of either fictional or mythical objects.
23. Cf. Salmon 1998, 315n38.

24. I am aware some philosophers see no significant difference between the paleontologist and the theoretical physicist. But they are asleep, or blind.
25. Linguistic evidence tends to support the general claim that if someone believes there is an F that is such-and-such when in fact there is no such thing, then there is a mythical F thereby believed to be such-and-such. It does not follow that whenever someone notionally believes an F is such-and-such, there is always something or someone (either an F or a mythical F) relationally believed to be such-and-such. That the latter is false is demonstrated by the believer who notionally believes some spy is shorter than all others. (Thanks to James Pryor and Robert Stalnaker for pressing me on this point.) If two believers notionally believe there is an F that is such-and-such when in fact there is no such thing, they may or may not believe in the same mythical F , depending on their interconnections. (This may help explain why it is more difficult to form beliefs about the shortest spy than about a mythical planet: Le Verrier and we are all *de re* connected to Vulcan.)

Mark Richard (1998, 262–64, 278–79/16) criticizes my account of mythical objects while defending a version of Kripke's. Richard objects (1998, 279) to the examples given here on the ground that, for example, the first quoted sentence is in fact untrue and is easily confused with a true variant that avoids attributing to Babinet and Le Verrier any ontological commitment to, or beliefs concerning, the mythical planet: 'It was hypothesized by Babinet that there is an intra-Mercurial planet, Vulcan₁, and it was believed by Le Verrier that Vulcan₁ affects Mercury's perihelion, but it has never been hypothesized that there is a planet whose orbit lies between Mercury and Venus'. (Richard denies, with Kripke, that Babinet and Le Verrier have beliefs concerning Vulcan₁.) Richard explains the alleged confusion as the product of an exportation inference from ' α believes that β is an F that is G ' to ' β is an F that α believes is a G ', where β is a proper name. Richard says this inference pattern is valid if, but only if, the name β , as used by the referent of α (e.g., 'Vulcan' as used by Babinet and Le Verrier), has a referent. This explanation is dubious. For one thing, the particular exportation-inference pattern is invalid regardless of the logico-grammatical status of β . Moreover, it does not yield the quoted sentence. As will be seen shortly, Geach's puzzle demonstrates that Richard's substitute sentence does not do justice to the data. Babinet's and Le Verrier's beliefs concern *something*; indeed they each concern the *same* thing.

26. Quasi-formally:

$$(\exists x)(x \text{ is a mythical-witch} \ \& \ \text{Hob thinks } \wedge x \text{ has blighted Bob's mare} \wedge \ \& \ \text{Nob wonders } \wedge x \text{ killed Cob's sow} \wedge),$$

where ' \wedge ' serves as a content-quotation mark. Note the quantification into both *ungerade* contexts. (Cf. note 13 above regarding the error of replacing 'mythical' with 'merely possible'.)

27. This may better capture Geach's intent. The first conjunct is notional. The second is relational, and entails that there is exactly one mythical witch whom Hob relationally thinks has blighted Bob's mare. Quasi-formally:

$$\text{Hob thinks } \wedge (\exists x)(x \text{ is a witch} \ \& \ x \text{ has blighted Bob's mare}) \wedge \ (\lambda y)[\text{Nob wonders } \wedge y \text{ killed Cob's sow} \wedge] (\exists x)(x \text{ is a mythical-witch} \ \& \ \text{Hob thinks } \wedge x \text{ has blighted Bob's mare} \wedge).$$

The principally intended reading of (5) is perhaps best captured by an equivalent formulation:

$$\text{Hob thinks } \wedge (\exists x)(x \text{ is a witch} \ \& \ x \text{ has blighted Bob's mare}) \wedge \ \& \ \text{Nob wonders } \wedge \textit{dthat}([\exists x)(x \text{ is a mythical-witch} \ \& \ \text{Hob thinks } \wedge x \text{ has blighted Bob's mare} \wedge)] \text{ killed Cob's sow} \wedge,$$

interpreting '*dthat*'-terms so that their content is their referent (cf. note 14 above). Elizabeth Harman has suggested (in conversation) a neutral reading on behalf of the speaker who remains cautiously agnostic on the question of witchcraft: replace ' x is a mythical-witch' with the disjunction, ' x is a witch \vee x is a mythical-witch'.

28. It can happen that Hob misidentifies Maggoty Meg with, say, her mythical sister. Hob

might thus notionally think that only one witch has blighted Bob's mare even though there are two mythical witches each of whom Hob relationally thinks has blighted Bob's mare.

One further note: The present analysis entails that (5) is committed to mythical witches. The analysis is not itself thus committed, and is consistent with the thesis that (5) is untrue precisely because of this commitment. Disbelief in mythical objects is insufficient ground for rejecting the analysis. It is a basis for rejecting the present solution to Geach's puzzle (which takes it that (5), so analyzed, can be true in the absence of witches, assuming Hob and Nob are real), but carries with it the burden of explaining the intuition that (5) can be true *sans* witches—a challenge that might be met by providing a plausible rendering of (5), as intended, that is free of mythical objects. (Good luck.)

REFERENCES

- Almog, J. 1998. "The Subject Verb Object Class." In J. Tomberlin (ed.), *Philosophical Perspectives 12: Language, Mind, and Ontology*. Cambridge, Mass.: Blackwell.
- Barwise, J., and J. Perry. 1983. *Situations and Attitudes*. Cambridge, Mass.: MIT Press.
- Burge, T. 1983. "Russell's Problem and Intentional Identity." In J. Tomberlin (ed.), *Agent, Language, and the Structure of the World*. Indianapolis: Hackett.
- Church, A. 1956. *Introduction to Mathematical Logic I*. Princeton: Princeton University Press.
- Clark, M. "Critical Notice of P. T. Geach, *Logic Matters*." *Mind* 74: 122–36.
- Dennett, D. 1968. "Geach on Intentional Identity." *Journal of Philosophy* 65: 335–41.
- Donnellan, K. 1966. "Reference and Definite Descriptions." *The Philosophical Review* 75: 281–304.
- Geach, P. (1967a). "Intentional Identity." *Journal of Philosophy* 64: 627–32. Reprinted in Geach *Logic Matters*. (Oxford: Basil Blackwell, 1972), .
- . 1967b. "The Perils of Pauline." *Review of Metaphysics* 23. Reprinted in Geach, *Logic Matters*, (Oxford: Basil Blackwell, 1972), 153–65.
- . 1976. "Two Kinds of Intentionality." *Monist* 59: 306–20.
- Kaplan, D. 1969. "Quantifying In." In D. Davidson and J. Hintikka (eds.), *Words and Objections: Essays on the Work of W. V. Quine*. Dordrecht: D. Reidel.
- Kripke, S. 1972. *Naming and Necessity*. Cambridge, Mass.: Harvard University Press.
- . 1973. "Reference and Existence: The John Locke Lectures for 1973." Oxford University Press (unpublished).
- . 1977. "Speaker's Reference and Semantic Reference." In P. French, T. Uehling, and H. Wettstein (eds.), *Contemporary Perspectives in the Philosophy of Language*. Minneapolis: University of Minnesota Press.
- McKinsey, M. 1986. "Mental Anaphora." *Synthese* 66: 159–75.
- Neale, S. (1990). *Descriptions*. Cambridge, Mass.: MIT Press.
- Partee, B. 1972. "Opacity, Coreference, and Pronouns." In D. Davidson and G. Harman (eds.), *Semantics of Natural Language*. Boston: D. Reidel.
- Quine, W.V. 1956. "Quantifiers and Propositional Attitudes." *Journal of Philosophy* 53. Reprinted in Quine's *The Ways of Paradox* (New York: Random House, 1966), 183–94.
- Richard, M. 1998. "Commitment." In J. Tomberlin (ed.), *Philosophical Perspectives 12: Language, Mind, and Ontology*. Cambridge, Mass.: Blackwell.
- Russell, B. 1905. "On Denoting." *Mind* 14: 479–93.
- . 1919. *Introduction to Mathematical Philosophy*. London: George Allen and Unwin.
- Saarinen, E. 1978. "Intentional Identity Interpreted: A Case Study of the Relations Among Quantifiers, Pronouns, and Propositional Attitudes." *Linguistics and Philosophy* 2: 151–223.
- Salmon, N. 1982. "Assertion and Incomplete Definite Descriptions." *Philosophical Studies* 42: 37–45.
- . 1991. "The Pragmatic Fallacy." *Philosophical Studies* : 83–97.
- . 1998. "Nonexistence." *Noûs*, 32: 277–319.
- van Inwagen, P. 1977. "Creatures of Fiction." *American Philosophical Quarterly* 14: 299–308.
- Wettstein, H. 1981. "Demonstrative Reference and Definite Descriptions." *Philosophical Studies* 40: 241–57.
- . 1983. "The Semantic Significance of the Referential-Attributive Distinction." *Philosophical Studies* 44: 187–96.

CHAPTER 7

Truth and Identity

Marian David
University of Notre Dame

ACCORDING TO A classical *correspondence theory* of truth, a proposition is true iff it corresponds to a fact. The approach has its competitors. One of them, the *identity theory* of truth, pushes for a surprising simplification. It says that true propositions do not correspond to facts, they *are* facts. Some find this view too bizarre to be taken seriously. Some are attracted to it because they worry that the correspondence theory opens a gap between our thoughts and reality—a gap that, once opened, will turn out to be unbridgeable, thus making it impossible for our thoughts to come into contact with reality and for us to attain knowledge. They think the identity theory will avoid these nasty consequences because it does not open *the gap* to begin with. The no-gap theme will play a role in the background of this chapter. It will surface at times, but the chapter is more concerned with a different theme, the collapse-charge. Opponents of the correspondence theory sometimes charge that the theory is unstable, that it must collapse into the identity theory because there is not enough play between true propositions and facts to leave room for a genuine relation to hold between them. Those who regard the identity theory as absurd might see this a *reductio* of the correspondence theory. Others might see it as an argument for the identity theory. After some exploration of the identity theory, I will present one form of the collapse-charge, then I will discuss what a correspondence theorist has to offer by way of a response.

AT THE BEGINNING of the twentieth century, G.E. Moore contributed an article on truth to Baldwin's *Dictionary of Philosophy and Psychology*. In this article, he claimed that there is no room between truth and reality for any relation other than identity:

It is commonly supposed that the truth of a proposition consists in some relation which it bears to reality; and falsehood in the absence of this

relation. The relation in question is generally called a ‘correspondence’ or ‘agreement’, and it seems to be generally conceived as one of partial similarity; but it is to be noted . . . that it is essential to the theory that a truth should differ in some specific way from the reality, in relation to which its truth is to consist . . . It is the impossibility of finding any such difference between a truth and the reality to which it is supposed to correspond which refutes the theory . . . Once it is definitely recognized that the proposition is to denote, not a belief or form of words, but an object of belief, it seems plain that a truth differs in no respect from the reality with which it was supposed merely to correspond¹

Moore does not mention facts, he talks about reality instead. However, since facts are the (bits of) reality that true propositions are typically supposed to correspond to (by correspondence theorists), I take the central claim of the identity theory to be that true propositions are (identical with) facts. But this claim offers only a necessary condition for a proposition’s being true. Other theories of truth typically provide sufficient conditions as well. Since the identity theory seems intended as an alternative, or rather a competitor, to other truth theories (especially to the correspondence theory), we should strengthen it into an equivalence claim, so that it is formally on a par with its peers. We could try the following as a first-shot formulation:

(IT*) For every x , x is a true proposition iff x is a fact.

Advocates of the identity theory often express their view in terms of the weaker claim that covers only the left-to-right direction of (IT*).² This is convenient because the weaker claim is easier to put into ordinary words and is already strong enough to bring out the intended contrast with the correspondence theory.

(IT*) does not actually mention the concept of *identity*. Why, then, refer to it as the identity theory? One could of course expand its right-hand side so that it says ‘ x is identical with a fact’. But then again, one could expand the right-hand side of the correspondence theory so that it says ‘ x is identical with something that corresponds to a fact’—and this would not make us think that the correspondence theory is really an identity theory. The point of referring to the identity theory as the identity theory is not so much to emphasize the *presence* of the concept of identity, which can be inserted into every predication anyway, it is rather to emphasize the *absence* of the concept of correspondence. This absence marks the contrast with the correspondence theory. The contrast, incidentally, can be described in two ways. If one holds that the meaning of ‘corresponds’ already entails the nonidentity of the corresponding items, then the identity theory is incompatible with the claim that a proposition is true iff it corresponds to a fact. If, on the other hand, one holds that the meaning of

‘corresponds’ does not by itself entail the nonidentity of the corresponding items, then the identity theory is not incompatible with the claim that a proposition is true iff it corresponds to a fact. There is still disagreement though, for the identity theory implies that identity is the only correspondence relation that can fill the bill, which goes very much against what the correspondence theorist wanted to say. The latter way of describing the contrast fits better with the charge that correspondence must collapse into identity. The former way is suggested by Moore in the passage quoted above. As far as I can see, it does not matter much which way we choose.

Our formulation of the identity theory is not quite satisfying: it does not cover *falsehood*. At first, one might think that falsehood is handled easily by simply negating the right-hand side of (IT*): For every x , x is a false proposition iff x is not a fact. But no—combined with (IT*), this would entail, absurdly, that everything there is is a proposition. Some might be tempted to “go Meinongian” and to say that a false proposition is identical with a fact that does *not exist* (and a true one is identical with a fact that does exist). But there are no facts that do not exist. To handle falsehood, the identity theorist has to say that a proposition is false iff it is identical, not with any old thing that is not a fact, but with a *proposition* that is not a fact. This avoids the unwelcome consequence that anything there is is a proposition. It also suggests that we should make a parallel adjustment in the clause for truth, so that the complete formulation now says: For every x , x is a true proposition iff x is a proposition that is a fact, and x is a false proposition iff x is a proposition that is not a fact; or equivalently

- (IT) For every proposition x , x is true iff x is a fact;
 For every proposition x , x is false iff x is not a fact.

One might wonder whether (IT) is properly called a ‘theory’. Isn’t this a rather grand-sounding label for what is little more than a one-liner? It is, but the label comes in handy, and minitheories of this sort are already familiar to philosophers. Moreover, one can think of (IT) as a core-theory that can be enriched in various ways by saying more about propositions and facts. One may also wonder whether (IT) is intended as a definition, or as an axiom, or as a principle governing truth and falsehood for propositions. My inclination is to set such issues aside: if the correspondence theory collapses into (IT), then the correspondence theory is in trouble—never mind whether (IT) is intended as a definition, or an axiom, or a principle, or whatever. Some will subscribe to the clauses under (IT) merely because they use ‘fact’ as a handy label for true propositions. Should they count as (genuine) advocates of the identity theory? I think they are better regarded as advocating an identity theory of *facts* rather than truth; for they want to emphasize that facts are true propositions, whereas the identity theorist wants to emphasize that true propositions are facts. Or better, they hold that calling

something a 'fact' amounts to saying that it is a true proposition, whereas the identity theorist holds (if her view is to be put into meta-linguistic terms at all) that calling a proposition 'true' amounts to saying that it is a fact. The identity theory is supposed to be a novel theory about truth, not a novel theory about facts and not a stipulation concerning the usage of the term 'fact'. Moore seems to express a similar sentiment in the passage quoted earlier. His words suggest that the identity theorist starts with some rough antecedent fix on the reality to which a truth is *supposed* to correspond (by the correspondence theorist) and that she then goes on to claim that a truth does not differ from this very reality to which it was antecedently supposed merely to correspond.³

Although it is not really my intention to defend the identity theory, I want to discuss some more worries and/or objections one might have. This should help bring the theory into sharper focus. Moreover, the question whether the theory is tenable is interestingly related to the collapse-theme: if it is tenable, and correspondence collapses into identity, then that is not good for the correspondence theorist; if it is untenable, and correspondence collapses into identity, then that is even worse for the correspondence theorist.

First, a rather general worry concerning propositions. The identity theory, as I understand it here, takes propositions seriously. It has no qualms quantifying objectually over propositions; and it is not one of those views that help themselves to proposition-talk while maintaining that they do not incur any ontological commitments to propositions. The worry is that propositions are problematic: their existence is contentious; they are said to be creatures of darkness; their identity conditions are said to be obscure; and so on. Indeed, if there are no propositions then the identity theory, as I understand it, is a theory without a subject matter; however, it still would not be false, for (IT) does not affirm the existence of propositions. The same, by the way, goes for the correspondence theory of truth for propositions, which is the brand of correspondence theory that is primarily relevant to our discussion.⁴ But what are propositions? Well, the identity theory, as I understand it, is committed to the traditional propositional analysis (PA) of belief. The PA characterizes propositions. It does not actually tell us what propositions are; at least, it does not tell us what proposition are made of. Rather, it characterizes propositions in terms of the role they are supposed to play. The characterization is fairly familiar, so I will be brief.

The PA: (i) If you believe that flies are insects, then *what you believe* is the proposition *that flies are insects*. Your belief state is a relational state; it involves a relation to a proposition. The proposition is the *object* of the belief relation and the *content* of the belief state: propositions are "content-objects." Moreover, propositions can be shared. If you and I both believe that flies are insects, then we believe the same proposition. In general, the same proposition can be believed by the same person at different times and by different persons at the same time. The PA extends to many other states and acts; here I will usually talk about

beliefs and thoughts. (ii) Propositions are primary bearers of truth and falsehood. That is, truth and falsehood as applied to beliefs (thoughts, statements, etc.) are parasitic on truth and falsehood as applied to propositions: a true belief is a belief that has a true proposition as its content; a false belief is a belief that has a false proposition as its content. This is the point where theories of truth for propositions connect with larger issues concerning belief and thought. Propositions are also bearers of broadly logical properties and relations, for these are all tied up with truth and falsehood. (iii) The ‘that’-clause in ‘John believes *that flies are insects*’ refers to the proposition expressed by its embedded sentence. It is a special “perspicuous” name of the proposition, for it allows us to specify the proposition referred to while referring to it. Unlike a proper name, or a label, or a description (like ‘Frege’s favorite proposition’), the ‘that’-clause tells us which proposition it is that is being referred to. ‘That’-clauses are often abbreviated schematically, as in ‘John believes that p’, where the dummy-letter ‘p’ can be replaced by any arbitrary declarative sentence that makes sense.

These are some of the salient features by which the PA characterizes propositions in terms of their role. One can then have different views about the nature of propositions, i.e., about what sort of thing, ontologically speaking, can play that role (and one can have the view that nothing can play the role, which would take us back to the worry that there might not be any propositions). Taken by itself, (IT) does not imply much about the nature of propositions; and the question how the PA constrains possible views about their nature is a notoriously tricky one. Similarly, taken by itself, (IT) does not imply much about the identity conditions for propositions; and the question how the PA constrains possible views about their identity conditions is again a notoriously tricky one. Of course, one would hope that a more fully worked out version of the identity theory will have something to say on these subjects. (All this holds equally for the correspondence theory.) However, even our rather undeveloped identity theory does imply a little bit about identity conditions. For example, one might wonder whether, according to (IT), the true proposition that Rome is south of Vienna might be identical with the fact that the capital of Austria is north of the capital of Italy (cf. Cartwright 1987, 74). (IT) does at least give a conditional answer. It says that the proposition is identical with that fact, only if that fact is identical with the fact that Rome is south of Vienna; that much, at least, follows from (IT). By implication (IT) also offers some conditionals about identity conditions for propositions: x and y are the same/different propositions if and only if x and y are the same/different facts.

Objection. Assume the proposition that p is contingently true. The proposition exists whether it be true or false. But the *fact* that p would not have existed, if the proposition had been false. Hence, the proposition that p is not identical with the fact that p —*reductio*. This might seem devastating at first, but

it assumes that ‘the fact that p’ functions like a *rigid designator*. The identity theorist will and must respond that, on the contrary, ‘the fact that p’ is *not* rigid, i.e., it refers to the proposition that p only in those worlds in which the proposition is true. According to the identity theorist, to say of a contingently true proposition that it is identical with a fact is to say that the proposition is (necessarily) identical with something, namely itself, that happens to be a fact—it is a bit like saying that Aristotle is identical with the author of the *Metaphysics*.⁵ I remark in passing that most advocates of the PA will hold that expressions of the form ‘the proposition that p’ are rigid (although I do not remember anyone mentioning this explicitly). They will think that there is no reading under which “the proposition that p might have failed to be the proposition that p” comes out true. This is because a ‘that’-clause is usually supposed to specify the essential nature of the proposition it refers to.⁶

Objection. The identity theory is committed to the claim that facts are true, which is absurd. The identity theorist will have to take this in stride. She will have to say that “facts are true” is literally true; it merely sounds odd because it amounts to the redundant claim that true propositions are true.

The identity theory is committed to the principle of *bivalence* for propositions (every proposition is either true or false). But bivalence is problematic. On the face of it, it seems to fail for vague propositions (maybe also for propositions that are referentially indeterminate and for propositions that suffer from complete reference failure). Bivalence is a tricky issue. I can do little more here than to register that its failure will create serious difficulties for the identity theory (but also for various other theories of truth, including some correspondence theories). It is sometimes asserted, flatly, that all propositions are either true or false—the idea being that, while bivalence does fail for declarative sentences, it never fails for propositions. Note that this move does not sit at all well with the PA, for it seems that we do have, say, vague beliefs and that we typically utter our vague sentences to express our vague beliefs. Since the PA introduces propositions as possible contents of our beliefs and thoughts, it suggests that there are vague propositions after all. There may be arguments to the effect that, contrary to appearances, bivalence never fails (neither for propositions nor for meaningful declaratives); but such arguments have to go far beyond mere assertion. An identity theorist might want to get around the problem by simply restricting (IT) to those propositions that are bivalent. But this is not satisfying either. Consider the first clause of (IT); it is equivalent to: For every x , x is a true proposition iff x is a proposition that is a fact. If we replace ‘proposition’ with ‘proposition that is either true or false’, the result will be explicitly circular.

I want to close this section with some remarks about the no-gap *motif* that is so close to the heart of some identity theorists. A friend of facts will hold with Wittgenstein (1921) that *the world* is the totality of facts. Of course, facts have

further constituents, viz., things and properties and relations; but, in the first instance, the world divides into facts. Now, combined with the PA, the identity theory tells us that, when you think what is true, then what you think, the content of your thought, *is* a fact—not some stand-in or representative of a fact, but a part of the world itself. So, when you think what is false but might have been true, then what you think is not a fact, but it may well have been a fact. False thoughts, according to the PA, have the same kind of content as true thoughts (namely, propositions). So, if the contents of true thoughts are facts, then the contents of false thoughts must be made from the same kind of worldly stuff that facts are made of; they must be just like facts, only *not* facts—unfacts. It is helpful to talk of *states of affairs* in this context. Unlike facts, states of affairs are “bipolar”; i.e., they can obtain or fail to obtain. The ones that obtain are facts. The ones that fail to obtain still exist of course (*pace* Meinong); they just are not facts. We could reformulate the identity theory as suggested by Chisholm (1976, chap. 4): true propositions are states of affairs that obtain (facts); and false propositions are states of affairs that do not obtain (unfacts). Propositions are states of affairs—it is just that, when we think of them as contents of beliefs and thoughts, we tend to call them ‘propositions’, whereas when we think of them as facts or unfacts, then we tend to call them ‘states of affairs’.⁷

The upshot of this is that we have to distinguish between *the world*, i.e., the totality of facts, and *the big-wide world*, i.e., the totality of states of affairs. (The term ‘reality’ could be used to refer to either one: to the world, on the grounds that reality should be everything that occurs or obtains; or to the big-wide world, on the grounds that reality should be everything there is.) One may now ask: What is the point of the no-gap *motif*? The answer to this question is not easy to discern. Although there is no gap between thought in general and the big-wide world (states of affairs), and no-gap between true thought in particular and the world (facts), there is ample space for talk about a gap between false thought and the world. Take the big-wide world and divide it into the world and the remainder: behold the gap. Unfortunately, many of our thoughts are on the wrong side of this gap, namely, the ones whose contents are false propositions.⁸ And what philosophical *Angst* might be assuaged by the two no-gap theses (no gap between thought in general and the big-wide world; no gap between true thought and the world)? It is hard to tell. Surely, they do not assuage the fear that, for all we know, many of our beliefs might be false.⁹

IF THE CORRESPONDENCE THEORY for propositions were to collapse into the identity theory, Would that be bad? Well, on the face of it, the identity theory does seem a bit bizarre. Combined with the PA, it evokes the picture of the mind stepping out of the head and into the world—the mind seems oddly externalized.

What the identity theory amounts to will depend very much on the underlying view of the nature of propositions and facts. Let us look at facts first. The facts, taken together, make up the world. Facts themselves are naturally thought of as composed of worldly objects, properties, and relations. But this must be qualified right away. Facts cannot be “composed” of their constituents in the same sense in which the world is composed of facts. The world is just the facts taken together; it is just the sum of facts. But a fact is more than its constituents taken together. John, Mary, and the relation of loving can enter into two facts at the same time: the fact that John loves Mary and the fact that Mary loves John. Consequently, each fact must be more than the sum of its constituents. Facts are complexes that are not reducible to their constituents: they enjoy a nonmereological mode of composition from objects, properties, and relations.

What about propositions? I will set aside Lewis’s (1986) view that propositions are sets of possible worlds—given a natural view of facts, Lewis-propositions do not go with the identity theory, but they do not go well with the correspondence theory either.¹⁰ Instead, let us think of propositions as having internal structure. Like facts, they will be composed in some nonmereological manner. But what are their constituents? What are propositions made of?¹¹ Let us look at some options and see what emerges when we combine them with the identity theory. If propositions are ordinary sentences, then, on the identity theory, the world is a text. If propositions are sentences in the mind/brain, then, on the identity theory, the world is in our heads. If propositions are (sequences of) immaterial ideas à la Descartes, then, on the identity theory, the world is a modification of our souls. Each one of these options yields a quick *reductio* of the identity theory.¹² Actually, propositions as sentences in the mind/brain or as Cartesian ideas are not even options according to the PA. They are private and cannot be shared, but the PA requires that different persons can believe one and the same proposition. Malebranche had an answer to this “privacy problem.” The contents of our thoughts, he held, are God’s thoughts. Propositions are divine ideas, and we think all things in God. Combine the identity theory with this and you get the view that the world is made of divine ideas, namely, the ones God assents to (as opposed to the ones He merely considers, which would comprise the big-wide world). Malebranche, I take it, would have rejected this as pantheistic heresy. The positions of Hegel and Bradley, however, appear to be of this general sort. They advocated an identity theory but replaced God with *the Absolute*: propositions are the ideas of the Absolute and, since truth is identical with fact, the world is made of the Absolute. (But this neglects the existence of falsehoods. The Absolute ought to be the big-wide world. The world had better be made of something like “the positive aspect” of the Absolute—I am unable to tell how the absolute idealists wanted to handle falsehood).

Frege (1892) and Moore (1899) addressed the privacy problem in a slightly different manner. They held that propositions are composed of *concepts*—where a concept is construed as an objective way of conceiving of things and properties.¹³ Frege-Moore concepts are rather similar to *types* of Cartesian ideas: different persons can have different token ideas of the same type. Since idea-types, or concepts, are neither mental nor physical, are not easily localized in space or in time, and exist independently of individual thinkers, concept-propositions are often called ‘abstract’. With a view such as this, one can hardly object to the identity theory on the grounds that it makes the contents of our thoughts *external* to our heads. Abstract concept-propositions (much like, incidentally, the immaterial idea-propositions of Descartes and Malebranche) are surely not *in* our heads. Still, if the identity theory is combined with this view, the result is a peculiar conception of facts and the world. Facts will be composed of objective concepts of objects and properties rather than being composed of the objects and properties themselves. There will be as many facts surrounding an object (property) as there are different ways of conceiving of the object (property). Concept-propositions à la Frege and Moore are individuated along the lines drawn by our patterns of thinking: different ways of thinking, different propositions. Combined with the identity theory, it follows that there cannot be different ways of thinking of the same scene, different perspectives, that are both accurate: different accurate perspectives already are different scenes. Intuitively, this is a far too mind-infected way of individuating facts, playing into the hands of the common complaint that philosophers tend to read the features of thought and language back into the world.

Unlike the Frege-Moore view, a Kripke-Putnam-Kaplan inspired view makes thought contents dependent on the objects and properties in the thinker’s spatio-temporal environment.¹⁴ On this view—let us call it “strong externalism” (though it isn’t entirely clear how it is stronger than Frege-Moore externalism)—propositions can be composed of objects and properties themselves, rather than concepts of objects and properties. Strong externalism seems made to order for the identity theory of truth. Strongly external propositions that are true must be facts (states of affairs that obtain); and strongly external propositions that are false must be states of affairs that do not obtain. The idea that Mount Everest with all its snow-fields, as well as Aristotle, fleas, avocados, and the like, are in some sense constituents of the contents of our thoughts may initially appear bizarre—still, there it is, vindicated by powerful arguments in the theory of content. But not so fast. Externalist arguments do not fully vindicate the identity theory, not by a long shot. Colin McGinn has reminded us that the arguments supporting strong externalism apply only to those propositional constituents that correspond to proper names, natural-kind terms, and indexicals. They do not work for artifact-concepts and other functional concepts; nor do they work for phenomenological and qualitative concepts. They do not even work for *com-*

plex concepts of natural kinds.¹⁵ In short, the arguments for strong externalism work only for some types of contents. They provide only a very partial vindication of the identity theory, which says, after all, that *every* true proposition is a fact (and every false proposition is made of worldly fact-stuff). The limited reach of externalist arguments leaves two options in the theory of content. (i) Hold that strong externalism applies to contents of all types, even though for most types it is not supported by externalist arguments. (ii) Hold that strong externalism applies only as far as externalist arguments reach; where they do not reach, contents ought to be individuated in the traditional manner, i.e., along the lines of Frege-Moore concept-propositions. The first option goes well with the identity theory, but it requires advocating a surprising view about content that remains largely unsupported. The second, and by my lights more plausible, option offers a mixed view of content, but for many types of contents it does not go well with the identity theory.¹⁶

It looks like the identity theory will have unpalatable consequences no matter what theory of content is in play. The correspondence theory had better not collapse into the identity theory. But what is the nature of the threat anyway? I think, it is this.

The identity theory emerges quite naturally from the way in which truth-talk and fact-talk interact with the use of ‘that’-clauses—it emerges quite naturally, that is, provided one has embraced the PA. According to the PA, ‘that’-clauses occupy *referential position* (subject position) even in contexts where surface grammar does not make this intuitively obvious, viz., ‘S believes that p’ and ‘it is true that p’. According to the PA, we should take surface grammar to be misleading here, for we want to capture valid inferences involving such contexts by quantifying into ‘that’-clause positions: “She believes that p; It is true that p; therefore: She believes something that is true, i.e., for some *x*, she believes *x* and *x* is true”—where the objectual variable ‘*x*’ ranges over propositions. So the form ‘it is true that p’ gets recast into subject-predicate form, ‘*x* is true’, which allows quantifying over propositions. Once this treatment of ‘that’-clauses is in place, there is a smooth transition from the use of ‘that’-clauses in truth-and-fact-talk to the identity theory:

- (a) it is true that p iff it is a fact that p,
- (b) that p is true iff that p is a fact,
- (c) *x* is true iff *x* is a fact,
- (IT) for every proposition *x*, *x* is true iff *x* is a fact.

I do not want to maintain that one will find this derivation actually laid out somewhere. But I do want to maintain that it makes explicit the natural progression of thought that underlies the collapse-charge and leads to the identity theory of truth.¹⁷

Our “derivation” of the identity theory assumes that the ‘that’-clauses in (b) stably refer to propositions. More precisely, the step from (b) to (c) assumes that in each substitution instance of schema (b) both ‘that’-clauses refer to the same thing. The step from (c) to (IT) assumes that the thing in question is a proposition. The assumption is a natural one to make for subscribers to the PA—they may even find themselves embracing the more general principle that *all* ‘that’-clauses refer to propositions (provided their embedded sentences make sense). However, it seems the correspondence theorist has to reject the stability assumption if she is to avoid seeing her theory collapse into the identity theory; she ought to argue that, in substitution instances of (b), the ‘that’-clauses shift reference from propositions to facts.

To see how this might go, compare expressions of the form ‘the proposition that p’ and ‘the fact that p’ with expressions like ‘the planet Jupiter’ and ‘the god Jupiter’ (and ‘the man Descartes’ and ‘the town Descartes’, etc.). The latter are definite descriptions, but they are not quite like ordinary definite descriptions. The ordinary description ‘the planet beyond Jupiter’ refers to a thing *other* than Jupiter by relating it to what the embedded name ‘Jupiter’ refers to. But ‘the planet Jupiter’ refers to the *very same* thing as the embedded name ‘Jupiter’ refers to (in the embedding context). The embedded name is referentially ambiguous and the description serves to disambiguate it: ‘the planet . . .’ and ‘the god . . .’ tell us how to take the ambiguous name ‘Jupiter’. Let us call such descriptions *disambiguating descriptions*. They have a second interesting feature. Unlike ordinary descriptions, they can be turned into subject-predicate sentences without much ado. Simply take the embedded name and use it as the subject: ‘the planet Jupiter’ turns directly into ‘Jupiter is a planet’, in which the predicate serves to disambiguate the name retroactively. Now, expression of the form ‘the proposition that p’ and ‘the fact that p’ can be understood as disambiguating descriptions of this sort.¹⁸ The ‘that’-clause, ‘that p’, is ambiguous; it refers to one type of thing when preceded by ‘the proposition’, namely to a proposition, and to another type of thing when preceded by ‘the fact’, namely to a fact. ‘The proposition that p’ refers to the proposition expressed by the sentence embedded in the ‘that’-clause. ‘The fact that p’ refers to whatever fact is the truthmaker for the proposition expressed by the sentence embedded in the ‘that’-clause. Since ‘the fact that p’ is a disambiguating description, the embedded ambiguous ‘that’-clause refers to the very same thing, if any, as the whole description, namely to the fact that p. Remember that disambiguating descriptions can be turned into subject-predicate sentences in which the predicate has to do the disambiguating work retroactively. Our two descriptions are readily transformed into ‘that p is a proposition’ and ‘that p is a fact’, in which ‘that’-clauses refer to different things. This is the form in which they appear in schema (b). Since the ‘that’-clause in ‘that p is a fact’ refers to a fact iff its embedded sentence expresses a

true proposition, 'that p is a fact' is equivalent to 'that p is a true proposition'. This takes care of schema (b).

So the correspondence theorist can give an account of why and how 'that'-clauses can switch referents from propositions to facts. This blocks the derivation of the identity theory: the step from (b) to (c) fails because it relies on the mistaken assumption that the 'that'-clauses occurring in truth-and-fact-talk must have stable reference to propositions.

So far, this is primarily a defensive move. It would be nice if we had some idea of how the correspondence theorist might go on from here. She might subscribe to the *atomistic* program, first proposed by Wittgenstein (1921) and Russell (1918), and later modified and developed by David Armstrong (1997) and others. Let me try to give a very condensed sketch. *First*, uphold what Armstrong calls the *truthmaker principle*: for every truth there must be something in the world that makes it true, i.e., every true truthbearer must have a truthmaker.¹⁹ *Second*, reject the tempting idea that correspondence is a one-one relation between truthbearers and truthmakers. Adopt a *sparse* theory of truthmakers instead. For example, a disjunctive proposition is true iff either one, or both, of its disjuncts are true; different disjunctive propositions can be made true by the same truthmaker: no need for disjunctive facts. Ideally, all molecular (logically complex) propositions should be handled in some such manner, so that there is no need for any facts but *atomic* facts (and aggregates of atomic facts). *Third*, reject the tempting idea that there is a one-one correspondence between predicative concepts and genuine universals. Adopt a *sparse* theory of universals instead. Most predicates we use express concepts rather than genuine universals. Genuine universals (properties and relations) are objective features of the world that ground the objective resemblances among particulars and explain their causal powers. What universals there are will have to be decided on the basis of total science. It is not to be decided by looking at what concepts there are: universals are not concepts. *Fourth*, atomic facts, the truthmakers, are composed of fundamental particulars and genuine universals.²⁰ Working out the atomist program is nontrivial and concessions may have to be made along the way. In particular, negative and/or universal truths cause difficulties. Atomists may be forced to count some negative and/or universal facts among the truthmakers. In general, they will try to keep things as sparse as possible—see Armstrong (1997) for a recent defense of atomism.

An atomist can subscribe to the PA view about 'that'-clauses, but with a rider. 'That'-clauses are perspicuous names of propositions, *provided* they are used in believe, truth, and proposition contexts. When used in fact contexts, say, 'that p is a fact' and 'the fact that p', they are typically all but perspicuous. In such contexts, 'that'-clauses will often have messy reference, referring to whatever atomic facts make true the proposition that p. Only when the proposition

that p is elementary will the ‘that’-clause in ‘the fact that p ’ have a nicely perspicuous reference to the atomic fact in question. In this case, ‘that p ’ in ‘the fact that p ’ will refer to the fact that p (remember that ‘Jupiter’ in ‘the planet Jupiter’ refers to Jupiter.)

If some types of contents are strongly externalist, then propositions may be a varied lot. One may have to distinguish between three different types: (i) Pure strongly externalist propositions, composed entirely of particulars and genuine universals; e.g., the proposition that Fido is a dog, and maybe the proposition that there is water on Mars. (ii) Pure concept-propositions, composed entirely of Frege-Moore concepts; e.g., the proposition that beverages are usually kept in containers, the proposition that doorknobs are cheaper than carburetors, the proposition that the tallest spy is a university professor. (iii) Mixed propositions, e.g., the proposition that water is a beverage, the proposition that Aristotle is a famous philosopher. Now, assuming there really are thought contents that are pure cases of strongly externalist propositions, the atomistic correspondence theory must make a concession to the identity theory. After all, pure strongly externalist propositions that are *true* must be facts (at least the ones that are taken to be elementary must be facts). So, in this case, correspondence must shrivel to identity, which is not a genuine relation (relational universal) according to a sparse theory of universals. Once more the correspondence theory turns out to be a somewhat messy affair. The relational concept x *corresponds to* y must be a generic concept that refers to (or is realized by?) a number of different setups. In some cases, namely, when x is a pure and true externalist proposition, all it takes for x to correspond to a fact y is for x to exist, for in such cases x is identical with y . In other cases, namely, when x is a pure and true concept-proposition, correspondence is a genuine relation between wholly distinct items. In yet other cases, namely, when x is a true mixed proposition, correspondence is in part identity and in part a relation between distinct items; for, when x is a true mixed proposition, it shares at least one constituent with the fact that makes it true. As far as I can see, the “messiness” of correspondence does not provide ammunition for an objection. It does make life difficult for the correspondence theorist; but life is difficult.

There are two basic forms of correspondence-to-fact theories for propositions. Let x range over propositions:

- (CF) x is true iff x corresponds to a fact;
 x is false iff x does not correspond to any fact.
- (CS) x is true iff x corresponds to a state of affairs that obtains;
 x is false iff x corresponds to a state of affairs that does not obtain.

An advocate of (CS) will hold that a fact is just a state of affairs that obtains (and a state of affairs that does not obtain is an unfact). So, like the identity the-

orist, he will distinguish between the world, i.e., the totality of states of affairs that obtain, and the big-wide world, i.e., the totality of states of affairs. But, unlike the identity theorist, the CSist will be an atomist about states of affairs, or at least, he will try to advocate as sparse a theory of states of affairs as possible. (CF) could be intended, and maybe sometimes is, as a condensed version of (CS). However, a significant number of correspondence theorists (including Russell and Armstrong) would want to be “genuine” CFists, embracing (CF) while rejecting (CS). They would hold that false propositions do not correspond to anything, especially not to nonobtaining states of affairs.²¹

(CS) tends to be regarded with a fair amount of suspicion. It is criticized because it invokes a new primitive, the concept of *obtaining*, which must be as fundamental to states of affairs as instantiation is to universals.²² It is also criticized on the grounds that nonobtaining states of affairs do not go at all well with a “vivid sense of reality” (Russell 1918, 223), that they are not worldly enough, too abstract.²³ (CF) may have advantages over (CS), but it also has some disadvantages. Like the identity theory, it is committed to bivalence, whereas (CS) is not: propositions that do not correspond to any state of affairs are neither true nor false. Also, (CF) tends to have difficulties finding atomic truthmakers for negative propositions; nonobtaining states of affairs might help with this. Also, we have seen earlier that for pure externalist propositions, correspondence will reduce to identity. We have seen that the identity theory has to identify false propositions with states of affairs that do not obtain. So it looks like any correspondence theory has to accept nonobtaining states of affairs anyway, provided it acknowledges pure externalist propositions.

A question can be raised about the response to the collapse-charge. I want to close with a brief discussion of how the two versions of the correspondence theory handle this question. Remember, I said that a correspondence theorist ought to block the derivation of (IT), and hence the collapse-charge, at the step *from* (b) *to* (c). The idea was that the ‘that’-clauses in schema

(b) that p is true iff that p is a fact

do not refer to the same thing: the one on the left refers to a proposition, while the one on the right refers to a fact. Now, assume that the proposition that p is false. Does it not follow, on this account, that the ‘that’-clause on the right-hand side suffers from reference failure? If so, would that not mean that the correspondence theorist cannot really account for (b)? After all, it seems he cannot evaluate it as true, since its left-hand side is false while its right-hand side comes out as neither true nor false.²⁴

I think both types of correspondence theorists should respond that, if the proposition that p is false, then (b) is true because its right-hand side is false too. The CSist has virtually no explaining to do here. He already reads (b) as equiv-

alent to ‘that p is true iff that p is a state of affairs that obtains’, where the ‘that’-clause on the left refers to a proposition and the one on the right to a state of affairs. The problem does not even arise here—although the CSist ought to rephrase our earlier account of how the derivation of (IT) is to be blocked in terms of states of affairs and obtaining. The CFist has a bit of explaining to do. He should say that, when the proposition that p is false, the right-hand side of (b) is false too, because it implies a false existence claim. I pointed out earlier that the disambiguating description ‘the fact that p’ turns easily into the subject-predicate sentence ‘that p is a fact’. Based on this, the CFist could hold that the ‘that’-clauses in (b) are truncated descriptions, so that (b) is a variant of ‘the proposition that p is true iff the fact that p is a fact’, where the right-hand side is false if there is no such fact. Alternatively, he could read (b) as a variant of ‘the proposition that p is true iff the fact that p exists’, where the ‘exists’ could be absorbed into the description in the usual manner. This account seems workable, although it is rather less smooth than the one available for (CS).

I do not know how to decide between the two versions of the correspondence theory. For the moment I would be satisfied to have shown how both can respond to the charge that the correspondence theory will collapse into the identity theory.

NOTES

Precursors of this paper were presented at the Third Inland Northwest Philosophy Conference at Washington State University, and at the Fifth Metaphysical Mayhem conference at Syracuse University. Thanks to George Bealer, Joe Campbell, Andrew Cortens, Tom Crisp, Rob Cummins, Delia Graff, and Timothy Williamson, for helpful comments.

1. Moore 1901–02, 20–21. The real father of the identity theory may have been Hegel (1830, §213): “Truth in the deeper sense consists in the identity between objectivity and the notion.” Compare also: Bradley 1893, 150–52, and 1907, 110–13; Moore 1899, 4–5; Russell 1904, 74–76; Meinong 1910, chap. 3; Frege 1918, 74, 60–61; Ducasse 1940; and Chisholm 1976, chap. 4. For more recent discussions, see Candlish 1989, which introduces the label “identity theory”; Baldwin 1991; McDowell 1994, 27; and Hornsby 1997.
2. Compare, for instance, Hornsby 1997, 2: “The identity theory is encapsulated in the statement that true thinkables are the same as facts.”
3. Thanks to Andrew Cortens for reminding me of the view that subscribes to the clauses under (IT) because it wants to advocate an identity theory of facts. Note that Frege sounds more like an identity theorist about facts rather than an identity theorist about truth: “‘Facts, facts, facts’ cries the scientist if he wants to bring home the necessity of a firm foundation for science. What is a fact? A fact is a thought that is true” (Frege 1918, 74).
4. I should note that it is misleading to speak of *the* correspondence theory. There is no such thing. Instead, there are various groups of such theories for different categories of truthbearers: e.g., sentences, utterances, statements, beliefs, thoughts, propositions. When I talk of *the* correspondence theory, I should be taken to mean one that applies to propositions. The danger of collapse exists first and foremost for this brand of correspondence theory because the identity theory is a theory of truth for proposition.
5. The objection is raised by Moore (1953, 308) and repeated by Kit Fine (1982, 46–47). The response is due to Richard Cartwright, see his 1987, 76–78.
6. What might a view look like on which ‘the proposition that p’ is not rigid? Well, someone

might hold, for instance, the following: the that-clause in 'S believes that p' refers to a brain sentence that has the content that p, and brain sentences do not have their contents essentially. This would allow him to say that the proposition that p (i.e., the brain sentence with the content that p) might have failed to be the proposition that p (i.e., might have failed to have the content that p). Most advocates of the PA will think that this view uses 'proposition' very oddly and that the term would have been used more appropriately to refer to the content that p rather than to the brain sentence that happens to have the content that p.

7. Terminology varies. Chisholm and many others, including myself, use 'state of affairs' to refer to bipolar entities for which nonobtaining, or nonoccurring, does not coincide with nonexistence. Wittgenstein (1921) and Armstrong (1997) use 'state of affairs' to refer to *facts* which are of course "unipolar"; so for their states of affairs nonobtaining does coincide with nonexistence. (Actually, in the *Tractatus* things are a bit more confusing because there are indications that Wittgenstein had Meinongian inclinations at times, cf. 1921, 4.25.)
8. McDowell says: "When one thinks truly, what one thinks *is* what is the case. So . . . there is no gap between thought, as such, and the world" (1994, 27). He does concede right away that thought can be "distanced from the world by being false." Still, the initial no-gap conclusion seems to rely on a fallacious inference from what holds only for *true* thought, as such, to thought as such. When one thinks falsely, what one thinks *is* what is *not* the case; so there is a gap between thought (false thought as such) and the world after all.
9. Note also that the identity theory does not imply that it is somehow easy to attain knowledge. Say, S believes that p, and it is a fact that p. This does not even begin to suggest that S knows that p—to think otherwise would be to confuse knowledge with true belief.
10. Given a natural view of facts, Lewis-propositions yield a *reductio* of the identity theory. A Lewis-proposition is a set of possible worlds. It is true at our world (i.e., true) iff our world (i.e., the world) is a member of the set. Since a fact is a part of our world, the identity theory would end up identifying a set with a part of one of its members.
11. Remember that the PA does not tell us anything about the inner makeup of propositions. It only provides us with relational properties of propositions: they, or at least very many of them, must be possible contents of belief states.
12. If propositions are "abstractions" from mental states, à la conceptualism, then the identity theory says, absurdly, that the world is an abstraction from the mind.
13. Frege would have talked of "modes of presentation" or of the "senses of words" instead; he used 'concept' (*Begriff*) in a different and somewhat strange way.
14. Cf. Kripke 1972; Putnam 1975; and Kaplan 1976.
15. If they have Hydrogen and Oxygen on Twin-Earth, and if they can do some rudimentary mental chemistry, then they can *make* thoughts about water even before they make water; see McGinn's discussion of externalism (1989, chap. 1) from which I have borrowed the term 'strong externalism'.
16. Thanks to Delia Graff for pointing out that there are indeed two options. Maybe I should be less dismissive of the first one. But note that even Twin-Earth arguments seem to presuppose that qualitative thoughts are individuated along traditional lines. The Earthling's and Twin-Earthling's qualitative thoughts are the same because they conceive of (experience) water/XYZ in the same way; their qualitative thoughts would have been different if they had conceived of (experienced) water/XYZ in different ways.
17. Those who take the identity theory to be absurd might use this progression of thought in the course of a *reductio* of the correspondence theory.
18. But are they definite descriptions at all? 'The proposition that p is F' does not seem to dissolve neatly, in Russellian manner, into 'there is exactly one proposition *x* such that *x* is a proposition that p and *x* is F'. The paraphrase is odd because it's unclear what to do with '*x* is a proposition that p'; surely, we don't want it to read '*x* is *the* proposition that p'. Timothy Williams reminded me that we get a similar situation with 'the tallest spy is F', which seems to turn into 'there is exactly one *x* such that *x* is *the* tallest spy and *x* is F'. I think the comparison suggests the solution. In case of the tallest spy, one uses an analysis

- of tallestness: 'There is exactly one spy x such that x is taller than every other spy and x is F '. So we could use the PA to help us with our case: 'There is exactly one proposition x such that for every S , S thinks that p iff S thinks x , and x is F '.
19. Cf. Armstrong 1997, chap. 8. In what follows I will often use Armstrong's terminology; however, where he talks of *states of affairs* I talk of *facts*—I use 'state of affairs' to refer to bipolar entities.
 20. Note that the fact corresponding to a proposition *about* a concept might contain that concept as its object-component, but only if concepts turn out to belong to the ultimate constituents of the world.
 21. Concerning 'the fact that p ', I think the genuine CFist should hold that such descriptions are rigid: the fact that p could not have failed to be the fact that p . The CSist, on the other hand, must hold that such descriptions are non-rigid whenever the state of affairs in question is contingent: the fact that p (i.e., the obtaining state of affairs that p) could have failed to be the fact that p (i.e., could have failed to obtain).
 22. Note that (CF)-facts and (CS)-states are made of very different types of "glue." Applying (CF)-glue to an object a and the universal being- F entails that it is true that a is F . Applying (CS)-glue does not; it merely results in the existence of the state of affairs that a is F . Obtaining, although fundamental to (CS), is not an ingredient of (CS)-glue.
 23. Then again, it seems that facts, containing universals, are not all that concrete in any case. The CFist might say that his atomic facts, at least the ones involving physical objects, have spatial location: the fact that a is F is where a is (and the fact that a - R - b is where a and b are). Could the CSist maintain that the state of affairs that a is F is where a is, even if the state does not obtain?
 24. The question was raised by Timothy Williams.

REFERENCES

- Armstrong, D.M. 1997. *A World of States of Affairs*. Cambridge: Cambridge University Press.
- Baldwin, J. 1901. *Dictionary of Philosophy and Psychology*. New York: Macmillan.
- Baldwin, T. 1991. "The Identity Theory of Truth." *Mind* 100: 35–52.
- Bradley, F.H. 1893. *Appearance and Reality*. Partly reprinted in J.W. Allard and G. Stock (eds.), *Writings on Logic and Metaphysics*. Oxford: Clarendon Press (1994). (Page references are to margin numbers in the reprint.)
- . 1907. "On Truth and Copying." *Mind* 16. In J.W. Allard and G. Stock (eds.), *Writings on Logic and Metaphysics*. Oxford: Clarendon Press (1994). (Page references are to margin numbers in the reprint.)
- Candlish, S. 1989. "The Truth About F. H. Bradley." *Mind* 98: 331–48.
- Cartwright, R.L. 1987. "A Neglected Theory of Truth." *Philosophical Essays*. Cambridge, Mass.: MIT Press.
- Chisholm, R.M. 1976. *Person and Object: A Metaphysical Study*. London: George Allen & Unwin.
- Ducasse, C.J. 1940. "Propositions, Opinions, Sentences, and Facts." *Journal of Philosophy* 37: 701–11.
- Fine, K. 1982. "First-Order Modal Theories III—Facts." *Synthese* 53: 43–122.
- Frege, G. 1892. "Über Sinn und Bedeutung." Translated in B. McGuinness (ed.), *Collected Papers on Mathematics, Logic, and Philosophy*. Oxford: Basil Blackwell (1984).
- . 1918. "Der Gedanke." Translated in B. McGuinness (ed.), *Collected Papers on Mathematics, Logic, and Philosophy*. Oxford: Basil Blackwell (1984). (Page references are to the original pagination.)
- Hegel, G.W.F. 1830. *The Science of Logic*. Translated by W. Wallace in *Hegel's Logic*. Oxford: Clarendon Press (1975).
- Hornsby, J. 1997. "Truth: The Identity Theory." *Proceedings and Addresses of the Aristotelian Society* 97: 1–24.
- Kaplan, D. 1976. *Demonstratives*. In J. Almog, J. Perry, and H. Wettstein (eds.), *Themes on Kaplan*. New York: Oxford University Press (1989).
- Kripke, S. 1972. *Naming and Necessity*. Cambridge, Mass.: Harvard University Press.
- Lewis, D. 1986. *On the Plurality of Worlds*. Oxford: Basil Blackwell.

- McDowell, J. 1994. *Mind and World*. Cambridge, MA: Harvard University Press.
- McGinn, C. 1989. *Mental Content*. Oxford: Basil Blackwell.
- Meinong, A. 1910. *Über Annahmen*, second edition (first edition, 1902). In R. Haller and R. Kindinger (eds.), *Gesamtausgabe*, Band iv. Graz: Akademische Druck-u. Verlagsanstalt (1977).
- Moore, G. E. 1899. "The Nature of Judgment." *Mind* 8. In T. Baldwin (ed.), *Selected Writings*. London and New York: Routledge (1993).
- . 1901–02. "Truth and Falsity." *Dictionary of Philosophy and Psychology*. In T. Baldwin (ed.), *Selected Writings*. London and New York: Routledge (1993).
- . 1953. *Some Main Problems of Philosophy* (lectures given in 1910–11). London: George Allen & Unwin.
- Putnam, H. 1975. "The Meaning of Meaning." *Mind, Language, and Reality: Philosophical Papers*, vol. 2. Cambridge: Cambridge University Press.
- Russell, B. 1904. "Meinong's Theory of Complexes and Assumptions." *Mind* 13. In D. Lackey (ed.), *Essays in Analysis*. New York: Braziller (1973).
- . 1918. "The Philosophy of Logical Atomism." In R.C. Marsh (ed.), *Logic and Knowledge: Essays 1901–1950*. London: George Allen and Unwin (1956).
- Wittgenstein, L. 1921. *Tractatus Logico-Philosophicus*. Translated by D.F. Pears and B.F. McGuinness. London: Routledge and Kegan Paul (1961).

What Is the Role of a Truth Theory in a Meaning Theory?

Kirk Ludwig
University of Florida

INTRODUCTION

WE ARE PREEMINENTLY linguistic beings. An understanding of our linguistic abilities is central to understanding our powers of thought and forms of social organization. One part of the project of understanding our linguistic abilities, has to do with the combinatorial structure of natural languages, which enables a finite supply of primitive terms to have infinite expressive powers, in the sense of grounding our ability to mean and understand an infinity of non-synonymous expressions. We gain an understanding of this feature of natural languages by providing a compositional meaning theory for them: a theory that, from a specification of meanings for a finite vocabulary and a finite set of rules, specifies the meaning of every sentence of the language.

Restricting our attention for the moment to a context insensitive language L , we can think of such a theory as aiming to meet the following two conditions. The first is that it prove true meaning theorems of the form (M) (henceforth M -theorems),

(M) ϕ means in L that p ,

where what replaces ' p ' translates in the language of the theory (the metalanguage) the sentence in L (the object language) denoted by what replaces ϕ . The second is that it does so from axioms in some sense specifying or giving the meanings of primitive expressions in L , in a way that exhibits how our understanding of the sentence depends on our understanding of its significant parts and their mode of combination. Such a theory would give us insight into the structure of our practical ability to speak and understand the languages we have mastered,¹ and how their infinite expressive powers rest on a finite base.

The goal of this chapter is to say what the relation is between *this* project, or its generalization to languages containing context-sensitive elements (henceforth ‘context-sensitive languages’) and so-called *truth-theoretic semantics*. The suggestion that a truth theory, in the style of Tarski (1983),² can play a central role in a compositional meaning theory is a familiar one, due to Donald Davidson (1984d). However, it is often not clear from Davidson’s work, or from that of his followers, exactly how we are to conceive of the connection. It has sometimes been thought that the truth theory is supposed to *replace* a meaning theory, to provide the most we could provide in the way of a compositional meaning theory. On this view, the truth theory does not serve as a compositional meaning theory, but as a more philosophically and scientifically respectable replacement of it (Stich 1976).

I do not believe that this was Davidson’s intent, though there are certainly things he says that would lead one to believe this, and I think many have been misled. This chapter explains what I think the connection is, and makes much more explicit the role of a truth theory in a compositional meaning theory than Davidson has. I do this by stating explicitly the form of a theory that entails all instances of (*M*) for a language in a way that makes central use of a truth theory, and which I believe meets at least in spirit the other requirements on an adequate compositional meaning theory. Even if I am mistaken in thinking that this represents Davidson’s line of thought (one will not find what I say explicitly in anything he has written), it illustrates one important way of seeing how to exploit the recursive machinery of a truth theory to meet the goal of a compositional meaning theory. And it has some unexpected benefits. It enables us to see how the theory can achieve *its* aims even though the language contains defects that spell trouble for a *truth theory simpliciter*. I have in mind specifically the semantic paradoxes and semantic vagueness, both of which present serious difficulties for truth theories for natural languages.

I first develop the account for a context insensitive language, where the issues will be clearest. Then I extend the account to context sensitive languages. Finally, I show how this helps us out of what have been taken to be some very serious difficulties for the truth-theoretic approach to meaning. In the appendix, I explain why a recursive translation theory cannot achieve the same aims.

FROM TRUTH TO MEANING

Davidson was driven to propose using a truth theory to do duty for a meaning theory by despair of providing a theory that more straightforwardly entailed *M*-theorems. We need not be concerned here with his arguments against more straightforward attempts to do this. Our question is whether there is some other route to the same result through a truth theory.

The first question to ask is why it would it so much as look as if a truth theory could do duty for a meaning theory? It is obvious that a truth theory is

not a meaning theory, and that its theorems (of the form of (T) below) do not tell us what sentences in the language for which it is a theory mean.

The key to seeing why we might nonetheless gain insight into meaning through knowledge of a truth theory is by seeing that if we knew enough about the theory we would know something that enabled us to interpret correctly the sentences of the language for which it was a theory. In particular, as Davidson noted, if we knew that a truth theory satisfied Tarski's Convention T (Tarski 1983, 187–88), we would be a long way toward knowing what we needed to know to interpret object language sentences. Convention T requires of a formally correct truth definition for a predicate 'is true-in- L ' for a language L that it entail for all sentences of the object language a theorem of the form (T) ,

$$(T) \phi \text{ is true-in-}L^3 \text{ iff } p,$$

where ϕ is replaced by a structural description of an object language sentence (a description in terms of the mode of combination of its meaningful constituents), and p is replaced by a translation of ϕ into the metalanguage. Let us call sentences of the form (T) that meet this condition T -sentences. Suppose we knew a truth theory for a language and we knew it met Convention T . If we had then some means of mechanically picking out which of its theorems were the T -sentences, we would be in a position to understand any sentence of the object language. Why is this? Consider (1).

$$(2) \text{'Gwyn' } \wedge \text{'yw' } \wedge \text{'eira' is true in Welsh iff snow is white.}$$

Suppose that this is a T -sentence for the Welsh sentence 'Gwyn yw eira'. We understand (1). Knowing it is a T -sentence tells us that 'snow is white' in English translates 'Gwyn yw eira' in Welsh. Thus, we know (2), which is all we need to know to interpret this sentence in Welsh.

$$(2) \text{'Gwyn' } \wedge \text{'yw eira' means in Welsh that snow is white,}$$

(I ignore complications having to do with tense.) As Davidson remarked at one point (and as far as I know only at one point), if we know that (1) is a T -sentence, what we know guarantees that if we replace 'is true in Welsh iff' with 'means in Welsh that' we will not go wrong (Davidson 1984c, 60). This observation is the key to seeing (i) how to make explicit a compositional meaning theory that relies on the recursive machinery of a truth theory in generating M -theorems, and (ii) how to generalize Convention T to a context sensitive language.

A truth theory is not a meaning theory. Yet, if we knew enough about a truth theory of the right sort, we would be in a position to interpret object lan-

guage sentences. How can we turn this into an explicit meaning theory? In order to formulate an explicit meaning theory, a theory that will have as consequences all true *M*-theorems (and no false *M*-theorems), we need to make explicit everything that we need to know *about* a truth theory in order to be able to interpret sentences of the object language on the basis of understanding their significant parts.

We want more from a compositional meaning theory than just the *M*-theorems for a language. We want in addition that it should inform us about how the meanings of complexes in the language depend on those of their parts in a way that enables us to understand the complexes on the basis of understanding the primitives and rules for their combination. That is why, even for a language with a finite number sentences, we cannot simply give a list of true *M*-theorems for the language. This is an important requirement, which is sometimes overlooked in discussions of the role of a truth theory in providing a compositional meaning theory (see the appendix for further discussion).

We can divide our task into two parts. First, we must answer the question of what constraints a truth theory must meet in order for it to have among its theorems all the *T*-sentences for the object language, and that it, in some sense to be specified, exhibit in (at least certain) proofs of the *T*-sentences how understanding of the sentences depends on understanding their parts. Second, we must say what we have to know in addition to that fact about a theory in order to be in a position to state *M*-theorems for each sentence of the object language; in particular, we have to say what we could know that would enable us to pick out the theorems that are the *T*-sentences.

Davidson at first thought that, for a natural language that contains such context-sensitive elements as demonstratives and tense, a formally correct truth theory that was simply true would *ipso facto* satisfy (a suitable analog) of Convention *T* (Davidson 1984d). This hope proved ill founded. From any extensionally adequate truth theory, we can generate another extensionally adequate truth theory that generates theorems giving truth conditions for sentences using nonsynonymous sentences in the metalanguage. This is obvious once we reflect that there can be nonsynonymous but extensionally equivalent satisfaction clauses for object language predicates.⁴ The position Davidson subsequently adopted was that the truth theory not be merely true, but be confirmable from the standpoint of a radical interpreter (Davidson 1984a, 139). The hope was that this would put enough additional constraints on the theory to ensure that it would satisfy Convention *T* (or a suitable analog).

I am not concerned currently, however, with the adequacy of these additional constraints. Davidson's aim was to provide constraints on a truth theory which would shed light on the relation between meaning and relatively more primitive concepts, particularly those employed in describing empirical evidence one could have for a truth theory for a speaker or community of speakers

(Davidson 1984a, 137). My interest is more limited. I want to say what knowledge we could have about a truth theory for a language that would enable us to use it to interpret speakers *of that language*. It is a separate issue how we could confirm that *a given speaker or speech community spoke that language*. Indeed, the question what knowledge we could have about a truth theory for a language that would enable us to interpret its speakers is conceptually prior to the question how we could confirm such a truth theory for a speaker or group of speakers. For the question of how and whether we could confirm it on the basis of certain evidence depends upon being able to specify independently what counts as success.⁵

What then could we know about a truth theory that would suffice for its meeting Convention *T*? We might simply say that we know that it meets Convention *T*. But this will not guarantee that appropriate proofs of the *T*-sentences exhibit how the meanings of those sentences are understood on the basis of understanding their parts and mode of combination.⁶ The solution is to require that the axioms of the theory themselves meet an analog of Convention *T*, which will then suffice to insure that the theory meets Convention *T* and that certain proofs of the *T*-sentences will exhibit how the meanings of those sentences are understood on the basis of their structures.

To see what we require, contrast predicate satisfaction clauses (3) and (4) (given in English).⁷

- (3) For all functions f , f satisfies in Welsh 'triongl yw x ' iff $f(x)$ is a triangle.
- (4) For all functions f , f satisfies in Welsh 'triongl yw x ' iff $f(x)$ is a trilateral.

In (3) we use a predicate in the metalanguage synonymous with the object language predicate; in (4) we use a predicate coextensive with (indeed, necessarily coextensive with) the object language predicate but not synonymous with it.⁸ Either could be used to provide a true truth theory for the language, but only (3) would do if we wanted the theory to have *T*-sentences among its theorems. (3) exemplifies the most straightforward way to give a truth theory for a language which we understand, namely, by using a sentence in the metalanguage which is of the same form as the object language sentence, and which employs a predicate, or recursive term, which is synonymous with the object language expression for which satisfaction conditions are being given.

This then is what we require: that axioms of the truth theory,⁹ reference axioms, predicate satisfaction axioms, and recursive axioms, use terms in the metalanguage in giving their reference or satisfaction conditions that *are synonymous with* the object language expressions for which they are used to give satisfaction conditions. More fully, for reference axioms, we require that the correct referents be assigned to referring terms in the object language, and, if there is more to their meaning than that, also that a metalanguage term synonymous with the

object language term be used (the metalanguage can be enriched as needed). For example, in ‘the referent of ‘Caesar’ in Welsh is Caesar’ we use a term in our metalanguage, ‘Caesar’, in given the referent of the object language expression ‘Caesar’, which is synonymous with it (this may only come to their having the same referent, but if more is required the convention requires that it be supplied). For predicate satisfaction clauses, we require a predicate in the metalanguage synonymous with the object language predicate be used in giving the satisfaction conditions, and that the sentence form on the right hand side of the quantified biconditional be the same in logical form as that in the object language. (3) provides an example. For recursive terms, we require that the metalanguage term (or structure) used in the recursion be synonymous with the object language term (or structure) the axiom discharges, and that the sentence form on the right hand side of the embedded biconditional be the same in logical form as the object language sentence for which satisfaction conditions are being given. (This will be qualified when we turn to context-sensitive languages.) Thus, for example, for truth functional connectives, we use in the metalanguage a synonymous truth functional connective in giving satisfaction conditions, as in (5).

- (5) For any function f , any formulas ϕ, ψ , f satisfies in Welsh $\phi \wedge \psi$ iff f satisfies in Welsh ϕ and f satisfies in Welsh ψ .¹⁰

Similarly for other connectives, and for quantifiers. One has only to think here about how we in fact standardly proceed to give an axiomatic truth theory for a language we understand. Let us call this requirement *Convention S*.¹¹ If we know that a formally correct truth theory meets Convention *S*, then we can be assured that it meets Convention *T*. Let us call a truth theory that meets Convention *S* an *interpretive truth theory*. This is a stronger condition in general than requiring that a truth theory meet Convention *T*.

The simplest proofs (speaking loosely) of *T*-theorems (theorems of the form (*T*)) in an interpretive truth theory produce *T*-sentences in a way that shows how the truth conditions of the sentence are determined from the reference and satisfaction conditions of their parts, using in the metalanguage expressions synonymous with those for which satisfaction conditions are given. This can then be fairly said to *show* how the meaning of the sentence depends on the meanings of its parts. This completes the first part of the task we set above.

If we know that a formally correct truth theory is interpretive, we know it meets Convention *T*, and that there are proofs of *T*-sentences that exhibit how the meanings of sentences depend on the meanings of their parts. To use such a theory for interpreting object language sentences, however, we need to know more than this.

First, we need to know some mechanical way of identifying the *T*-sentences

among the theorems of the theory. If we allow the theory a rich enough logic, we will be able to prove T -theorems that are not T -sentences.¹² What we need is to define a predicate that, relative to a formal interpretive truth theory, applies to all and only T -sentences, and whose extension can be determined mechanically, at least in the sense that for any sentence we are given of the language, we can mechanically determine for it a T -theorem which falls in the extension of the predicate.¹³ Intuitively, given an interpretive truth theory, proofs that draw solely on the content of the axioms in proving T -theorems will yield T -sentences. Let us call such theorems *canonical T-theorems*. This is not itself a syntactical notion. But for a given theory with its logic, we can characterize a syntactical notion that aims to be coextensive with this intuitive notion. We do this by characterizing a canonical theorem as a T -theorem that is the last sentence of a proof meeting certain constraints that ensure that only the content of the axioms is drawn on in proving it. This can be accomplished by restricting the rules we can appeal to in proofs and what we can apply them to. We can call proofs that satisfy the constraints *canonical proofs*. We can call a set of rules for constructing a canonical proof for a given object language sentence a *canonical proof procedure* (following Davidson). A canonical proof procedure, for an interpretive truth theory, has a T -sentence as its conclusion; given how it is constructed, it reveals in its structure also the semantic structure of the object language sentence.

There can be no general syntactical characterization of these notions simply because there are many different logical systems we could employ in the theory. For any given theory and logic, it would be straightforward, if somewhat tedious, to write out what restrictions were required. Once we had a characterization of the restrictions required in some logical system, we could in fact weaken the system so that it consisted of only the moves so allowed. In this case every T -theorem of the theory would also be a T -sentence. What then do we need to know about an interpretive truth theory in order to pick out its canonical theorems? We can put it this way: we need to know a canonical proof procedure for the theory or that its logic permits only canonical T -theorems. (In the appendix a simple example is given in the course of the discussion of recursive translation theories.)

Second, we also need to know what the theory says, for we might know that a truth theory *in Italian for Welsh* meets Convention S without being in a position to interpret Welsh, because we don't know Italian. Moreover, we don't just need to know what the theory expresses, we need to know that the theory for which we know a canonical proof procedure says what the theory expresses. That is to say, we need to know enough to be able to understand the theory. Otherwise, our knowledge of how to pick out T -theorems that are T -sentences is not connected with our knowledge of what the theory expresses in a way that allows us to interpret object language sentences. Our semantic and syntactical knowledge must be matched.¹⁴

This completes the second part of our task. What remains is to state all of this explicitly, that is, to write out the propositions that we must know in order to use a truth theory to interpret object language sentences. If we identify the meaning theory with the body of knowledge that is required in order to interpret object language sentences on the basis of knowledge of the meanings of their parts, then the result will be the meaning theory itself, as distinct from the truth theory we exploit in formulating it. And this will also make explicit the relation between the truth theory and the compositional meaning theory.

Let us suppose that we have an interpretive truth theory \mathfrak{T} for a language \mathcal{L} , in a metalanguage \mathfrak{M} , with axioms $A_1 \dots, A_2 \dots, \dots$, and a specification of a canonical proof procedure (CP) for \mathfrak{T} . In addition to the usual vocabulary required in a truth theory for a given object language, we will require \mathfrak{M} to contain a predicate, μ , which is a translation of ‘ x means in \mathcal{L} that’. A *meaning theory*, \mathfrak{M} for \mathcal{L} can then be stated in the following form:

1. \mathfrak{T} in \mathfrak{M} is an interpretive truth theory for \mathcal{L} .
2. The axioms of \mathfrak{T} are $A_1 \dots, A_2 \dots, \dots$.
3. A_1 means in \mathfrak{M} that \dots ; A_2 means in \mathfrak{M} that \dots ; \dots ; μ means in \mathfrak{M} x means in \mathcal{L} that \dots .¹⁵
4. CP is a canonical proof procedure for \mathfrak{T} .
5. For any sentence t , any language L , any interpretive truth theory T for L , if t is the last line of a canonical proof in T , then the corresponding M -sentence is true in \mathfrak{M} .

The M -sentence corresponding to a canonical theorem t in an interpretive truth theory is the result of replacing the translation in \mathfrak{M} of ‘is true in \mathcal{L} iff’ in t with the translation of ‘means in \mathcal{L} that’.

Suppose we knew 1–5 for some suitable theory. 1, 2, and 4 suffice for us to be able to identify the T -sentences of the theory and to know that we have identified the T -sentences. 3 ensures that we will understand them. 5 states the knowledge we have that enables us to infer from the T -sentences the truth of corresponding meaning theorems. Semantic descent allows us to infer the theorems themselves. The instantiation of 5 to \mathfrak{T} is something we could deduce from 1–4 given the meanings of ‘is an interpretive truth theory’ and ‘is a canonical proof procedure for’; so in a sense it is redundant, but it helps to make explicit how the connection is made between T -sentences and true M -theorems. The canonical proof of a T -sentence in \mathfrak{T} exhibits (for someone who understands the language of the theory) how the meaning of the object language sentence depends on the meanings of its parts. Thus, someone who knows \mathfrak{M} knows how to interpret any sentence in \mathcal{L} on the basis of knowledge sufficient to understand each of the primitive terms of \mathcal{L} and rules for their combination.¹⁶

Notice that \mathfrak{M} contains statements about the truth theory \mathfrak{T} and its axioms,

but it does not include the axioms of \mathfrak{T} . Indeed, it is clear that \mathcal{M} and \mathfrak{T} need not be in the same language. Thus, surprisingly, we have been led to the conclusion that the truth theory itself is not *part of* the meaning theory.¹⁷ This turns out, as we see below, to be a virtue when we come to some worries about the coherence of defining truth for many natural language sentences. Before we come to that, however, I want to sketch how this approach extends naturally to context sensitive languages.

EXTENSION TO CONTEXT-SENSITIVE LANGUAGES

The extension to context-sensitive languages requires two things. First, we need to explain the appropriate form of the analog of T -sentences and M -sentences for context sensitive sentences. I will continue to call these T -sentences and M -sentences for convenience. Second, we need to say what it is for a truth theory that issues in such T -sentences to be interpretive. Once we have done this, we simply reinterpret 1–5 above according to the notions appropriate for a context-sensitive language; all the morals will carry over straightforwardly.

There is more than one way to adapt a Tarski-style truth theory to a context-sensitive language. One approach is to shift from a predicate of sentences to a predicate of utterances, and this has many advantages. It makes clear that in a context-sensitive language, the primary unit of semantic evaluation is the speech act using a sentence. However, it also entails certain technical complexities that I wish to avoid. I therefore adopt the alternative of introducing a truth predicate with additional argument places for contextual parameters that are relevant to the determination of the semantic contribution of context sensitive elements of the language. For present purposes, I suppose we can get by with just two: speaker and time (place can be reduced to the speaker's location; the contribution of demonstratives can, at a first pass, be secured relative to the speaker's demonstrative intentions).¹⁸ The predicate I introduce is 'x is true in L taken as if spoken by s at t '.¹⁹ That is, in asking whether a certain sentence is true, relative to a speaker and a time, we ask relative to the interpretation it would have, fixing the language, if its context-sensitive elements were assigned semantic values in accordance with rules in the language, given the speaker and time as input. We make the parallel modification in the case of 'x means in L that', to get 'x taken as if spoken by s at t means in L that'. I abbreviate these as 'x is true_[s, t] in L ' and 'x means_[s, t] in L that', respectively.

The form of T -sentences and M -sentences, then, for context-sensitive languages, will be (TCS) and (MCS):

(TCS) For all times t , all speakers s , ϕ is true_[s, t] in L iff p .

(MCS) For all times t , all speakers s , ϕ means_[s, t] in L that p .

The next question is how to say what it is for a truth theory for a context-sensitive language to be interpretive. Convention T no longer applies, since in

(*TCS*) when ϕ is a context-sensitive sentence we will have bound variables in ' p ', and so it would be inappropriate to require that ϕ be translated by ' p '. For example, consider (6), intuitively the *T*-sentence for 'Rydw i yn darllen', Welsh for 'I am reading'.

- (6) For any time t , speaker s , 'Rydw i yn darllen' is true_[s, t] in Welsh iff s is reading at t .²⁰

If a speaker Σ of Welsh utters 'Rydw i yn darllen' at a time τ , instantiating (6) to him yields a specification of its truth conditions that expresses the proposition expressed by the sentence interpreted relative to the occasion of utterance, namely, that Σ is reading at τ . And that is just what we want. Clearly we *don't* want to say that 'Rydw i yn darllen' means the same as ' s is reading at t ', with its free variables. We see what we want: but how can we express the requirement in general terms?

Here is the clue. Convention *T* can be *restated* in the following way:²¹

An adequate truth theory for a context-sensitive language L must be formally correct and entail for all sentences of the object language a theorem of the form (*T*), where ' ϕ ' is replaced by a structural description of an object language sentence,

(*T*) ϕ is true in L iff p ,

such that the result of replacing 'is true in L iff' with 'means in L that' yields a true sentence in the metalanguage.²²

This is equivalent to the original because we can replace 'is true in L iff' with 'means in L that' *salva veritate* if, and only if, the sentence that replaces ' p ' translates that denoted by ϕ . This is, of course, precisely the fact that allows us to move from *T*-sentences to *M*-sentences. (*TCS*) and (*MCS*) are our analogs for (*T*) and (*M*) for a context-sensitive language. To generalize Convention *T* to a context-sensitive language, *we need merely generalize our reformulated statement of it*. Notice that in (6) we can replace 'is true_[s, t] in Welsh' with 'means_[s, t] in Welsh' to yield a true sentence. And that is precisely what it is for the truth conditions assigned relative to a speaker and time to express the proposition a use of the sentence would express in the language relative to the speaker and time. Thus, our modified Convention *T*, which I'll call Convention *TCS*, can be stated as follows:

An adequate truth theory for a context-sensitive language L must be formally correct and entail for all sentences of the object language a

theorem of the form (T) , where ‘ ϕ ’ is replaced by a structural description of an object language sentence,

(T) ϕ is true_[s, t] in L iff p ,

such that the result of replacing ‘is true_[s, t] in L iff’ with ‘means_[s, t] in L that’ yields a true sentence in the metalanguage.

To complete our characterization of an interpretive truth theory, we now need merely to modify Convention S in a similar way. For recursive terms, which are not context sensitive, no modification is needed. For context-sensitive referring terms, we require simply that the reference clause provide the correct referent (if any) relative to a use of the referring term. For context-sensitive predicates, we employ a variant of the device we used for sentences. We will say that an axiom for a predicate, with free variables ‘ x_1 ’, ‘ x_2 ’, ... ‘ x_n ’, denoted by ‘ $Z(x_1, x_2, \dots, x_n)$ ’, which is context sensitive relative to speaker s , and time t ,

For all f , f satisfies_[s, t] $Z(x_1, x_2, \dots, x_n)$ iff $\zeta(f(x_1), \dots, f(x_n), s, t)$,

meets Convention S just in case the corresponding relativized meaning statement is true in the metalanguage:²³

For all f , $Z(x_1, x_2, \dots, x_n)$ means_[f, s, t] that $\zeta(f(x_1), \dots, f(x_n), s, t)$.

Here we introduce a meaning relation that holds between a formula, speaker, time, and function, if the formula interpreted relative to the assignments made by the function to its free variables and taken as if uttered by s at t means what the sentence in the complement clause means (taking the ‘ $f(x_i)$ ’ to be directly referring terms). The revised convention is Convention SCS . A truth theory that meets Convention SCS is interpretive; clearly it will meet Convention TCS if it meets Convention SCS .

This completes the extension of the results of the previous section to context-sensitive languages.

APPLICATION TO SEMANTIC DEFECTS

Explicitly formulating a meaning theory that makes use of a truth theory has some important benefits. This becomes apparent when we consider certain kinds of objections to truth-theoretic semantics that are based on the assumption that in order for a truth theory to aid in the work of giving a compositional meaning theory, the theory must minimally be true and consistent. On this assumption, certain semantic defects in natural languages that make formulating a true truth theory for them problematic threaten to restrict or undermine altogether the possibility of a truth-theoretic semantics for them.

The most obvious difficulty is the possibility of formulating semantic paradoxes in natural languages. Consider the sentence labeled (*L*).

(*L*) The sentence labeled (*L*) in “What Is the Role of a Truth Theory in a Meaning Theory?” is false.

Empirical investigation shows this sentence to say of itself that it is false; if true, then, it is false, and if false, then true, so it is true iff false, which is a contradiction. This will follow from any truth theory that meets Convention *T* (*TCS*) for the language together with the relevant empirical facts. No consistent truth theory can be given for the whole language then. This looks at the least to put some limitations on the use of truth-theoretic semantics in application to natural languages (see Chihara 1976, for example).

An even more serious problem is raised by the fact that many, even most, natural language predicates are vague. In my view, no vague sentence is either true or false, since vague predicates fail a presupposition of our semantic vocabulary, namely, that they are semantically complete and have extensions.²⁴ But problems arise even if one simply accepts that vagueness engenders truth-value gaps. It looks as if one’s truth theory itself will inherit the gaps because one is forced to use metalanguage predicates synonymous with object language predicates in giving truth conditions to meet Convention *T* (or *TCS*).

However, once we recognize that the meaning theory itself, which exploits a truth theory, does not embed a truth theory, then we see that *the truth theory need not be true* in order for it to serve its function. It serves its function by meeting Convention *S* (*SCS*). That is what guarantees that it has as its canonical theorems *T*-sentences; and that is what allows us to infer corresponding *M*-sentences, which are the output of the meaning theory. We do not need to assert the truth theory in order to use its recursive machinery to (a) reveal compositional semantic structure and (b) generate true *M*-theorems.

Take the semantic paradoxes. Even (*L*) will have its *T*-sentence. Ignoring context-sensitive elements, an adequate truth theory would yield (*TL*) as the canonical theorem for (*L*).

(*TL*) ‘The sentence labeled (*L*) in “What Is the Role of a Truth Theory in a Meaning Theory?” is false’ is true in English iff the sentence labeled (*L*) in “What Is the Role of a Truth Theory in a Meaning Theory?” is false.

(*TL*) is problematic. But we do not have to assert it as part of the meaning theory. The meaning theory will generate (*ML*).

(*ML*) ‘The sentence labeled (*L*) in “What Is the Role of a Truth Theory in a Meaning Theory?” is false’ means in English that the sentence labeled

(*L*) in “What Is the Role of a Truth Theory in a Meaning Theory?” is false.

(*ML*) is *true*, just as (7) is.

(7) ‘ $2 + 2 = 5$ ’ means that $2 + 2 = 5$.

Thus, the meaning theory bypasses the difficulties that afflict the truth theory!

The same is true when we turn to semantic vagueness. Again, even if the truth theory uses vague predicates, the meaning theory need not, since it does not include the truth theory, as opposed to statements about the truth theory. Consider (ignoring tense) the *T*-sentence for a sentence about a borderline case for ‘bald’.

(*TB*) ‘Barring Pate is bald’ is true in English iff Barring Pate is bald.

Everyone except epistemicists will agree that (*TB*) is neither true nor false. The truth theory that generates it is therefore defective. Many of its axioms dealing with vague predicates will likewise be neither true nor false. However, the meaning theory is not committed to asserting (*TB*) (or the axioms that lead to it), but rather (*MB*).

(*MB*) ‘Barring Pate is bald’ means that *Barring Pate is bald*.

The trouble with vague terms arises when they have to contribute their *extensional* properties to the truth conditions of sentences in which they are used. While ‘bald’ is in some sense used in (*MB*) in the complement (we do not understand (*MB*) unless we understand the complement), it is clear that it does not contribute its extensional properties to the truth conditions of (*MB*). Thus, even if (*TB*) is without a truth-value, (*MB*) comes out true.²⁵

The point is general. No semantic defect that undercuts the possibility of giving a true truth theory for a language need thereby undermine its use in a compositional meaning theory, since such a theory need not assert the content of the truth theory itself.

CONCLUSION

In conclusion, I have aimed to do three things in this chapter. The first was to make explicit the connection, or at least one sensible connection, between a recursive truth theory for a non–context-sensitive language and the project of giving a compositional meaning theory for it. The second was to show how to extend this result to context-sensitive languages. The third was to show that, once we are clear about the connection between the truth theory and meaning theory, a number of what have been thought to be serious difficulties for truth-

theoretic semantics for natural languages turn out to be impotent, because commitment to the truth of the meaning theory does not entail commitment to that of the truth theory.

APPENDIX

In this appendix, I address the claim that insofar as a truth theory enables us to provide something like a meaning theory for a language, it can in principle be dispensed with in favor of a translation theory.²⁶ This claim is based on the assumption that the only purpose of the truth theory in the meaning theory is to provide us with a way of matching object language sentences with metalanguage sentences that translate them. This assumption is mistaken. It fails to pay attention to the desiderata on a *compositional* meaning theory, which it is our aim to provide by appeal to the mechanism of a truth theory. A compositional meaning theory must exhibit both how our understanding of complex expressions depends on our understanding of their parts and their mode of composition, and how we determined the meaning of a sentence relative to a context of utterance. A translation theory does neither of these. A compositional meaning theory employing a truth theory as sketched above does both.

Consider context sensitivity first. For a natural language, a meaning theory should exhibit how we determine the meaning of a sentence as uttered on a particular occasion. If we consider the form of a truth theory of the sort introduced in “Extension to Context-Sensitive Languages,” it is clear that its function is not at all to provide us with a way of matching sentences with sentences that translate them, but rather to provide context-relativized truth conditions, which then enable us to specify context-relativized statements of what they mean. A translation theory does not issue in any statements about what a sentence means as used by a speaker at a time. A translation theory takes ‘Rydw i yn darllen’ in Welsh blandly into ‘I am reading’ in English, with no hint that its truth may vary from context to context, or that what it means in the mouth of one speaker is different from what it means in the mouth of another. If we understand one of the languages, we can understand an utterance of a sentence from the other, at least as a whole; but we still have no *theory* that reveals anything about their context sensitivity. We might as well say we understand natural languages and be done with it. A translation theory obviously does not do the same job as a truth theory in the case of a context sensitive language, for its job is *not* just to match sentences of the object language with translating sentences of the metalanguage.

Let us turn to our second concern: even a recursive translation theory fails to meet the central desideratum on a compositional meaning theory, namely, that it exhibits how our understanding of complex expressions depends on our understanding of their parts and mode of combination. To see this, consider a translation theory for a simple context insensitive language *L*. The vocabulary and symbols of *L* consist of the following expressions:

$\langle a \rangle, \langle R \rangle, \langle F \rangle, \langle \& \rangle, \langle \sim \rangle, \langle (, \rangle$

The first we call a singular term, the second two we call predicates, the fourth and fifth the conjunction and negation signs, respectively, and the right and left parentheses we call grouping elements. The sentences of L are given by the following rules.

If α is a singular term and ϕ a predicate, then $\alpha \frown \phi$ and $\phi \frown \alpha$ are sentences.

If ψ is a sentence, then $\langle \sim \rangle \frown \psi$ is a sentence.

If ϕ and χ are sentences, then $\langle (\rangle \frown \phi \frown \langle \& \rangle \frown \chi \frown \langle) \rangle$ is a sentence.

(The point of allowing concatenation of a singular term with a predicate in either order will become clear in the sequel, where it will be used to highlight a limitation of the recursive translation theory.) Let $\langle \text{Tr}(x, y) \rangle$ be short for ' x in L^* is translated by y in L^* '. We will allow that L^* has formation rules homologous to those for L . We let ' s ' together with subscripts range over sentences of L , and ' P ' range over predicates of L , and ' n ' over singular terms in L . We give the translation theory as 1–8.

1. $\text{Tr}\langle a \rangle, \langle \alpha \rangle$.
2. $\text{Tr}\langle R \rangle, \langle \rho \rangle$.
3. $\text{Tr}\langle F \rangle, \langle \zeta \rangle$.
4. $\text{Tr}\langle \& \rangle, \langle \wedge \rangle$.
5. $\text{Tr}\langle \sim \rangle, \langle \neg \rangle$.
6. For any n, P , $\text{Tr}(n \frown P, \text{Tr}(n) \frown \text{Tr}(P))$ and $\text{Tr}(P \frown n, \text{Tr}(P) \frown \text{Tr}(n))$.
7. For any s_1, s_2 , $\text{Tr}\langle \langle \sim \rangle \frown s_1 \frown \langle \& \rangle \frown \langle \sim \rangle \frown s_2 \frown \langle \sim \rangle \rangle, \langle \langle \sim \rangle \frown \text{Tr}(s_1) \frown \langle \wedge \rangle \frown \langle \sim \rangle \frown \text{Tr}(s_2) \frown \langle \sim \rangle \rangle$
8. For any s , $\text{Tr}\langle \langle \sim \rangle \frown s, \langle \neg \rangle \frown \text{Tr}(s) \rangle$

1–5 provide translation axioms for primitive expressions term by term. 6–8 provide a recursive procedure for producing translations of complex expressions built from them. So far so good.

The claim we want to examine is whether this theory can serve the ends of a compositional meaning theory. If it can, then it should exhibit how understanding of complex expressions in the language depends on understanding of the primitive expressions and their modes of combination. As stated, of course, it does not say anything that tells us what the primitive expressions of either language mean. So it must obviously be supplemented. Since we are interested in whether it tells us what complex expressions mean on the basis of understanding their primitive components and their mode of combination, we should add axioms that state the meaning of primitive terms in L^* , and we will say, for concreteness, that ' α ' means *Alfred*, ' ρ ' means *is round*, ' ζ ' means *is red*, ' \neg ' means *not*, and ' \wedge ' means *and*.

Yet, even once we have done this, we are *not* in a position to interpret complex expressions in the language. For we have not yet been told anything about how the meanings of the simples contribute to those of the complexes in which they appear, that is to say, the contribution of the mode of combination remains opaque, so far as the information we have been provided with goes. We cannot assume anything about this on the basis of knowing just one word translations for the primitive expressions of the language, even if we are tempted, in the case above, by analogies with familiar artificial languages. An easy way to see this is to notice that it is compatible with the form of the theory given above that concatenation of a *singular term* in L^* with a *predicate* is to be understood as equivalent in English to concatenation of a singular term together with the predicate negation, while the concatenation of a *predicate* with a *singular term* (the reverse order) is understood as simple predication. It is also compatible with the information we have available that these two forms are simple variants of one another and that both represent simple predication, or that both represent application of predicate negation. The translation theory given above doesn't exhibit which way L^* works, even if we know the meanings of its primitive terms. Similar remarks apply to axioms 7 and 8. Knowledge of the translation theory and knowledge of the meanings of the primitive expressions does not automatically give us knowledge of the meanings of complex expressions (i.e., it doesn't put us in a position to understand them, and it doesn't exhibit how an understanding of the complexes would rest on understanding the primitives and their mode of combination). Of course, if we already understood L^* , then we could translate L . But this shows that the theory itself is impotent to give us knowledge of how the meanings of the simple expressions in it contribute to determining the meanings of the complexes in which they can appear. Thus it does not do the same job as a compositional meaning theory that appeals to a truth theory in the manner sketched above. Even translation into a language we understand leaves unarticulated what the truth theory makes plain, how what the parts of complex expressions mean contribute, together with their mode of combination, to determining what we mean in using them. Translation theories then, even recursive translation theories, cannot replace truth theories in the project of providing a compositional meaning theory for a language.²⁷

We can contrast the ineffectual translation theory with a simple truth theory for the language L , which can be used in the fashion indicated above to provide a compositional meaning theory, and work through a sample proof. The truth theory is given by A1–A5.

- A1. ' a ' \frown ' R ' is true in L iff $\text{ref}('a')$ is round.
- A2. ' R ' \frown ' a ' is true in L iff $\text{ref}('a')$ is not round.
- A3. ' a ' \frown ' F ' is true in L iff $\text{ref}('a')$ is red.

- A4. ' $F \wedge a$ ' is true in L iff $\text{ref}(a)$ is not red.
 A5. For any sentence s_1, s_2 , ' $(s_1 \wedge \& \wedge s_2)$ ' is true in L iff (s_1 is true in L and s_2 is true in L).
 A6. For any sentence s , ' $\sim s$ ' is true in L iff it is not the case that s is true in L .
 A7. $\text{Ref}(a) = \text{Alfred}$.

(Here, of course, the puzzle that the translation theory left us with is removed immediately; it is enough to know just that the theory is true, though we will stipulate also that it is interpretive.) A *canonical theorem* is a sentence of the form ' s is true in L iff p ' in which ' s ' is replaced by a structural description of a sentence of L and ' p ' is replaced by a metalanguage sentence which contains no semantic terms and which is the last line of a proof that employs only the following rules in application to Axioms A1–A7 and the results of the following applications:

- R1. *Universal Quantifier Instantiation (UQI)*: For any sentence ϕ , variable v , singular term β : $\text{Inst}(\phi, v, \beta)$ may be inferred from $\text{UQuant}(\phi, v)$.
 R2. *Replacement (RPL)*: For any sentences $\phi, \psi, S(\phi)$: $S(\psi)$ may be inferred from $\text{Eq}(\phi, \psi)$ and $S(\phi)$.
 R3. *Substitution (SUB)*: For any singular terms α, β , sentence $S(\alpha)$: $S(\beta)$ may be inferred from $S(\alpha)$ and $\text{Ident}(\alpha, \beta)$.

' $\text{UQuant}(\phi, v)$ ' means 'the universal quantification of ϕ with respect to v '. ' $\text{Inst}(\phi, v, \beta)$ ' means 'the result of replacing all instances of the free variable v in ϕ with the singular term β '. Note that we count structural descriptions of object language terms, and terms of the form ' $\text{ref}(x)$ ', as singular terms for the purposes of this rule of inference. ' $\text{Eq}(\phi, \psi)$ ' means 'the biconditional linking ϕ with ψ (in that order)'. ' $S(x)$ ' stands for a sentence containing the grammatical unit x , which may be a word, phrase, or sentence. ' $\text{Ident}(\alpha, \beta)$ ' means 'the identity sentence linking α with β (in that order)'.

Sample proof:

1. ' $(a \wedge F) \wedge \& \wedge a \wedge R$ ' is true in L iff ' $a \wedge F$ is true in L and ' $a \wedge R$ ' is true in L '. {from A5 by two applications of *UQI*}
2. ' $a \wedge F$ ' is true in L iff $\text{ref}(a)$ is red. {from A3 by *UQI*}
3. ' $a \wedge R$ ' is true in L iff $\text{ref}(a)$ is round. {from A1 by *UQI*}
4. ' $a \wedge F$ ' is true in L iff Alfred is red. {from 2 and A7 by *SUB*}
5. ' $a \wedge R$ ' is true in L iff Alfred is round. {from 3 and A7 by *SUB*}
6. ' $(a \wedge F) \wedge \& \wedge a \wedge R$ ' is true in L iff (Alfred is red and Alfred is round). {from 1, 4, 5, by two applications of *RPL*}

Suppose we know the canonical proof procedure, the axioms of the theory, and what they mean, and that they meet Convention *S*. Then we know the theory is interpretive. Given this, we know that every instance of the following schema is true:

If (*s* is true in *L* iff *p*), then (*s* means in *L* that *p*),

when the antecedent is instantiated to a canonical theorem. Thus, we can introduce an inference rule, which I will call *MR*:

MR: '*s* means in *L* that *p*' may be inferred from the corresponding canonical theorem of *T*, '*s* is true in *L* iff *p*'.

The rest of the story goes as follows:

- 1'. 6 is a canonical theorem of *T*. {Inspection of the proof and definition of 'canonical theorem'}
- 2'. ('*a*'*F*' & '*a*'*R*') means in *L* that (Alfred is red and Alfred is round) {6, 1', by *MR*}.

The role the truth theory plays is not just to get us to 1'. Its most important role lies in *how* it gets us to 1'. It does so by a method that exhibits how the meaning of the complex sentences of the object language depends on the meanings of its parts, that is, it shows how to understand the complexes on the basis of understanding the parts. It does not state this of course. But someone who understands the theory and knows it meets Convention *S* can see how it works. The mistake of thinking a truth theory does no more than a recursive translation theory rests on thinking that our only aim is to match object language sentences with metalanguage sentences that translate them. But that misses the main point, and interest, of a compositional meaning theory.

NOTES

I would like to thank Professor Reinaldo Elugardo for helpful comments on an earlier version of this paper, which was delivered at the "Truth and Meaning" conference sponsored by the University of Idaho and Washington State University, March 24–26, 2000. I would like to thank the audience at that presentation as well, and to thank Andrew Cortens and Nathan Salmon in particular for questions that have improved, I hope, the presentation of the material in this paper.

1. It is not part of this claim that the ability is instantiated by propositional knowledge of the theory, only that its structure mirror the structure of the complex set of interlocking dispositions required of someone competent in speaking and understanding the language.
2. I have in mind here recursive axiomatic truth "definitions," which use roughly the sorts of devices Tarski introduced in Tarski 1983. See note 9 for further remarks on the relevant class of theories.

3. Tarski was concerned to define a predicate (actually, membership in a set) that applied to the object language only. The shift to using a truth theory in pursuit of a meaning theory will free us from this constraint, and we can regard the predicate as expressing truth in a language, where the place of L is a genuine argument place. We then cease to regard the truth theory as defining a truth predicate and regard it instead as a theory about the language, which may be right or wrong. When we regard it as a theory about a language described as the language of a given speech community or speaker, it is an empirical theory.
4. See Foster 1976; Loar 1976; Evans and McDowell 1976, introduction; Davidson 1984b; and Wallace 1978, 51. More recent discussions include Soames 1989, 1992; Higginbotham 1992; and Richard 1992.
5. This calls into question the intelligibility of Davidson's apparent answer to the question of what constraints a truth theory would have to meet to be useable as a central component of an interpretation theory. To require that a radical interpreter confirm the truth theory leaves his goal underspecified. It is not enough to merely confirm a true truth theory. What about a truth theory does a radical interpreter have to confirm then? It cannot be *that he has confirmed it*, since this *leaves open* what it is about it that he must confirm.
6. We can illustrate the problem with a very simple truth theory for a language without quantifiers. Suppose we have as axioms A1–A4 (we could add more for a richer language without affecting the point to be made), where 'is T ' is our truth predicate for the language, and suppose we have as our logic a suitable complete natural deduction system (I omit formal characterization of the syntax of the object and metalanguages). We suppose further that 'Caesar thrice refused the crown and Caesar was ambitious' in the object language means what that sentence does in English (ignoring, again, tense).

A1. $\text{Ref}(\text{'Caesar'}) = \text{Caesar}$.

A2. For any name α , α 'thrice refused the crown' is T iff $\text{ref}(\alpha)$ thrice refused the crown and for every x , $x = x$.

A3. For any name α , α 'was ambitious' is T iff $\text{ref}(\alpha)$ was ambitious and it is not the case that any thing is both ambitious and not ambitious.

A4. For any ϕ , ψ , ' ($\neg\phi$ 'and' $\neg\psi$ ')' is T iff ϕ is T and ψ is T .

A1–A4 together will be adequate to prove the T -sentence, (T) , since we can prove (i), and then (T) , since 'for every x , $x = x$ ' and 'it is not the case that any thing is both ambitious and not ambitious' are logical truths (counting the identity sign as a logical constant).

(i) ' (\neg 'Caesar' 'thrice refused the crown' 'and' 'Caesar' 'was ambitious' ')' is T iff Caesar thrice refused the crown and for every x , $x = x$ and Caesar was ambitious and it is not the case that any thing is both ambitious and not ambitious.

(T) ' (\neg 'Caesar' 'thrice refused the crown' 'and' 'Caesar' 'was ambitious' ')' is T iff Caesar thrice refused the crown and Caesar was ambitious.

Yet it is evident that we have not gone from axioms that show the structure and meaning of the object language sentence to the T -sentence, and so we have not revealed the compositional structure of the object language sentence in the proof.

7. Rather than sequences, I use functions from variables to objects as satisfiers. Tarski's sequences can be represented using sets of ordered pairs of positive integers and objects, the integers representing the order of the objects in the sequence. To associate an object with a variable in an open formula, Tarski associated variables in a predetermined order with the integers representing the order of objects in a sequence. This simply represents an assignment of objects to the variables, and we can dispense with the sequences, which can be seen to play merely a heuristic role in Tarski's discussion.
8. This shows, incidentally, that not even knowing that the theory is analytically true suffices to know that it is interpretive.
9. The class of theories we apply the convention to must be circumscribed so as to exclude introduction of extraneous materials that might cause difficulties. Thus, we require that theory be minimal in a certain sense: we want only axioms needed for giving satisfaction conditions for object language expressions and nothing not needed for this purpose, and we want the axioms that give satisfaction conditions not to include anything in them that

is not necessary for giving satisfaction conditions for object language expressions or sentence forms. Thus, for example, we would want one axiom for each primitive expression of the object language, and every satisfaction axiom should be a quantified biconditional; in addition to such axioms we would need only axioms for our theory of functions and reference axioms, which themselves would have a standard form (‘the referent of $A = x$ ’). Thus,

((for all functions f , f satisfies in Welsh ‘triongl yw x ’ iff $f(x)$ is a triangle) and $2 + 2 = 4$),

would not be an axiom of a truth theory of the form under consideration here. This could be spelled out more precisely, but it should be clear enough for present purposes what form of theory is intended.

10. There are some additional complications in Welsh I overlook here: ‘ac’ is the form required when what follows begins with a vowel. Strictly speaking, then, there should be a restriction on ψ in (5) to sentences or formulas that do not require the ‘ac’ form.
11. Convention S can be made more precise, of course, relative to a precise specification of the forms of axioms employed in a theory for a given kind of language.
12. Take any T -sentence, say (1) in the text, and any logical truth, λ . From these we can prove (1’).

(1’) ‘Gwyn’-‘yw’-‘eira’ is true in Welsh iff snow is white and λ .

See Soames 1992, 28 and Foster 1976 for the objection. It must be said that Davidson was aware of the need, and invoked the idea of a canonical proof procedure to meet it, without, however explaining exactly how we were to think of it. Perhaps he thought it was too obvious to be worth remark.

13. I include the qualification because there are reasons to think that the grammar of English is not finitely recursively specifiable. The reason I have in mind is that the quotation name of any symbol is a symbol of English, whether the symbol itself is or not, and it is doubtful that there is any way to recursively enumerate every possible symbol, since it is plausible that there are an infinite number of primitive symbols.
14. This meets the criticisms leveled by Loar 1976 and Foster 1976.
15. I am assuming that knowledge of what is stated here will suffice for understanding \mathfrak{M} ; if not, we will just add as much as we need to state knowledge sufficient to understand the language of the truth theory. This violates none of our constraints, since obviously understanding the language of the truth theory does not by itself suffice for understanding the object language.
16. The account given here can be extended to a generalization of the truth-theoretic approach to handle imperatives and interrogatives that are not assigned truth conditions but rather compliance conditions. See Ludwig 1997 for an outline of the approach.
17. Thus, it is not surprising that this proposal avoids objections to truth-theoretic semantics that presuppose that the truth theory itself is the meaning theory. Critics looking at the truth theory and wondering where the meaning theory was were looking in the wrong place.
18. I do not claim here that speaker and time determine by themselves the referents of demonstratives. The suggestion is that we can describe the referent of a demonstrative in terms of speaker and time; roughly, it is the object demonstrated by the speaker at the time using the demonstrative. In fact, there is one additional complexity in the case of demonstratives: we must make reference also to the speech act in which a demonstrative is used. This is necessary to accommodate the possibility of someone using a single token of a demonstrative ambiguously in two speech acts, directed, e.g., at different audiences at the same time, with different demonstrative intentions with respect to the different audiences (‘Bring me that’). With this in mind, we can give the following reference clause for simple demonstratives:

For all speakers s , times t , speech acts u , and objects x ,

if s demonstrates x at t using ‘that’ in u ,

then $\text{ref}_{[s, t, u]}(\text{‘that’}) = x$.

Here $\text{ref}_{[s, t, u]}(\text{'that'}) = x'$ is read 'the referent of 'that' as used by s at t in u '. The relativization to speech acts will require that our semantic predicates likewise be relativized to speech acts, but this additional relativization will not affect any of the points made in the text. The fact that we must conditionalize on a speaker's demonstrating something in the reference axioms for demonstratives means that we cannot discharge the reference axioms in proofs of T -sentences for object language sentences that contain demonstratives until we apply them to speakers using the object language sentence, so as to allow the antecedent to be satisfied by some object. See Lepore and Ludwig 2000, appendix, for further discussion.

19. For a discussion of why we cannot read it as ' x would be true if it were spoken by s at t ', see Evans 1985, 359–60.
20. Strictly, we should interpret ' s is . . . at t ' as a primitive metalanguage verb relating a speaker to a time. See Lepore and Ludwig 2000 for an explanation.
21. Recall from note 3 that we have shifted to thinking of truth as a primitive in our theories, and so treat it as relating sentences to languages.
22. This requires that the metalanguage contain the meaning predicate as well as the truth predicate. Of course, the exact form of the convention will vary depending on the metalanguage.
23. Again, we require the metalanguage to have such a predicate.
24. See Ludwig and Ray 2001.
25. See Ludwig and Ray 1998 for a sententialist account of such contexts that makes good on the claim that terms in that-clauses and similar contexts do not contribute their extensional properties to determining the truth conditions of containing sentences.
26. See Harman 1974 and Soames 1992. Soames claims, for example, that "the role of truth theories in specifying the knowledge that is supposed to be sufficient for understanding sentences is essentially heuristic, and in principle, dispensable" (27), for "all that is needed for the derivation is that we be provided with sentences of the form, ' S ' is F iff p , where the sentence replacing ' p ' is guaranteed to be a translation of the sentence replacing ' S '" (ibid.). He goes on to say that "beyond this, it is not important how these sentences are produced, what they say, or even whether they are true" (ibid.). Yet, as will be shown, it is important how these sentences are produced if we wish to accomplish the aim of a compositional meaning theory, even though, if the main thesis of this paper is correct, there is truth in the claim that it does not matter whether or not they are true, within certain limits. In fairness to Soames, it should be said that he is not taking the proponent of the utility of a truth theory in the theory of meaning to be aiming for more than is required to put someone in a position to interpret each sentence of another language.
27. It has been suggested to me that one could mimic in a translation theory the features of a truth theory that enable it to exhibit the meaning of, e.g., connectives in the language. Take conjunction as an example. Instead of 7, one could propose 7'.

7'. For any s_1, s_2 , of L , for any s_3, s_4 of L^* , $\text{Tr}('(\neg s_1 \neg \&' \neg s_2 \neg)'), (' \neg \text{Tr}(s_3) \neg \wedge' \neg \text{Tr}(s_4) \neg)'$ iff $\text{Tr}(s_1, s_3)$ and $\text{Tr}(s_2, s_4)$.

The suggestion is that this shows that '&' and '\wedge' mean 'and' in the same way that a recursive axiom for conjunction in a truth theory shows that the object language term means 'and' (see axiom A5 below). But this is illusory. 7' will be true iff '&' translates '\wedge' and the concatenation of a sentence of L with '&' and another sentence in L translates the concatenation of the translation of the first with '\wedge' with the translation of the second into L^* . Nothing else matters, and in particular it doesn't matter what '&' and '\wedge' mean: they could mean 'and', 'or', 'iff', 'only if', 'because', and so on, as long as they are sentential connectives. To learn more we must be told two things: what '&' or '\wedge' mean, and what the significance of their pattern of combination is in these sentences: it is the latter in particular that a translation theory will never give us any information about. It is also worth noting that the strategy suggested backfires when we turn to some other connectives, such as negation. For in the case of negation, the axiom mimicking the negation axiom for the truth theory would be 5'.

5'. For all s_1 of L , s_2 of L^* , $\text{Tr}(\sim \ulcorner s_1, \ulcorner \urcorner s_2)$ iff it is not the case that $\text{Tr}(s_1, s_2)$.

An embarrassing result, to be sure. The parallel axiom for disjunction yields a similarly odd result.

REFERENCES

- Chihara, C. S. 1976. "Truth, Meaning, and Paradox." *Nous* 10: 305–12.
- Davidson, D. 1984a. "Radical Interpretation." *Inquiries into Truth and Interpretation*. New York: Clarendon Press.
- . 1984b. "Reply to Foster." *Inquiries into Truth and Interpretation*. New York: Clarendon Press.
- . 1984c. "Semantics for Natural Languages." *Inquiries into Truth and Interpretation*. New York: Clarendon Press.
- . 1984d. "Truth and Meaning." *Inquiries into Truth and Interpretation*. New York: Clarendon Press.
- Evans, G. 1985. "Does Tense Logic Rest on a Mistake?" *Collected Papers*. Oxford: Clarendon Press.
- Evans, G., and J. McDowell, eds. 1976. *Truth and Meaning: Essays in Semantics*. Oxford: Clarendon Press.
- Foster, J. A. 1976. "Meaning and Truth Theory." In G. Evans and S. McDowell (eds.), *Truth and Meaning: Essays in Semantics*. Oxford: Clarendon Press.
- Harman, G. 1974. "Meaning and Semantics." In M. Munitz and P. Unger (eds.), *Semantics and Philosophy*. New York: New York University Press.
- Higginbotham, J. 1992. "Truth and Understanding." *Philosophical Studies* 65, 3.
- Lepore, E., and K. Ludwig. 2000. "The Semantics and Pragmatics of Complex Demonstratives." *Mind* 109: 199–240.
- . 2002. "Outline of a Truth Conditional Semantics for Tense." In Q. Smith (ed.), *Tense, Time and Reference*. Cambridge, Mass.: MIT Press.
- Loar, B. 1976. "Two Theories of Meaning." In G. Evans and J. McDowell (eds.), *Truth and Meaning: Essays in Semantics*. Oxford: Clarendon Press.
- Ludwig, K. 1997. "The Truth about Moods." *Protosociology* 10 (*Cognitive Semantics I - Conceptions of Meaning*): 19–66.
- Ludwig, K., and G. Ray. 1998. "Semantics for Opaque Contexts." *Philosophical Perspectives* 12: 141–66.
- . 2001. "Vagueness and the Sorites Paradox." Unpublished Manuscript.
- Richard, M. 1992. "Semantic Competence and Disquotational Knowledge." *Philosophical Studies* 65: 37–52.
- Soames, S. 1989. "Semantics and Semantic Competence." *Philosophical Perspectives*, 3: *Philosophy of Mind and Action Theory, 1989*. Atascadero, Calif.: Ridgeview.
- . 1992. "Truth, Meaning, and Understanding." *Philosophical Studies* 65: 17–36.
- Stich, S. 1976. "Davidson's Semantic Program." *Canadian Journal of Philosophy* 6: 201–27.
- Tarski, A. 1983. "The Concept of Truth in Formalized Languages." *Logic, Semantics, Metamathematics*, second edition. Indianapolis: Hackett.
- Wallace, J. 1978. "Logical Form, Meaning, Translation." In M. Guenther-Reutter (ed.), *Meaning and Translation*. London: Duckworth.

A New Argument against Modesty

Jonathan Sutton
Southern Methodist University

INTRODUCTION

MICHAEL DUMMETT (1975) argued that T-theories cannot serve as adequate theories of meaning as Davidson wished them to (Davidson 1984) since they do not explain what it is to possess the concepts that speakers associate with a language's primitive terms. Few have disputed Dummett's claim; what has been disputed is whether explaining possession of such concepts, on which speakers' semantic competence undoubtedly rests, is a proper task for a theory of meaning. The first section of this chapter reviews Dummett's criticism and the standard response to it. Although I agree with orthodoxy that Dummett fails to establish that T-theories are inadequate theories of meaning, I believe that there is another argument that yields Dummett's conclusion—constructing an adequate theory of meaning requires giving an account of what it is to possess many of the concepts that speakers associate with the primitive terms of their language, and T-theories as standardly conceived fail to give such an account. I will provide this argument, illuminating it by contrast it with Dummett's failed argument. Unlike Dummett's own argument, my argument does not rely on denying that thought can be understood independently of language.

The Dummettian conclusion turns out to be consistent with the claim that knowing a homophonic T-theory suffices for understanding English sentences, a feature of a theory of meaning that I call *epistemic adequacy*. However, homophonic T-theories fail to meet a further constraint on theories of meaning that I call *factual adequacy*.

DUMMETT'S ARGUMENT

Knowing a T-theory for English is alleged by Davidsonians to be sufficient for understanding all sentences of English (although not all *utterances* thereof, which will require pragmatic knowledge), and, indeed, by the bolder Davidsonians to

account for actual speakers' knowledge of the semantics of their language (e.g., Larson and Segal 1995).¹ A T-theory for English consists of base clauses that specify the semantic contribution that primitive expressions make to the semantic properties of complex expressions, and recursive clauses that specify how the semantic properties of complex expressions are determined by the semantic properties of their constituents. The most important complex expression is the sentence, and the T-theory enables a speaker to derive T-sentences for English of the form 'S is true iff *p*', knowledge of which suffices for understanding English sentences.² The account is sufficiently familiar by now that I will not explain it in any more detail. Our concern lies (as did Dummett's) entirely with the base clauses—we will not need to examine the recursive apparatus of a T-theory to draw the conclusions that we wish to draw.

As standardly conceived, a T-theory whose metalanguage and object language are identical will contain base clauses (and yield T-sentences) that *use* the terms for which they are clauses to provide the semantics of those same terms—it will be *homophonic*. (Of course, to cope with indexical and demonstrative expressions, a T-theory will have to depart from homophony. However, primitive expressions whose semantic value is not context sensitive are standardly thought to have homophonic base clauses; I shall ignore the complexities introduced by expressions with context-sensitive semantic values since they are not germane to my discussion.) If the metalanguage is distinct from the object language, a base clause for a term '*t*' will use what we pretheoretically regard as a translation of '*t*' into the metalanguage to provide the semantics of '*t*'. If a T-theory gives the semantics of primitive terms in the object language by using primitive terms of the metalanguage, I shall say that it has the *minimal structure property*, and when I use the term 'T-theory' I will mean a T-theory with the minimal structure property unless I state otherwise. (Homophonic T-theories are hence one kind of T-theory with the minimal structure property.) In the Conclusion, I shall extend the scope of my objections to cover T-theories that lack the minimal structure property, arguing that the only theories lacking the minimal structure property that avoid the problems that I raise are obviously full-blooded.³

One of the base clauses of a T-theory for English whose metalanguage is English might be '“snow” refers to snow'; that same theory stated in German will use the word '*Schnee*' and mention the word 'snow' in the counterpart base clause. (Nothing hangs on the particular example.) Part of what would suffice to understand English, then, is knowing that 'snow' refers to snow. In order to have such knowledge, an English speaker must be capable of entertaining the thought 'SNOW' REFERS TO SNOW, and so must possess the concept SNOW.⁴ It is reasonable to assume that a thinker possesses the concept SNOW in virtue of having acquired either a body of propositional knowledge (much or all of which will be about snow) or a set of practical abilities (knowledge *how to x* rather than knowledge *that p*) or both. A speaker of English could not under-

stand sentences containing the word ‘snow’ unless he had that knowledge, for if he lacked it, he would not possess the concept SNOW, and so could not know that ‘snow’ refers to snow. That knowledge, however, is not specified by a T-theory for English—as Dummett (1975, 112) puts it, a T-theory fails to specify what “constitutes” knowledge that ‘snow’ refers to snow.) Consequently, a T-theory for English fails to specify some knowledge that speakers rely on in interpreting English sentences, and so is crucially incomplete in its account of English speakers’ semantic competence—a T-theory is too *modest*, in Dummett’s words. A modest theory “gives[s] the interpretation of the language to someone who already has the concepts required [to understand the language].” Dummett’s argument is supposed to show that a theory of meaning must be “full-blooded”; such a theory “seeks actually to explain the concepts [such as SNOW] expressed by primitive terms of the language [such as ‘snow’]” (Dummett 1975, 102). For Dummett, a concept is individuated by its possession conditions, by what is required for a thinker to grasp it. Consequently, to “explain” a concept is to explain the propositional knowledge and cognitive abilities possessed by a thinker in virtue of which he possesses the concept.

The problem with Dummett’s argument, as is well-known, is that not all knowledge that speakers rely on in interpreting English sentences need be *semantic* knowledge. A speaker can rely on both his knowledge of the semantics of English and nonsemantic knowledge to understand English sentences. A theory of meaning should only be required to specify the properly semantic knowledge that English speakers *qua* English speakers possess. It is far from clear that the knowledge that underlies competence with the concept SNOW is semantic knowledge. To specify that knowledge would be to give an account of what it is to possess the concept SNOW, and that looks like part of a theory of *thought*, not a theory of *language*. It is hardly surprising that to understand language use properly, we need to understand thought; but that does not make a theory of concept possession part of a theory of meaning.

Dummett appreciates this point, I think, but it is not clear what his response to it is. It *is* clear that his response has something to do with the thesis that language is prior to thought. Whatever role that thesis plays in Dummett’s argument, it will play no role in my own argument.

A BETTER ARGUMENT AGAINST MODESTY

L—A Language without a T-theory Stateable in English

Constructing a theory of meaning for a language, like constructing any theory in any domain, will involve framing and deciding between competing hypotheses—in this case, hypotheses about the semantics of various linguistic constructions. If a semantic theorist claims that a certain kind of theory is of the right *form* to be a theory of meaning for a natural language, then debates about the

semantics of various constructions should be describable as debates over which particular theories of that form are correct. In particular, if a speaker's semantic competence is said to consist in knowledge of a T-theory, then competing accounts of the semantics of English should be expressible in terms of competing T-theories. I shall call this "the competition constraint" on proposals about the form of a theory of meaning. Actual debate about the semantics of various natural language constructions among theorists committed to the Davidsonian program does involve the submission and advocacy of rival T-theories or kinds of T-theory, and the competition constraint states that this is as it should be. Building on Dummett's failed argument, in this section I argue that T-theories with the minimal structure property stated in English or any other natural language violate the competition constraint. Further, it follows from this that there are facts genuinely concerning semantic competence that T-theories do not tell us.

At first blush, it seems that there could be a language, call it *L*, that demanded of its speakers that they know that proper names for people *are* names of people.⁵ We suppose that the vocabulary of *L* is character for character identical to that of English, and, *L*'s distinctive characteristics aside, there is great similarity in the semantic roles of the languages' lexical items.⁶ To understand 'Genghis Khan' in *L*,⁷ one must know that it refers to a person.⁸ A T-theory for *L* whose metalanguage is English will contain the base clause ' "Genghis Khan" refers to Genghis Khan', and so part of what suffices for understanding *L* is knowing that 'Genghis Khan' refers to Genghis Khan. *L* makes no similar demand on speakers' understanding of other proper names—one can understand 'arsenic' without knowing that it refers to a metal.

Will the T-theory for *L* specify that speakers of *L* must know that 'Genghis Khan' refers to a person? *Prima facie*, it will specify no such thing. The T-theory requires of speakers that they know that 'Genghis Khan' refers to Genghis Khan. It also requires of speakers that they know that 'arsenic' refers to arsenic. Since the presence of the 'arsenic' base clause does not entail that speakers know that 'arsenic' refers to a metal, the presence of the 'Genghis Khan' base clause cannot entail that speakers know that 'Genghis Khan' refers to a person. Therefore, there is an aspect of semantic competence in *L* that its T-theory does not capture, or so one might argue—the T-theory for *L* is inadequate, and inadequate because it does not explain the concept that one must associate with 'Genghis Khan' to understand it.

This is not the better argument against modesty that I advertised, for it suffers from the same defects as Dummett's argument. We have no right to assume that it is a (semantic) fact *about L* that speakers must know that names of persons refer to persons. It might be a fact about name *concepts* that refer to persons.

I argue that we can understand *L* in such a way that this story cannot be

the right explanation of what it takes to understand ‘Genghis Khan’ in *L*, and so construct a better argument against modesty.

Consider two name concepts that refer to Genghis Khan—call them GENGHIS_1 and GENGHIS_2 . GENGHIS_1 requires of its possessor that he know GENGHIS_1 IS A PERSON. GENGHIS_2 makes no such requirement—all it requires of its possessors is that they know, so to speak, *what object* Genghis_2 is, but not what kind of thing that object is. (My use of the term ‘concept’ here is rather loose; we need not assume that GENGHIS_1 and GENGHIS_2 are possible components of thoughts, although, for convenience, I will talk that way every now and then. We can take them to be distinct collections of beliefs about Khan, one of which contains the belief that Khan is a person.) We stipulate that *L* requires of its speakers that they associate GENGHIS_1 with ‘Genghis Khan’; other languages (such as English, perhaps) require only that speakers associate GENGHIS_2 with proper names of Genghis Khan. It remains a fact about *concepts* that GENGHIS_1 requires of its possessor that he know GENGHIS_1 IS A PERSON, and this is not the kind of fact that a T-theory must specify. But it is a *further fact about semantic competence in L* that speakers of *L* must associate GENGHIS_1 with ‘Genghis Khan’—associating GENGHIS_2 therewith is not enough to understand the name. Consequently, a theory of meaning for *L* must specify that fact if it is to be a full account of semantic competence in *L*.

I claim that a T-theory for *L* with the minimal structure property stated in a metalanguage that does not require its speakers to know that names of persons *are* names of persons does not specify that fact. Let us suppose, for convenience, that English is such a metalanguage. English can also be used to state its own T-theory. The base clause for ‘Genghis Khan’ in the T-theory for *L* with English as the metalanguage will be ‘“Genghis Khan” refers to Genghis Khan’. The base clause for the English homonym of *L*’s ‘Genghis Khan’ in the T-theory for *English* with English as the metalanguage will be ‘“Genghis Khan” refers to Genghis Khan’. Consequently, the ‘Genghis Khan’ base clause in *L*’s T-theory does not specify any knowledge required for understanding ‘Genghis Khan’ that the corresponding base clause in English’s T-theory does not require for understanding the English homonym of ‘Genghis Khan’. Both theories require only that speakers know that the name refers to Genghis Khan. The T-theory for English does not require speakers to know that ‘Genghis Khan’ refers to a person (because the T-theory is, we suppose, adequate and *that* is not required for speaking English), so neither can the T-theory for *L*.

Why T-theories with the Minimal Structure Property Stated in English Fail the Competition Constraint

English cannot express a T-theory adequate as a theory of meaning for *L*—so what? It does not obviously follow that English cannot express a T-theory that will serve as an adequate theory of meaning for English. For all that I have said

and will say in this chapter, English can express its own T-theory in the sense that one can state a T-theory for English in English knowledge of which would suffice for understanding English (and, in particular, names of persons in English). Let us call this conception of adequacy in a theory of meaning *epistemic adequacy*. An epistemically adequate theory of meaning for a language specifies all the genuinely semantic knowledge a speaker needs to understand the language.

Davidsonians regard epistemic adequacy as adequacy *tout court*. Dummett did not dispute that a T-theory could be epistemically adequate; he claimed that a truly adequate theory of meaning needs to explain the conceptual competence that linguistic understanding relies on, and a theory that is epistemically adequate need not and typically does not do that. I also argue that epistemic adequacy does not amount to adequacy *tout court* by distinguishing a second notion of adequacy which I shall call *factual adequacy*. A theory of meaning for a language is factually adequate just in case it specifies all facts genuinely about *semantic* competence in the language. For example, a factually adequate theory for a language will tell us whether a name for Genghis Khan must be understood as a name of a person or not, something we desire in a theory of meaning. I argue that a homophonic T-theory for English stated in English is not factually adequate by way of arguing that it fails the competition constraint. A factually adequate theory of meaning need not, *ipso facto*, be a theory that explicates the conceptual competence upon which semantic competence rests. Consequently, to demand that a theory of meaning be factually adequate is not to demand illegitimately that a theory of meaning incorporate a theory of thought—the standard criticism of Dummett’s argument will not apply to my argument.⁹

How do T-theories with the minimal structure property stated in English fall afoul of the competition constraint? One hypothesis about the semantics of English names is that they function as do names in *L*—to understand a name of a person, one must know that it is a name of a person. Debate over that hypothesis and the contrary hypothesis that understanding English names requires no such thing should, according to the competition constraint, be expressible as a debate over the merits of rival T-theories. Those rival T-theories cannot both be expressed in English.¹⁰ Call the T-theory that we need to express the hypothesis that understanding English names of persons requires knowing that they are names of persons ‘ T_1 ’. Call the T-theory that we need to express the hypothesis that understanding English names requires no such knowledge ‘ T_2 ’. Recall that T_1 and T_2 , by hypothesis, both have the minimal structure property and can be stated in English. If English is not *L*, then T_1 cannot be expressed in English, since T_1 would be a T-theory for *L*, which cannot be expressed in a language that does not require those who understand its names of persons to know that they are names of persons, as we have already argued. If English is *L*, then we cannot express T_2 . The base clause for ‘Genghis Khan’ in any T-theory

with the minimal structure property whose metalanguage is English will be ‘“Genghis Khan” refers to Genghis Khan’. If English is L , knowing a T-theory that includes the base clause ‘“Genghis Khan” refers to Genghis Khan’ will entail knowing that ‘Genghis Khan’ refers to a person since knowledge of such a base clause (along with its containing T-theory) is supposed to provide understanding of the name. (If the entailment did not hold, then a homophonic T-theory would not be epistemically adequate: if knowing a T-theory that includes the base clause ‘“Genghis Khan” refers to Genghis Khan’ did not entail knowing that ‘Genghis Khan’ refers to a person, then we could not express T_1 in English while retaining the minimal structure property.) Hence a base clause of that form will not be appropriate for inclusion in T_2 since it will not adequately represent what is required to understand a name on the hypothesis that T_2 is supposed to express. Since all natural languages are L -like or not L -like, we can generalize our conclusion to the claim that T-theories with the minimal structure property stated in *any* natural language violate the competition constraint.

The fact that we might, for all we have said, be able to state an epistemically adequate homophonic T-theory for English in English even though such T-theories violate the competition constraint has a peculiar consequence. Even when one has an epistemically adequate homophonic T-theory for English in hand and one knows it to be such, one will not thereby be in a position to say whether understanding names of persons in English requires knowing that they are names of persons. For example, one will not be in a position to know whether knowing the T-theory entails knowing that ‘Genghis Khan’ refers to a person—even though, for all we have said, there will be a fact of the matter about whether that entailment holds. This is so precisely because producing a homophonic T-theory for English is not enough to settle the question of whether English is L ; one’s homophonic T-theory for English will be either T_1 or T_2 , but one will not know which it is, and so not know whether English is or is not L . And yet it is a fact *about semantic competence with English names* that one must, or need not, associate a concept like $GENGHIS_1$ with ‘Genghis Khan’ to understand it. There are facts concerning genuinely semantic competence that T-theories do not tell us. A homophonic T-theory for English is not *factually* adequate even though it is epistemically adequate.

CONCLUSION

So, T-theories with the minimal structure property whose metalanguage is English do not meet the competition constraint and hence are not factually adequate. Are there other varieties of T-theory that do meet the constraint? One might suggest that T_1 incorporate the base clause ‘“Genghis Khan” refers to x iff $x =$ Genghis Khan and x is a person’ and that T_2 incorporate the base clause ‘“Genghis Khan” refers to x iff $x =$ Genghis Khan’. This proposal accepts that some T-theories, such as T_1 , do not have the minimal structure property. In do-

ing so, it admits that a theory of meaning must *explain* some of the ways of thinking of referents (concepts, as Dummett would say) that language users associate with the primitive terms of their language and in virtue of which they understand those terms; for example, T_1 explicitly states that speakers understand ‘Genghis Khan’ by thinking of Khan as a person. Consequently, it is a concession to Dummett’s view that theories of meaning should be full-blooded.

However, it is perhaps not enough of a concession to Dummett. T_1 and T_2 express the hypotheses that English is L and that English is not L , respectively, only if English is not in fact L (that is, only if T_1 is an [epistemically] inadequate T-theory for English—inadequate in that it requires *too much* of English speakers). Assume, however, that English is L . Assume further that the truth of a knowledge ascription of the form ‘ S knows that . . . Genghis Khan . . .’ requires that S have a concept of Khan that is sufficient for understanding ‘Genghis Khan’ in English (provided an English speaker associates that concept with the name) and that that concept figure in S ’s state of knowledge that makes the ascription true—let us call this ‘The Fregean Assumption’. (Note that this assumption allows that such a knowledge ascription can be true of someone who does not speak English and is hence not acquainted with the name ‘Genghis Khan’.) This assumption is not uncontroversial, but it seems to be needed to justify the claim that homophonic T-theories (such as T_2) are epistemically adequate—in particular, the claim that knowledge of a T-theory containing ‘“Genghis Khan” refers to Genghis Khan’ could suffice for understanding ‘Genghis Khan’. For if the assumption is false, it looks like one could have knowledge of that base clause without understanding ‘Genghis Khan’.

Proceeding on these assumptions, a speaker knows that ‘Genghis Khan’ refers to x iff $x =$ Genghis Khan *just in case* the speaker knows that ‘Genghis Khan’ refers to x iff $x =$ Genghis Khan and x is a person. The right-to-left component of the biconditional is trivial. To see that the left-to-right component is true, observe that a speaker who knows that ‘Genghis Khan’ refers to x iff $x =$ Genghis Khan has a concept of Khan that suffices for understanding ‘Genghis Khan’, and hence a concept of Khan *as* a person, and it is that concept that figures in the knowledge under discussion. Consequently, the speaker will know that if $x =$ Genghis Khan, then x is a person, and our claim is established modulo elementary logical competence on the part of the speaker. We can now conclude that if English is L , a speaker knows T_2 just in case the speaker knows T_1 . T_2 fails to express the hypothesis that English is *not* L , since a speaker who knows T_2 *would* know that names of persons are names of persons. The competition constraint remains unsatisfied.

Further, we do not yet have a way of meeting the competition constraint even if English is not L . The constraint does not demand simply that we express the competing hypotheses that English is L and that English is not L in the form of competing T-theories; it requires that we can formulate the *debate* over these

hypotheses in the form of a debate over competing T-theories. This requires that the competing hypotheses be expressible as T-theories however the debate turns out—whether or not English is *L*. Abandoning the minimal structure property in the limited way that T_1 does is not sufficient to meet the constraint.

In closing, I wish to suggest that not only will a factually adequate truth theory be a full-blooded theory of meaning that departs greatly from homophony, but that we have good reason to be pessimistic that such a theory can be stated in a natural language.

To what extent must a T-theory depart from the minimal structure property to be factually adequate, supposing still that such a theory can be stated in a natural language? Perhaps we can raise our central worry about the concept of a person that a speaker must have to know an epistemically adequate theory incorporating the base clause ‘Genghis Khan’ refers to x iff $x =$ Genghis Khan and x is a person’. If the rest of the theory is largely homophonic, it will contain a clause along the lines of ‘ $(\forall x)$ (x satisfies “is a person” iff x is a person)’. Perhaps there are several concepts of a person we might hypothesize that a speaker must associate with the term ‘person’ to understand that term. The theory will not tell us which of those concepts is required to understand either the ‘person’ base clause or the ‘Genghis Khan’ base clause (given the Fregean Assumption, it is natural to suppose that the same concept of a person is required for understanding each clause). Adding structure to the right-hand sides of base clauses will produce a factually adequate T-theory only if all terms employed in right-hand sides throughout the theory are such that there is only one concept that, when associated by a speaker with the term, could (in the sense of epistemic possibility—see footnote 10) suffice for understanding it (we can even allow that ‘person’ is such a term). Consequently, my central argument has as its target not only (largely) homophonic T-theories, but any T-theory stated in a natural language that does not, so to speak, lay bare the structure of all concepts that a competent speaker associates with a language’s primitive terms. For some primitive terms, laying bare conceptual structure can be accomplished with a homophonic base clause; for many, such as names, it cannot.

This is a very strong requirement. Given the competition constraint, it is close to supposing that there is a set of primitive concepts *expressible in a natural language* from which all concepts required for understanding the primitive terms of any language can be constructed. Such a hypothesis is certainly not without precedent—atomic empiricists and others (including, perhaps, some lexical semanticists) have endorsed some such claim. A T-theory whose base clauses are specified in the indicated fashion is surely full-blooded, however—it goes as far as one can in explicating the concepts possession of which is required for understanding a language’s primitive terms. Moreover, it is not at all obvious that any natural language will contain a set of primitive terms that will suffice for framing such a factually adequate theory. That is, it is not at all obvious

that we will not encounter the same problem we encountered above in trying to frame a T-theory for a non-*L*-like language in *L*. Names in *L* are too *rich* in what is required to understand them to use them on the right-hand sides of base clauses in a T-theory for a language that does not require speakers to know that names of persons are names of persons. There is no guarantee that an *L*-like language will contain any other terms, primitive or complex, that can express the hypothesis that understanding ‘Genghis Khan’ requires a speaker to know only *which* object he is, and not what kind of object. Perhaps every natural language will contain some primitive term for which one cannot frame a base clause expressible in the language itself or any other natural language that reveals enough “conceptual structure” to figure in a factually adequate T-theory. There is no *a priori* reason to think that a natural language can express its own semantics in the form of a factually adequate theory of meaning.

NOTES

Thanks are due to the audience at the Inland Northwest Philosophy Conference 2000 on Truth and Meaning for their helpful comments; I am particularly grateful to Peter Ludlow, Mark McCullagh, and Kirk Ludwig.

1. I will often speak as though it is the bolder Davidsonian claim that is under scrutiny, but my arguments are easily reformulated to apply to the weaker Davidsonian claim.
I take it that, strictly speaking, even the meeker Davidsonians would accept that knowing a T-theory for English is not merely sufficient for understanding English, but that one who knew such a theory would understand English in virtue of that knowledge.
2. Davidsonians typically acknowledge that a speaker must also know that the T-sentences are consequences of a T-theory in light of the concerns raised by Foster (1976); in some sense, speakers must know not only a T-theory, but also that what they know is a T-theory. There are several ways to explicate the extra knowledge required for semantic competence. One might claim that it is sufficient for a speaker to know in addition that the theory conforms to certain “formal and empirical constraints” such as the canons of radical interpretation (Davidson 1976). Or, one might claim that we need not suppose that a speaker has any propositional knowledge beyond knowledge of the T-theory—he is able to interpret utterances of his language because his knowledge of a T-theory is localized within the mind’s “semantics module” (Larson and Segal 1995, 34–42). These suggestions are all concerned with establishing the epistemic adequacy of T-theories. My argument allows that T-theories are epistemically adequate, and so it is not important for my purposes to decide among these suggestions. I will speak for simplicity as though knowledge of a T-theory is supposed to suffice for semantic competence in a language; my argument is unaffected if a stronger claim prompted by Foster’s problem is substituted throughout. None of the responses to Foster’s problem that appear in the literature are of any obvious help in establishing that T-theories as standardly conceived are factually adequate (nor are they intended to provide such help).
3. For convenience, I use the term ‘T-theory’ ambiguously to denote either a theory expressed in written or spoken form by using a particular metalanguage—that is, the content of such an expression—or the expression itself. In the former sense of ‘T-theory’, the theory itself can be expressed in indefinitely many metalanguages and hence homophony is a property of a particular expression of a theory, not of the theory itself. (In that sense of ‘T-theory’, I will mean ‘T-theory that is stateable homophonically’ by ‘T-theory’ unless I indicate otherwise.) In the latter sense, we can simply talk of homophonic truth theories. Which sense I have in mind should be clear from context.
4. I follow the convention of using uppercase letters to denote concepts and thoughts, and will be sloppy about use and mention since which is required should be clear from context.

5. In footnote 10, I suggest that the assumption that *L* is a possible language is dispensable. My argument proceeds on the basis of how names are understood in a hypothetical language simply because of the prominence of names in various debates in the philosophy of language. Other categories of primitive term would serve my purpose equally well.
6. I very much doubt that *L* is English. ‘Genghis Khan was a mindless automaton controlled from Mars’ could be true, and that is very good evidence against the hypothesis that *L* is English; if understanding ‘Genghis Khan’ required believing that any bearer it has is a person, then ‘Genghis Khan is F’ should entail ‘Genghis Khan is a person’ for more or less any ‘F’.
7. I will drop the qualifier ‘in *L*’ for the remainder of this section; ‘Genghis Khan’ is a name of the name for Khan in *L* and not a name of the English name. Uses of the homonym are uses of the English name; this paper is not written in *L*.
8. Or perhaps: it refers to a person if it refers at all.
9. Besides, a theory of thought in conjunction with an epistemically adequate theory of meaning would not necessarily amount to a factually adequate theory of meaning. A complete theory of thought might explicate all possible concepts of Genghis Khan, but it would not explain which of those concepts one must associate with names of Khan to understand them in a given language—how one must think of Khan to understand a name for him. I will argue that epistemically adequate T-theories with the minimal structure property stated in English or another natural language will not provide that explanation either. If neither theory provides the needed explanation of name comprehension, nor does their conjunction.
10. It is here that we can dispense with the assumption that *L* is a possible language in any sense but the epistemic. Even if no language could contain names that are understood as the names of *L* are alleged to be, the contrary hypothesis makes sense—there could be genuine debate over its truth-value. And so T-theories for *L*-like languages should be formulable by the competition constraint. From this it follows that there is nothing in my argument with which a direct reference theorist should disagree. It takes more to understand names in *L* than a direct reference theorist is likely to accept is needed to understand names in actual languages. Perhaps the direct reference theory is true of names in all possible languages. It remains the case that the contrary hypothesis makes sense, and so we should be able to formulate T-theories that express the contrary hypothesis.

REFERENCES

- Davidson, D. 1976. “Reply to Foster.” In G. Evans and J. McDowell (eds.), *Truth and Meaning*. Oxford: Oxford University Press.
- . 1984. “Truth and Meaning.” *Inquiries into Truth and Interpretation*. Oxford: Oxford University Press.
- Dummett, M. 1975. “What Is a Theory of Meaning? (I).” In S. Guttenplan (ed.), *Mind and Language*. Oxford: Oxford University Press.
- Evans, G., and J. McDowell, eds. 1976. *Truth and Meaning*. Oxford: Oxford University Press.
- Foster, J. 1976. “Meaning and Truth Theory.” In G. Evans and J. McDowell (eds.), *Truth and Meaning*. Oxford: Oxford University Press.
- Larson, R., and G. Segal. 1995. *Knowledge of Meaning*. Cambridge, Mass.: MIT Press.

A New Argument against Modesty

Jonathan Sutton
Southern Methodist University

INTRODUCTION

MICHAEL DUMMETT (1975) argued that T-theories cannot serve as adequate theories of meaning as Davidson wished them to (Davidson 1984) since they do not explain what it is to possess the concepts that speakers associate with a language's primitive terms. Few have disputed Dummett's claim; what has been disputed is whether explaining possession of such concepts, on which speakers' semantic competence undoubtedly rests, is a proper task for a theory of meaning. The first section of this chapter reviews Dummett's criticism and the standard response to it. Although I agree with orthodoxy that Dummett fails to establish that T-theories are inadequate theories of meaning, I believe that there is another argument that yields Dummett's conclusion—constructing an adequate theory of meaning requires giving an account of what it is to possess many of the concepts that speakers associate with the primitive terms of their language, and T-theories as standardly conceived fail to give such an account. I will provide this argument, illuminating it by contrast it with Dummett's failed argument. Unlike Dummett's own argument, my argument does not rely on denying that thought can be understood independently of language.

The Dummettian conclusion turns out to be consistent with the claim that knowing a homophonic T-theory suffices for understanding English sentences, a feature of a theory of meaning that I call *epistemic adequacy*. However, homophonic T-theories fail to meet a further constraint on theories of meaning that I call *factual adequacy*.

DUMMETT'S ARGUMENT

Knowing a T-theory for English is alleged by Davidsonians to be sufficient for understanding all sentences of English (although not all *utterances* thereof, which will require pragmatic knowledge), and, indeed, by the bolder Davidsonians to

account for actual speakers' knowledge of the semantics of their language (e.g., Larson and Segal 1995).¹ A T-theory for English consists of base clauses that specify the semantic contribution that primitive expressions make to the semantic properties of complex expressions, and recursive clauses that specify how the semantic properties of complex expressions are determined by the semantic properties of their constituents. The most important complex expression is the sentence, and the T-theory enables a speaker to derive T-sentences for English of the form 'S is true iff *p*', knowledge of which suffices for understanding English sentences.² The account is sufficiently familiar by now that I will not explain it in any more detail. Our concern lies (as did Dummett's) entirely with the base clauses—we will not need to examine the recursive apparatus of a T-theory to draw the conclusions that we wish to draw.

As standardly conceived, a T-theory whose metalanguage and object language are identical will contain base clauses (and yield T-sentences) that *use* the terms for which they are clauses to provide the semantics of those same terms—it will be *homophonic*. (Of course, to cope with indexical and demonstrative expressions, a T-theory will have to depart from homophony. However, primitive expressions whose semantic value is not context sensitive are standardly thought to have homophonic base clauses; I shall ignore the complexities introduced by expressions with context-sensitive semantic values since they are not germane to my discussion.) If the metalanguage is distinct from the object language, a base clause for a term '*t*' will use what we pretheoretically regard as a translation of '*t*' into the metalanguage to provide the semantics of '*t*'. If a T-theory gives the semantics of primitive terms in the object language by using primitive terms of the metalanguage, I shall say that it has the *minimal structure property*, and when I use the term 'T-theory' I will mean a T-theory with the minimal structure property unless I state otherwise. (Homophonic T-theories are hence one kind of T-theory with the minimal structure property.) In the Conclusion, I shall extend the scope of my objections to cover T-theories that lack the minimal structure property, arguing that the only theories lacking the minimal structure property that avoid the problems that I raise are obviously full-blooded.³

One of the base clauses of a T-theory for English whose metalanguage is English might be '“snow” refers to snow'; that same theory stated in German will use the word '*Schnee*' and mention the word 'snow' in the counterpart base clause. (Nothing hangs on the particular example.) Part of what would suffice to understand English, then, is knowing that 'snow' refers to snow. In order to have such knowledge, an English speaker must be capable of entertaining the thought 'SNOW' REFERS TO SNOW, and so must possess the concept SNOW.⁴ It is reasonable to assume that a thinker possesses the concept SNOW in virtue of having acquired either a body of propositional knowledge (much or all of which will be about snow) or a set of practical abilities (knowledge *how to x* rather than knowledge *that p*) or both. A speaker of English could not under-

stand sentences containing the word ‘snow’ unless he had that knowledge, for if he lacked it, he would not possess the concept SNOW, and so could not know that ‘snow’ refers to snow. That knowledge, however, is not specified by a T-theory for English—as Dummett (1975, 112) puts it, a T-theory fails to specify what “constitutes” knowledge that ‘snow’ refers to snow.) Consequently, a T-theory for English fails to specify some knowledge that speakers rely on in interpreting English sentences, and so is crucially incomplete in its account of English speakers’ semantic competence—a T-theory is too *modest*, in Dummett’s words. A modest theory “gives[s] the interpretation of the language to someone who already has the concepts required [to understand the language].” Dummett’s argument is supposed to show that a theory of meaning must be “full-blooded”; such a theory “seeks actually to explain the concepts [such as SNOW] expressed by primitive terms of the language [such as ‘snow’]” (Dummett 1975, 102). For Dummett, a concept is individuated by its possession conditions, by what is required for a thinker to grasp it. Consequently, to “explain” a concept is to explain the propositional knowledge and cognitive abilities possessed by a thinker in virtue of which he possesses the concept.

The problem with Dummett’s argument, as is well-known, is that not all knowledge that speakers rely on in interpreting English sentences need be *semantic* knowledge. A speaker can rely on both his knowledge of the semantics of English and nonsemantic knowledge to understand English sentences. A theory of meaning should only be required to specify the properly semantic knowledge that English speakers *qua* English speakers possess. It is far from clear that the knowledge that underlies competence with the concept SNOW is semantic knowledge. To specify that knowledge would be to give an account of what it is to possess the concept SNOW, and that looks like part of a theory of *thought*, not a theory of *language*. It is hardly surprising that to understand language use properly, we need to understand thought; but that does not make a theory of concept possession part of a theory of meaning.

Dummett appreciates this point, I think, but it is not clear what his response to it is. It *is* clear that his response has something to do with the thesis that language is prior to thought. Whatever role that thesis plays in Dummett’s argument, it will play no role in my own argument.

A BETTER ARGUMENT AGAINST MODESTY

L—A Language without a T-theory Stateable in English

Constructing a theory of meaning for a language, like constructing any theory in any domain, will involve framing and deciding between competing hypotheses—in this case, hypotheses about the semantics of various linguistic constructions. If a semantic theorist claims that a certain kind of theory is of the right *form* to be a theory of meaning for a natural language, then debates about the

semantics of various constructions should be describable as debates over which particular theories of that form are correct. In particular, if a speaker's semantic competence is said to consist in knowledge of a T-theory, then competing accounts of the semantics of English should be expressible in terms of competing T-theories. I shall call this "the competition constraint" on proposals about the form of a theory of meaning. Actual debate about the semantics of various natural language constructions among theorists committed to the Davidsonian program does involve the submission and advocacy of rival T-theories or kinds of T-theory, and the competition constraint states that this is as it should be. Building on Dummett's failed argument, in this section I argue that T-theories with the minimal structure property stated in English or any other natural language violate the competition constraint. Further, it follows from this that there are facts genuinely concerning semantic competence that T-theories do not tell us.

At first blush, it seems that there could be a language, call it *L*, that demanded of its speakers that they know that proper names for people *are* names of people.⁵ We suppose that the vocabulary of *L* is character for character identical to that of English, and, *L*'s distinctive characteristics aside, there is great similarity in the semantic roles of the languages' lexical items.⁶ To understand 'Genghis Khan' in *L*,⁷ one must know that it refers to a person.⁸ A T-theory for *L* whose metalanguage is English will contain the base clause ' "Genghis Khan" refers to Genghis Khan', and so part of what suffices for understanding *L* is knowing that 'Genghis Khan' refers to Genghis Khan. *L* makes no similar demand on speakers' understanding of other proper names—one can understand 'arsenic' without knowing that it refers to a metal.

Will the T-theory for *L* specify that speakers of *L* must know that 'Genghis Khan' refers to a person? *Prima facie*, it will specify no such thing. The T-theory requires of speakers that they know that 'Genghis Khan' refers to Genghis Khan. It also requires of speakers that they know that 'arsenic' refers to arsenic. Since the presence of the 'arsenic' base clause does not entail that speakers know that 'arsenic' refers to a metal, the presence of the 'Genghis Khan' base clause cannot entail that speakers know that 'Genghis Khan' refers to a person. Therefore, there is an aspect of semantic competence in *L* that its T-theory does not capture, or so one might argue—the T-theory for *L* is inadequate, and inadequate because it does not explain the concept that one must associate with 'Genghis Khan' to understand it.

This is not the better argument against modesty that I advertised, for it suffers from the same defects as Dummett's argument. We have no right to assume that it is a (semantic) fact *about L* that speakers must know that names of persons refer to persons. It might be a fact about name *concepts* that refer to persons.

I argue that we can understand *L* in such a way that this story cannot be

the right explanation of what it takes to understand ‘Genghis Khan’ in *L*, and so construct a better argument against modesty.

Consider two name concepts that refer to Genghis Khan—call them GENGHIS₁ and GENGHIS₂. GENGHIS₁ requires of its possessor that he know GENGHIS₁ IS A PERSON. GENGHIS₂ makes no such requirement—all it requires of its possessors is that they know, so to speak, *what object* Genghis₂ is, but not what kind of thing that object is. (My use of the term ‘concept’ here is rather loose; we need not assume that GENGHIS₁ and GENGHIS₂ are possible components of thoughts, although, for convenience, I will talk that way every now and then. We can take them to be distinct collections of beliefs about Khan, one of which contains the belief that Khan is a person.) We stipulate that *L* requires of its speakers that they associate GENGHIS₁ with ‘Genghis Khan’; other languages (such as English, perhaps) require only that speakers associate GENGHIS₂ with proper names of Genghis Khan. It remains a fact about *concepts* that GENGHIS₁ requires of its possessor that he know GENGHIS₁ IS A PERSON, and this is not the kind of fact that a T-theory must specify. But it is a *further fact about semantic competence in L* that speakers of *L* must associate GENGHIS₁ with ‘Genghis Khan’—associating GENGHIS₂ therewith is not enough to understand the name. Consequently, a theory of meaning for *L* must specify that fact if it is to be a full account of semantic competence in *L*.

I claim that a T-theory for *L* with the minimal structure property stated in a metalanguage that does not require its speakers to know that names of persons *are* names of persons does not specify that fact. Let us suppose, for convenience, that English is such a metalanguage. English can also be used to state its own T-theory. The base clause for ‘Genghis Khan’ in the T-theory for *L* with English as the metalanguage will be ‘“Genghis Khan” refers to Genghis Khan’. The base clause for the English homonym of *L*’s ‘Genghis Khan’ in the T-theory for *English* with English as the metalanguage will be ‘“Genghis Khan” refers to Genghis Khan’. Consequently, the ‘Genghis Khan’ base clause in *L*’s T-theory does not specify any knowledge required for understanding ‘Genghis Khan’ that the corresponding base clause in English’s T-theory does not require for understanding the English homonym of ‘Genghis Khan’. Both theories require only that speakers know that the name refers to Genghis Khan. The T-theory for English does not require speakers to know that ‘Genghis Khan’ refers to a person (because the T-theory is, we suppose, adequate and *that* is not required for speaking English), so neither can the T-theory for *L*.

Why T-theories with the Minimal Structure Property Stated in English Fail the Competition Constraint

English cannot express a T-theory adequate as a theory of meaning for *L*—so what? It does not obviously follow that English cannot express a T-theory that will serve as an adequate theory of meaning for English. For all that I have said

and will say in this chapter, English can express its own T-theory in the sense that one can state a T-theory for English in English knowledge of which would suffice for understanding English (and, in particular, names of persons in English). Let us call this conception of adequacy in a theory of meaning *epistemic adequacy*. An epistemically adequate theory of meaning for a language specifies all the genuinely semantic knowledge a speaker needs to understand the language.

Davidsonians regard epistemic adequacy as adequacy *tout court*. Dummett did not dispute that a T-theory could be epistemically adequate; he claimed that a truly adequate theory of meaning needs to explain the conceptual competence that linguistic understanding relies on, and a theory that is epistemically adequate need not and typically does not do that. I also argue that epistemic adequacy does not amount to adequacy *tout court* by distinguishing a second notion of adequacy which I shall call *factual adequacy*. A theory of meaning for a language is factually adequate just in case it specifies all facts genuinely about *semantic* competence in the language. For example, a factually adequate theory for a language will tell us whether a name for Genghis Khan must be understood as a name of a person or not, something we desire in a theory of meaning. I argue that a homophonic T-theory for English stated in English is not factually adequate by way of arguing that it fails the competition constraint. A factually adequate theory of meaning need not, *ipso facto*, be a theory that explicates the conceptual competence upon which semantic competence rests. Consequently, to demand that a theory of meaning be factually adequate is not to demand illegitimately that a theory of meaning incorporate a theory of thought—the standard criticism of Dummett’s argument will not apply to my argument.⁹

How do T-theories with the minimal structure property stated in English fall afoul of the competition constraint? One hypothesis about the semantics of English names is that they function as do names in *L*—to understand a name of a person, one must know that it is a name of a person. Debate over that hypothesis and the contrary hypothesis that understanding English names requires no such thing should, according to the competition constraint, be expressible as a debate over the merits of rival T-theories. Those rival T-theories cannot both be expressed in English.¹⁰ Call the T-theory that we need to express the hypothesis that understanding English names of persons requires knowing that they are names of persons ‘ T_1 ’. Call the T-theory that we need to express the hypothesis that understanding English names requires no such knowledge ‘ T_2 ’. Recall that T_1 and T_2 , by hypothesis, both have the minimal structure property and can be stated in English. If English is not *L*, then T_1 cannot be expressed in English, since T_1 would be a T-theory for *L*, which cannot be expressed in a language that does not require those who understand its names of persons to know that they are names of persons, as we have already argued. If English is *L*, then we cannot express T_2 . The base clause for ‘Genghis Khan’ in any T-theory

with the minimal structure property whose metalanguage is English will be ‘“Genghis Khan” refers to Genghis Khan’. If English is L , knowing a T-theory that includes the base clause ‘“Genghis Khan” refers to Genghis Khan’ will entail knowing that ‘Genghis Khan’ refers to a person since knowledge of such a base clause (along with its containing T-theory) is supposed to provide understanding of the name. (If the entailment did not hold, then a homophonic T-theory would not be epistemically adequate: if knowing a T-theory that includes the base clause ‘“Genghis Khan” refers to Genghis Khan’ did not entail knowing that ‘Genghis Khan’ refers to a person, then we could not express T_1 in English while retaining the minimal structure property.) Hence a base clause of that form will not be appropriate for inclusion in T_2 since it will not adequately represent what is required to understand a name on the hypothesis that T_2 is supposed to express. Since all natural languages are L -like or not L -like, we can generalize our conclusion to the claim that T-theories with the minimal structure property stated in *any* natural language violate the competition constraint.

The fact that we might, for all we have said, be able to state an epistemically adequate homophonic T-theory for English in English even though such T-theories violate the competition constraint has a peculiar consequence. Even when one has an epistemically adequate homophonic T-theory for English in hand and one knows it to be such, one will not thereby be in a position to say whether understanding names of persons in English requires knowing that they are names of persons. For example, one will not be in a position to know whether knowing the T-theory entails knowing that ‘Genghis Khan’ refers to a person—even though, for all we have said, there will be a fact of the matter about whether that entailment holds. This is so precisely because producing a homophonic T-theory for English is not enough to settle the question of whether English is L ; one’s homophonic T-theory for English will be either T_1 or T_2 , but one will not know which it is, and so not know whether English is or is not L . And yet it is a fact *about semantic competence with English names* that one must, or need not, associate a concept like $GENGHIS_1$ with ‘Genghis Khan’ to understand it. There are facts concerning genuinely semantic competence that T-theories do not tell us. A homophonic T-theory for English is not *factually* adequate even though it is epistemically adequate.

CONCLUSION

So, T-theories with the minimal structure property whose metalanguage is English do not meet the competition constraint and hence are not factually adequate. Are there other varieties of T-theory that do meet the constraint? One might suggest that T_1 incorporate the base clause ‘“Genghis Khan” refers to x iff $x =$ Genghis Khan and x is a person’ and that T_2 incorporate the base clause ‘“Genghis Khan” refers to x iff $x =$ Genghis Khan’. This proposal accepts that some T-theories, such as T_1 , do not have the minimal structure property. In do-

ing so, it admits that a theory of meaning must *explain* some of the ways of thinking of referents (concepts, as Dummett would say) that language users associate with the primitive terms of their language and in virtue of which they understand those terms; for example, T_1 explicitly states that speakers understand ‘Genghis Khan’ by thinking of Khan as a person. Consequently, it is a concession to Dummett’s view that theories of meaning should be full-blooded.

However, it is perhaps not enough of a concession to Dummett. T_1 and T_2 express the hypotheses that English is L and that English is not L , respectively, only if English is not in fact L (that is, only if T_1 is an [epistemically] inadequate T-theory for English—inadequate in that it requires *too much* of English speakers). Assume, however, that English is L . Assume further that the truth of a knowledge ascription of the form ‘ S knows that . . . Genghis Khan . . .’ requires that S have a concept of Khan that is sufficient for understanding ‘Genghis Khan’ in English (provided an English speaker associates that concept with the name) and that that concept figure in S ’s state of knowledge that makes the ascription true—let us call this ‘The Fregean Assumption’. (Note that this assumption allows that such a knowledge ascription can be true of someone who does not speak English and is hence not acquainted with the name ‘Genghis Khan’.) This assumption is not uncontroversial, but it seems to be needed to justify the claim that homophonic T-theories (such as T_2) are epistemically adequate—in particular, the claim that knowledge of a T-theory containing ‘“Genghis Khan” refers to Genghis Khan’ could suffice for understanding ‘Genghis Khan’. For if the assumption is false, it looks like one could have knowledge of that base clause without understanding ‘Genghis Khan’.

Proceeding on these assumptions, a speaker knows that ‘Genghis Khan’ refers to x iff $x =$ Genghis Khan *just in case* the speaker knows that ‘Genghis Khan’ refers to x iff $x =$ Genghis Khan and x is a person. The right-to-left component of the biconditional is trivial. To see that the left-to-right component is true, observe that a speaker who knows that ‘Genghis Khan’ refers to x iff $x =$ Genghis Khan has a concept of Khan that suffices for understanding ‘Genghis Khan’, and hence a concept of Khan *as* a person, and it is that concept that figures in the knowledge under discussion. Consequently, the speaker will know that if $x =$ Genghis Khan, then x is a person, and our claim is established modulo elementary logical competence on the part of the speaker. We can now conclude that if English is L , a speaker knows T_2 just in case the speaker knows T_1 . T_2 fails to express the hypothesis that English is *not* L , since a speaker who knows T_2 *would* know that names of persons are names of persons. The competition constraint remains unsatisfied.

Further, we do not yet have a way of meeting the competition constraint even if English is not L . The constraint does not demand simply that we express the competing hypotheses that English is L and that English is not L in the form of competing T-theories; it requires that we can formulate the *debate* over these

hypotheses in the form of a debate over competing T-theories. This requires that the competing hypotheses be expressible as T-theories however the debate turns out—whether or not English is *L*. Abandoning the minimal structure property in the limited way that T_1 does is not sufficient to meet the constraint.

In closing, I wish to suggest that not only will a factually adequate truth theory be a full-blooded theory of meaning that departs greatly from homophony, but that we have good reason to be pessimistic that such a theory can be stated in a natural language.

To what extent must a T-theory depart from the minimal structure property to be factually adequate, supposing still that such a theory can be stated in a natural language? Perhaps we can raise our central worry about the concept of a person that a speaker must have to know an epistemically adequate theory incorporating the base clause ‘Genghis Khan’ refers to x iff $x =$ Genghis Khan and x is a person’. If the rest of the theory is largely homophonic, it will contain a clause along the lines of ‘ $(\forall x)$ (x satisfies “is a person” iff x is a person)’. Perhaps there are several concepts of a person we might hypothesize that a speaker must associate with the term ‘person’ to understand that term. The theory will not tell us which of those concepts is required to understand either the ‘person’ base clause or the ‘Genghis Khan’ base clause (given the Fregean Assumption, it is natural to suppose that the same concept of a person is required for understanding each clause). Adding structure to the right-hand sides of base clauses will produce a factually adequate T-theory only if all terms employed in right-hand sides throughout the theory are such that there is only one concept that, when associated by a speaker with the term, could (in the sense of epistemic possibility—see footnote 10) suffice for understanding it (we can even allow that ‘person’ is such a term). Consequently, my central argument has as its target not only (largely) homophonic T-theories, but any T-theory stated in a natural language that does not, so to speak, lay bare the structure of all concepts that a competent speaker associates with a language’s primitive terms. For some primitive terms, laying bare conceptual structure can be accomplished with a homophonic base clause; for many, such as names, it cannot.

This is a very strong requirement. Given the competition constraint, it is close to supposing that there is a set of primitive concepts *expressible in a natural language* from which all concepts required for understanding the primitive terms of any language can be constructed. Such a hypothesis is certainly not without precedent—atomic empiricists and others (including, perhaps, some lexical semanticists) have endorsed some such claim. A T-theory whose base clauses are specified in the indicated fashion is surely full-blooded, however—it goes as far as one can in explicating the concepts possession of which is required for understanding a language’s primitive terms. Moreover, it is not at all obvious that any natural language will contain a set of primitive terms that will suffice for framing such a factually adequate theory. That is, it is not at all obvious

that we will not encounter the same problem we encountered above in trying to frame a T-theory for a non-*L*-like language in *L*. Names in *L* are too *rich* in what is required to understand them to use them on the right-hand sides of base clauses in a T-theory for a language that does not require speakers to know that names of persons are names of persons. There is no guarantee that an *L*-like language will contain any other terms, primitive or complex, that can express the hypothesis that understanding ‘Genghis Khan’ requires a speaker to know only *which* object he is, and not what kind of object. Perhaps every natural language will contain some primitive term for which one cannot frame a base clause expressible in the language itself or any other natural language that reveals enough “conceptual structure” to figure in a factually adequate T-theory. There is no *a priori* reason to think that a natural language can express its own semantics in the form of a factually adequate theory of meaning.

NOTES

Thanks are due to the audience at the Inland Northwest Philosophy Conference 2000 on Truth and Meaning for their helpful comments; I am particularly grateful to Peter Ludlow, Mark McCullagh, and Kirk Ludwig.

1. I will often speak as though it is the bolder Davidsonian claim that is under scrutiny, but my arguments are easily reformulated to apply to the weaker Davidsonian claim.
I take it that, strictly speaking, even the meeker Davidsonians would accept that knowing a T-theory for English is not merely sufficient for understanding English, but that one who knew such a theory would understand English in virtue of that knowledge.
2. Davidsonians typically acknowledge that a speaker must also know that the T-sentences are consequences of a T-theory in light of the concerns raised by Foster (1976); in some sense, speakers must know not only a T-theory, but also that what they know is a T-theory. There are several ways to explicate the extra knowledge required for semantic competence. One might claim that it is sufficient for a speaker to know in addition that the theory conforms to certain “formal and empirical constraints” such as the canons of radical interpretation (Davidson 1976). Or, one might claim that we need not suppose that a speaker has any propositional knowledge beyond knowledge of the T-theory—he is able to interpret utterances of his language because his knowledge of a T-theory is localized within the mind’s “semantics module” (Larson and Segal 1995, 34–42). These suggestions are all concerned with establishing the epistemic adequacy of T-theories. My argument allows that T-theories are epistemically adequate, and so it is not important for my purposes to decide among these suggestions. I will speak for simplicity as though knowledge of a T-theory is supposed to suffice for semantic competence in a language; my argument is unaffected if a stronger claim prompted by Foster’s problem is substituted throughout. None of the responses to Foster’s problem that appear in the literature are of any obvious help in establishing that T-theories as standardly conceived are factually adequate (nor are they intended to provide such help).
3. For convenience, I use the term ‘T-theory’ ambiguously to denote either a theory expressed in written or spoken form by using a particular metalanguage—that is, the content of such an expression—or the expression itself. In the former sense of ‘T-theory’, the theory itself can be expressed in indefinitely many metalanguages and hence homophony is a property of a particular expression of a theory, not of the theory itself. (In that sense of ‘T-theory’, I will mean ‘T-theory that is stateable homophonically’ by ‘T-theory’ unless I indicate otherwise.) In the latter sense, we can simply talk of homophonic truth theories. Which sense I have in mind should be clear from context.
4. I follow the convention of using uppercase letters to denote concepts and thoughts, and will be sloppy about use and mention since which is required should be clear from context.

5. In footnote 10, I suggest that the assumption that *L* is a possible language is dispensable. My argument proceeds on the basis of how names are understood in a hypothetical language simply because of the prominence of names in various debates in the philosophy of language. Other categories of primitive term would serve my purpose equally well.
6. I very much doubt that *L* is English. ‘Genghis Khan was a mindless automaton controlled from Mars’ could be true, and that is very good evidence against the hypothesis that *L* is English; if understanding ‘Genghis Khan’ required believing that any bearer it has is a person, then ‘Genghis Khan is F’ should entail ‘Genghis Khan is a person’ for more or less any ‘F’.
7. I will drop the qualifier ‘in *L*’ for the remainder of this section; ‘Genghis Khan’ is a name of the name for Khan in *L* and not a name of the English name. Uses of the homonym are uses of the English name; this paper is not written in *L*.
8. Or perhaps: it refers to a person if it refers at all.
9. Besides, a theory of thought in conjunction with an epistemically adequate theory of meaning would not necessarily amount to a factually adequate theory of meaning. A complete theory of thought might explicate all possible concepts of Genghis Khan, but it would not explain which of those concepts one must associate with names of Khan to understand them in a given language—how one must think of Khan to understand a name for him. I will argue that epistemically adequate T-theories with the minimal structure property stated in English or another natural language will not provide that explanation either. If neither theory provides the needed explanation of name comprehension, nor does their conjunction.
10. It is here that we can dispense with the assumption that *L* is a possible language in any sense but the epistemic. Even if no language could contain names that are understood as the names of *L* are alleged to be, the contrary hypothesis makes sense—there could be genuine debate over its truth-value. And so T-theories for *L*-like languages should be formulable by the competition constraint. From this it follows that there is nothing in my argument with which a direct reference theorist should disagree. It takes more to understand names in *L* than a direct reference theorist is likely to accept is needed to understand names in actual languages. Perhaps the direct reference theory is true of names in all possible languages. It remains the case that the contrary hypothesis makes sense, and so we should be able to formulate T-theories that express the contrary hypothesis.

REFERENCES

- Davidson, D. 1976. “Reply to Foster.” In G. Evans and J. McDowell (eds.), *Truth and Meaning*. Oxford: Oxford University Press.
- . 1984. “Truth and Meaning.” *Inquiries into Truth and Interpretation*. Oxford: Oxford University Press.
- Dummett, M. 1975. “What Is a Theory of Meaning? (I).” In S. Guttenplan (ed.), *Mind and Language*. Oxford: Oxford University Press.
- Evans, G., and J. McDowell, eds. 1976. *Truth and Meaning*. Oxford: Oxford University Press.
- Foster, J. 1976. “Meaning and Truth Theory.” In G. Evans and J. McDowell (eds.), *Truth and Meaning*. Oxford: Oxford University Press.
- Larson, R., and G. Segal. 1995. *Knowledge of Meaning*. Cambridge, Mass.: MIT Press.

Truth and Meaning

Robert Cummins
University of California@Davis

INTRODUCTION

DONALD DAVIDSON'S "Meaning and Truth," revolutionized our conception of how truth and meaning are related (Davidson 1967). In that famous article, Davidson put forward the bold conjecture that meanings are satisfaction conditions, and that a Tarskian theory of truth for a language is a theory of meaning for that language. In "Meaning and Truth," Davidson proposed only that a Tarskian truth theory is a theory of meaning. But in "Theories of Meaning and Learnable Languages," he argued that the finite base of a Tarskian theory, together with the now familiar combinatorics, would explain how a language with unbounded expressive capacity could be learned with finite means (Davidson 1965). This certainly seems to imply that learning a language is, in part at least, learning a Tarskian truth theory for it, or, at least, learning what is specified by such a theory. Davidson was cagey about committing to the view that meanings actually *are* satisfaction conditions, but subsequent followers had no such scruples.

We can sum this up in a trio of claims:

Davidson's Conjecture

- (1) A theory of meaning for L is a truth-conditional semantics for L.
- (2) To know the meaning of an expression in L is to know a satisfaction condition for that expression.
- (3) Meanings are satisfaction conditions.

For the most part, it does not matter in what follows which of these claims is at stake. I simply take the three to be different ways of formulating what I call Davidson's Conjecture (or sometimes just The Conjecture).

Davidson's Conjecture was a very bold conjecture. I think we are now in a position to see that it is probably false, but I do not expect many to agree with me about this. Since the publication of "Meaning and Truth," truth-conditional semantics has been pretty much all the semantics there is. In the current climate, therefore, it is something of a challenge to get philosophers of language to realize that the Conjecture is not obviously *true*. Generations of philosophers have been trained to regard The Conjecture as a truism. What else could semantics be? Surely, to understand an expression, one must know the conditions under which it is satisfied!

Prior to Davidson, semantics, at least in Philosophy, was speech act theory: Austin, Grice, and their followers (Austin 1962; Grice 1957). That tradition either died, or was co-opted. Here is how the co-option went. The Gricean program, in the hands of Bennett (1973, 1976), Bach and Harnish (1979), Lewis (1969), Shiffer (1981, 1982), and their followers, reduces linguistic meaning to intentional psychology—i.e., to propositional attitudes. Fodor (1975), Schiffer (1981), and others then introduced what I call the *representational theory of intentionality* (RTI hereafter): the idea that an intentional attitude is a mental representation in a cognitive role—e.g., a belief is realized as a sentence in Mentalese available as a premise in inference but not as a goal specification. So, meaning for public language reduces to the attitudes, and the attitudes reduce to cognitive psychology and a theory of mental representation. A theory of mental representation, in this tradition, is, in Fodor's words, supposed to tell us where truth conditions come from (Fodor 1987, 1990). And that brings us back to Davidson's Conjecture. Meanings for Mentalese are to be given by a truth-conditional semantics, and the content of a propositional attitude is just the truth-condition of its associated mental representation. Meanings for a natural language, then, are specified finally in terms of the truth conditions of the Mentalese constituents of the attitudes involved in linguistic communication.¹ Thus Gricean speech act theory ultimately rests on truth-conditional semantics. The substantive content of Speech Act Theory was relegated to "pragmatics"—the business of distinguishing promises from threats, and specifying the contextual factors involved in determining truth-conditions.

Of course, you do not need to be a Gricean about meaning to get to this point. All you really need is the view that understanding² an expression E in a language L requires a mental state—either a representation or an intentional attitude—that has the same content as E. This reduces the theory of meaning and understanding the expressions of a language—the semantics anyway—to the theory of mental content. You then assume that a theory of mental content signs truth/satisfaction conditions to mental states, either directly, or via the RTI. And that brings you back to Davidson's Conjecture.

So the philosophy of language turned into truth-conditional semantics, and the philosophy of mind labored to explain how mental representations could come to have the satisfaction conditions required. Thus it is that “Meaning and Truth” set the agenda for the philosophy of language and the philosophy of mind, linking the two tightly together in the process.

The link is more important that it might first appear. Once you have a Davidsonian story about the semantics of natural language, it is nearly irresistible to conclude that intentional states or mental representations (or both) must have a truth-conditional semantics as well. How else could we hope to get a grip on how it is possible to mean and understand the expressions of a language? If the meanings of linguistic expressions are satisfaction conditions, and someone knows the meanings of those expressions, then surely they know satisfaction conditions for those expressions. The knowledge is tacit, of course, but can be tapped by suitable queries about what is “intuitively” true under some specified set of hypothetical or actual circumstances. This is how we get the conclusion that mental representation must be, “classical” (Fodor and Pylyshyn 1988; Fodor and McLaughlin 1990). It is worth setting this out explicitly.

- Davidson’s Conjecture: the meaning of a linguistic expression is a satisfaction condition for it.
- To understand a linguistic expression that means M, you must be able to token a mental representation that means M. (For example, to have the thought that p you must be able to token a mental representation that means that p.)
- Hence, mental representations must have a truth-conditional semantics, i.e., they must be “classical.”

This inference from the Conjecture to the nature of mental content carries a price.³ To see what it is, we need to begin with a basic constraint on any theory of linguistic meaning.

Communicative Constraint: The meaning of a natural language expression is whatever it is you have to know to understand that expression.

What I have just called the communicative constraint on linguistic meaning says, in effect, that linguistic meanings are whatever it is that have to be grasped or possessed for linguistic communication to be successful. Ultimately, a theory of meaning for natural language must dovetail with the psychology of linguistic communication.⁴

We can now see why the inference from Davidson’s Conjecture to the na-

ture of mental representation could be pricey. There are good reasons to think that the mental structures required for language understanding do not have a truth-conditional semantics. It is the burden of this chapter to argue this point. If you accept the point, and you accept the Communicative Constraint on linguistic meaning, you will think that a theory of language understanding will make no use of truth-conditional semantics. It doesn't follow from this that natural languages don't *have* a truth-conditional semantics. But it does follow that there is no good reason to think that a truth-conditional semantics for natural language will have any place in a mature psycholinguistics.

So here is the bottom line: I think that Davidson's Conjecture is a mistake. I think that truth has little to do with meaning. Or rather, so that we won't simply slide into arguing over the word, I think that truth has little to do with speaking and understanding a language.

COMMUNICATIVE VS. REFERENTIAL MEANING

Let's begin with some terminology. By the *communicative meaning* of a term in a language I mean whatever you have to have in your head to understand it.⁵ By the *truth-conditional meaning* of a term in a language I mean its satisfaction condition, or its role in generating one in the pragmatic context of some particular production of it. We can now express the central question thus:

- Are communicative meanings truth-conditional meanings?

OK. So what do you have to have in your head to understand, say, 'elevator'? Well, you have to have a more or less adequate concept of an elevator. But this just names the problem. What do you have to have in your head to have a concept of elevators? I think it is pretty clear that what you need is some basic knowledge of elevators. If you ask someone what 'elevator' means, they will tell you what an elevator *is*. They might, if they are very forthcoming and articulate, say something like this:

Imagine a little room like a closet that moves up and down in a vertical shaft in a building. You get in on one floor, and the thing moves up or down to other floors where you can get off. Faster and easier than stairs. I think it is done with pulleys. Modern ones are controlled with buttons inside, and you can summon it with a button by the door leading to it on any floor.

And they draw a diagram:

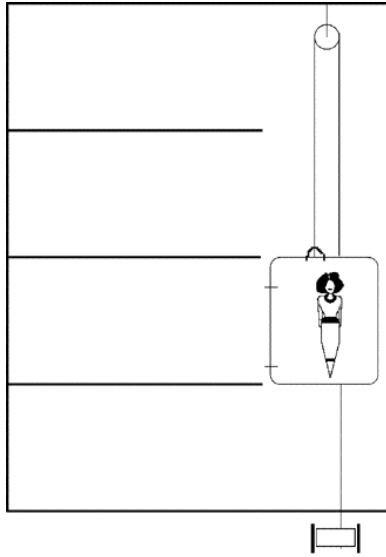


Figure 10.1. Drawing of an elevator.

This much, I think, would be plenty in ordinary life or a psychology experiment to demonstrate that the “subject” has the (or a) concept of an elevator. And it would be enough precisely because it would demonstrate basic knowledge of elevators. So it seems clear that one can be said to have concepts in virtue of having a basic knowledge of their instances. If you know what an elevator is, you have the concept of an elevator. Moreover, if you ask someone what ‘elevator’ means, the same answer will do the trick. If, in answer to the question, “What does the word ‘elevator’ mean?” they demonstrate possession of a ‘pretty good’ concept of an elevator, then they know what ‘elevator’ means.

All of this knowledge one has that comes into play in connection with elevators is not just about elevators, of course. It is also about buildings and pulleys, for example. But the *topic* of the knowledge that one accesses when, as we say, one applies the concept of an elevator, is: *elevators*. Similarly, one can have an entire book about elevators. That book will also be about lots of other things, but the topic is elevators. I have no general analysis of what it is that determines the topic of a book or a body of knowledge. I don’t think it is a very tight notion. Psychologically, the knowledge that gets accessed when a certain concept comes into play will vary from occasion to occasion and from person to person. My knowledge differs from yours, and my own is constantly changing. Moreover, which parts or aspects of my knowledge of a particular topic I happen to access on a given occasion will depend on the cues and on prior activation. But, given a reasonable amount of shared knowledge and stability over time, we can expect, in ordinary cases, a large overlap of core knowledge across persons and (reasonable) times.

On this view of things, the concept of a horse, and hence the communicative meaning of the word ‘horse’, is not a mental representation the reference of which is horses or the property of being a horse. It is, rather, a body of knowledge loosely identified by its topic. Just as a book about horses has horses as its topic, but not its referent, so a concept of horses has horses or the property of being a horse as its topic rather than its referent. With some trepidation, I’m going to sum this up by saying that a concept (of horses, say) is a *theory* (of horses),

the idea being that theories are organized bodies of knowledge that we identify in much the way we identify concepts—viz., by specifying a topic. One can have a theory of motion or a concept of motion; one can have a theory of pain or a concept of pain; one can have a theory of success or a concept of success. Theories, like concepts, are identified by their topics, not by their referents.⁶ And they are, at least on the hoof, blessed with fuzzy boundaries that overlap other theories identified by other topics. Indeed, the identification of theories by topic, while useful, is a kind of heuristic in just the way I think the standard identification of concepts is a heuristic: it points you in the right direction if you want to look it up, but not much more.⁷

Concepts, conceived as (perhaps tacit) theories, are pretty clearly what you need to have in your head to understand terms such as ‘elevator’ and ‘brown’ and ‘horse’ and ‘galloping’. They are also just what you need, along with the relevant sensory apparatus, to recognize elevators, horses, brown, and the gallop. And they are what you need to reason about such things. All of this is as it should be, since, when someone tells you not to ride the brown horse as he is likely to gallop, you don’t want to avoid riding elevators. An immediate consequence of this view of concepts, and hence of communicative meanings, however, is the following:

Concepts do not semantically combine in the way required by truth-conditional semantics.

The standard Tarskian combinatorics (Tarski 1956) suggests a mechanical process for combining a Mentalese term for being striped with a Mentalese term for being a fish, a process that yields a complex Mentalese term for being a striped fish. But no Tarskian process will semantically combine a theory of stripes with a theory of fish to yield a theory of striped fish. Even more obviously, the denial of a theory of fish is not a representation applying to all nonfish in the way that the denial of a Mentalese term for fish is (or would be if there were such a thing), precisely, a term applying to the nonfish. Tarskian combinatorics are hopeless in connection with the sorts of psychological structures concepts must be to do their jobs.

This is an instance of a widespread problem. The kinds of mental representations that are subject to Tarskian semantics are what Fodor and Pylyshyn (1988) call ‘classical’ representations: language-like concatenations of arbitrary primitive symbols whose syntactic rules of formation are directly exploitable by truth-conditional combinatorics. No one would dream of trying to exploit Tarskian truth-theory to cope with the semantic complexity and productivity of pictures, maps, graphs, or activation vectors. It only works for language-like schemes. Yet there is little reason to think that classical, language-like schemes have any real representational virtues. This is because there are basically just three ways that arbitrary mental symbols of the Language of Thought (LOT) va-

riety can enter into cognitive explanations: As *triggers* for procedures, as *cues* for stored knowledge, and as *constituents* of complex representations.

The point can be brought out by a simple example. You are asked to go milk the cow. You make a plan to carry out this request. Among your early sub-goals is the subgoal to find the cow. You decide to look in the barn. When you get to the barn, you walk around inside looking for the cow. You look in a stall, and token a |cow|—a mental symbol that refers to cows. But just how did this cow recognition work? To recognize cows, you need to know something about them. You need, at least, to know how they look. A mental symbol does not contain any information about how cows look, and so it is not what psychologists would call a concept. You need to deploy your knowledge of cows in order to recognize a cow. It is your knowledge of cows, including tacit knowledge about the sort of retinal projections they tend to produce, that makes it possible for you to token a |cow| when you encounter a cow. So the Mentalese |cow| did no work for the object recognition system, its just signaled its output.

But that is not all. Having tokened a |cow|, where do you stand in the great scheme of things? The |cow| tokening *triggers* the next step in the plan. Now that you have located the cow and are on the spot, you need to locate the udder. Here, something like a picture of a cow, an image, say, would be very helpful, whereas a mental word is totally useless unless it happens to function as a retrieval *cue* for some stored knowledge about cows. Faced with actually having to deal with a cow, the burden therefore shifts again from the symbol to your stored knowledge, because the symbol, being arbitrary, tells you nothing about cows. So it turns out that it is not because you have a Mentalese term for cows that you get the milking done, it is because you have a route—activated by a cue—to something else, some stored knowledge about cows. Mentalese |cow|s could play a role in stored knowledge about cows only as pointers to it, or as constituents of complex representations—|cows have udders between their back legs|, for example—that are, on the Mentalese story, implicated in the possession of stored knowledge about cows.

I do not think this should come as any real surprise to LOTers, for I think the view is widespread among them that it is really stored knowledge that does the explanatory work anyway. But it is worth emphasizing that there is a big difference between appealing to the fact that one has a primitive mental symbol referring to cows, and appealing to the fact that one has a lot of knowledge about cows. LOT commits one to the view that representations of cows don't tell you anything about cows.

Perhaps it is not so bad that LOT entails that the representations that are satisfied by cows have only an indirect role in the explanation of cow cognition, for there are always mental sentences to tell us about cows. But let us just be clear about what LOT is committed to here: The view we have arrived at is that cognition is essentially the application of a linguistically expressed theory. All the

serious work gets done by sets of sentences that are internal tacit theories (ITTs) about whatever objects of cognition there happen to be. As far as cognizing cows goes, your |cow|s really don't matter; it is your ITT of cows that does the work.

But, of course, ITTs are not subject to Tarskian combinatorics. Indeed, it is pretty obvious that no psychological structure can play the roles characteristic of both a Tarskian term and concept. Concepts, for example, subserve object recognition. A concept of a fish (a FISH) is what enables one to recognize fish. To recognize fish, you need to know something about fish—you need a theory of fish, in short. Having a Mentalese term is of no use at all; you have to learn to token that term in response to fish, and that is just what knowing something about fish allows you to do, and what you cannot hope to do if you don't know anything about fish. Similarly, to understand the word 'fish', you need to know something about fish. Having a mental term, by itself, would be no help at all, since having a mental term referring to something is not the same thing as knowing anything about it. You cannot understand 'fish' if you do not have a FISH, and your understanding of 'fish' is exactly as good as your FISH.

Mental terms in a language of thought, if there is such a thing, have satisfaction conditions: something counts as a |fish| just in case it is satisfied by fish. Consequently, mental terms in a LOT would be subject to semantic combination: you can combine a |striped| and a |fish| and get a |striped fish|. But having a |fish| at your disposal does not, by itself, endow you with any knowledge of fish, and hence does not enable you to recognize fish, or understand the word, or reason about fish. Expressions in a LOT might have the same truth-conditional meanings as the expressions of a natural language, but activating (tokening) a LOT expression that is truth-conditionally equivalent to an expression in a natural language could not possibly constitute *understanding* that natural language expression. To repeat, the story has to be that the Mentalese terms cue the corresponding theories.

MENTAL MERGING

I have been urging that communicative meanings are rather like theories. Since theories are not candidates for relevant sort of Tarskian combinatorics, it follows that a Tarskian truth theory cannot be a theory of communicative meaning. As I pointed out earlier, this does not refute Davidson's Conjecture, but it strips Davidson's Conjecture of most of its relevance to Cognitive Science. Even if a natural language could be fitted with a truth-conditional semantics, that would not help explain how it is learned or understood. Since natural language is a biological adaptation whose function is enabling communication—a fact philosophers of language sometimes forget and almost always neglect—the interest in such a semantics would be largely or completely orthogonal to the problem of understanding how we understand a language.

But if concepts do not have a Tarskian semantics, how do we combine our

understanding of ‘brown’ and ‘horse’ to get an understanding of ‘brown horse’? Theories do not simply merge, and the denial of a theory of horses is not a theory of nonhorses. Davidson’s Conjecture, and its implications for language understanding, gave us a story to tell about how our understanding of complex expressions could be constructed from our understanding of their constituents. What shall we put in its place?

This problem would need facing even if you believed in a language of thought with a truth-conditional semantics. For suppose you have uttered, ‘The man holding the brown shoe is my brother,’ and my language understanding system has constructed a truth-condition for it. What it has, in effect, is a Mentalese translation of your sentence, containing Mentalese terms like |man|, |brown|, |shoe|, and |holding|. We can assume, for the sake of argument, that each of these activates the corresponding concepts, |man|s cuing MANs, |brown|s cuing BROWNS, and so on. But this is a far cry from having a conception of the state of affairs expressed by your sentence. How does one build up that conception from MANs, BROWNS, SHOES, and so on, together with the truth-conditional combinatorics? Building a |brown shoe| from a |brown| and a |shoe| does not automatically give you a BROWN SHOE.

It is glaringly obvious, once the question is raised, that symbolically represented theories are not subject to Tarskian combinatorics. Truth-conditional combinatorics, therefore, allows you to explain how the truth-conditional meaning for a complex expression can be built up from the truth-conditional meanings of its components and its syntax, but it leaves untouched how the communicative meanings of complex expressions could be built up from the communicative meanings of their components. A truth-condition for a complex expression provides no clue as to how one might build up the conception of the situation that expression so readily conveys to the mind of a mature speaker. We are thus led to ask whether there is some other way of representing the relevant knowledge—some nonlinguistic way of representing the knowledge involved in BROWNS and SHOES, for example—which does allow the kind of relatively straightforward concept-merging that real-time language understanding so obviously requires.

In connectionist networks, long-term knowledge is stored in the connection weights. Whatever such a system knows about shoes and brown resides somehow in the pattern of connectivity and the associated weights.⁸ It is, in the present state of play, a mystery how we should “read” a pattern of connection weights. No one knows how to take a verbally expressed body of knowledge and express it as a pattern of connection weights. Indeed, if John Haugeland (1990) is right, and I think he is, this is impossible. According to Haugeland, different genera of representational schemes allow for the expression of characteristically different contents. Pictures and sentences are intertranslatable only in the very roughest way. We should expect the same for sentences and patterns of connec-

tion weights. However, this message of incomensurability between verbal and connectionist representation is a *hopeful* message in the present context, because we know that the problem facing us has no ready solution—perhaps no solution at all—in its verbal form: logically combining verbally expressed theories, to repeat, has no hope of giving us what we want. This, perhaps, is enough to justify a bit of wild speculation in spite of our ignorance of the semantics of weight matrices.

Think, then, of a weight matrix as an encoding (doubtless idiosyncratic) of a kind of know-how. It might be knowledge of how to retrieve an item from memory given a cue of some sort. This is what we have in the famous Jets and Sharks network of McClelland and Rumelhart (1988). Or it might be knowledge of how to pronounce English text, as in Sejnowski and Rosenberg's NetTalk. Know-how, it seems, is naturally captured in a weight matrix. Can we think of concepts as know-how? Certainly. To possess the concept of a shoe is, to a first approximation, to know how to recognize one, to know how they are worn, and, if one is a linguistic creature, to know how to describe one. Knowing how to describe a shoe is, of course, know-how like any other. In particular, we should not assume that knowing how to describe a shoe requires a sort of "declarative memory," where this is conceived as a stored Mentalese description. The stored-description account has many failings, not the least of which is that we do not always describe the same thing in the same way. We get a more realistic account if we imagine a network that generates descriptions as outputs, with the description generated depending on the details of the input and the current state of activation—*set*, as it used to be called in psychology. In a similar vein, having a conception of the color brown is being able to recognize it, being able to give instances of brown things, being able to relate brown to other colors (e.g., darker than yellow and lighter than black), and so on.

Can we assemble the connectionist know-how that goes with SHOE and the connectionist know-how that goes with BROWN into the know-how that goes with BROWN SHOE? Notice that this is not a question in semantics at all, but a question about the mechanics of network building. We need a design that exploits the presence of a BROWN network and a SHOE network and generates, on the fly, and temporarily, a structure that exhibits the kind of know-how characteristic of BROWN SHOE possession.

It must be confessed that we are nowhere near to understanding how this might be done. But we do, I think, have a pretty good beginning on how the problem should be posed.

We start with a brief consideration of representation in connectionist networks, beginning with simple three-layer feed forward cases. Following Paul Churchland (1998), consider a network that learns to discriminate hillbilly families in terms of facial resemblance. Figure 10.2 depicts a simplified version of such a network, with the activation space at the hidden layer contracted to al-

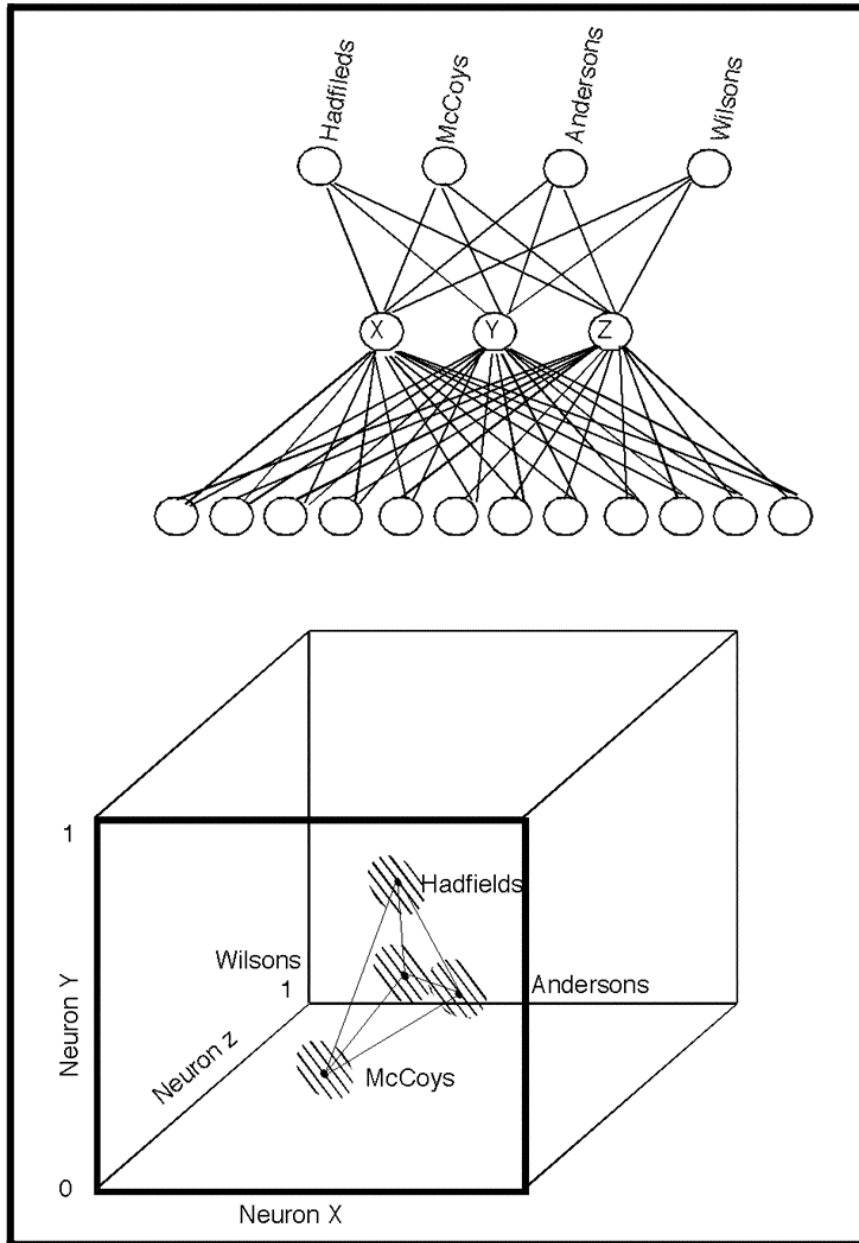


Figure 10.2. The network has acquired four prototype regions corresponding to facial family resemblance.

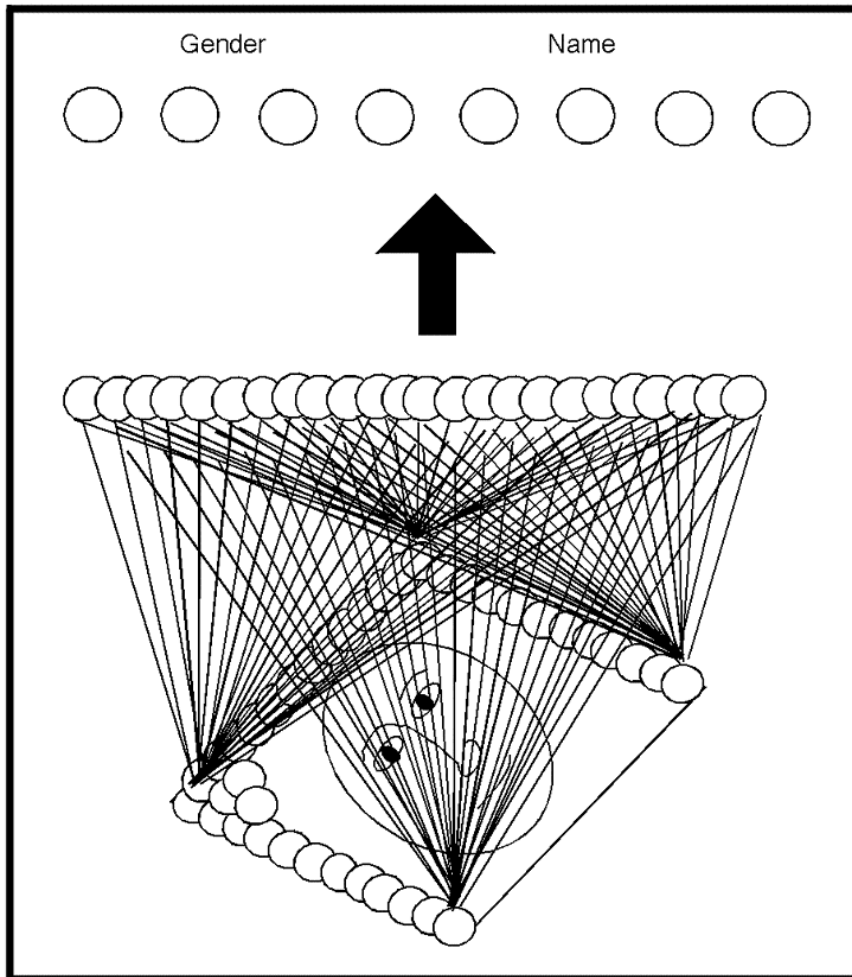


Figure 10.3. Schematic of Cottrell's face recognition network.

low three-dimensional illustration. The points in the space are what Churchland calls prototype points. They are centers of gravity around which cluster the points corresponding to related family members. They are a revealing way to represent the way that training the network partitions up the relevant activation space. The geometry thus revealed will be remarkably constant across different networks trained to the same task, including ones with differing input codings and even ones with differently dimensioned hidden layers (Laasko and Cottrell 2000). We are thus led to the idea that there is an objective structure to the relevant face space, and that trained networks discover this and represent it via an

isomorphic activation space. In such a space, it seems reasonable to think of the prototype points as something like individual concepts in a conceptual space. This perspective becomes more compelling as the networks face more complex tasks. Cottrell's tiny face recognition network (retina of 64 x 64 pixels, figure 10.3) implicitly partitions its activation space in such a way that female faces tend to be closer to each other than to male faces and vice versa (Cottrell 1991a).

Simple recurrent networks of the sort pictured in figure 10.4 pose a different case because they allow for dynamic representation. They are probably best conceived in terms of paths in activation space rather than points. This approach seems to work nicely for Elman's well-known grammar network, for example (Elman 1992).

Connectionist theory thus provides a compelling example of the kind of representation by structural similarity that I recommended in *Representations, Targets and Attitudes* (Cummins 1996). It provides representations that are structurally rich, representations that themselves guide cognition rather than func-

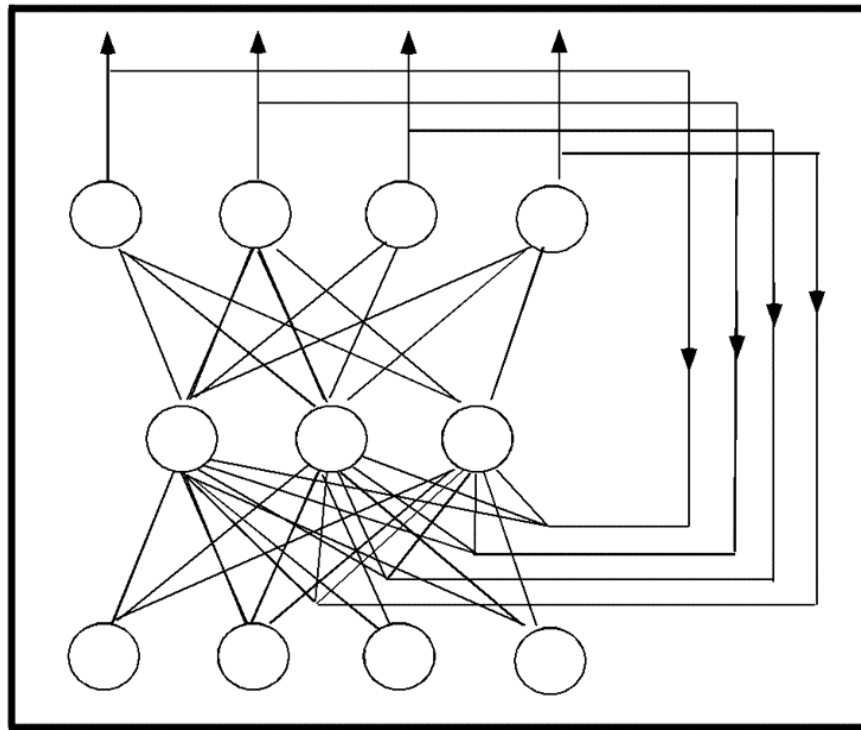


Figure 10.4. Simple recurrent network. Projections from the output layer to the hidden layer give the network a kind of short term memory of its immediate history.

tion as mere indicators in detection. Unlike the representations posited by LOT theories, these representations are plausible candidates for concepts.

They are not, as yet, however, plausible candidates for the sort of fleeting merges that seem to underlie language understanding. No cross-network associations between, e.g., a color network and a shape network, will fill the bill here because, first, associations have to be learned, and, second, because they have to be unlearned to go away. A reference to yellow dogs early in your discourse makes it easier to understand talk of brown dogs later, not more difficult. There are powerful connectionist techniques for representing hierarchical bindings of the sort found in parse trees (Smolensky et al. 1992). It is tempting to suppose that vectors representing a parse could somehow be used to orchestrate the kind of conceptual liaisons we are after, but I think it is fair to say that no one currently knows how to do this.

THE COMMUNICATIVE FUNCTION OF LANGUAGE

A novel conception of the function of language emerges from the foregoing discussion. Davidson's Conjecture implies that language is a medium for the expression of propositions and their constituents. It serves its communicative function when the hearer figures out what proposition the speaker expressed (or perhaps which proposition the speaker intended to express). The approach I have been urging implies that language is primarily in the communication business, and only secondarily, if at all, in the expression business. Sentences, on this view, are like recipes for assembling chunks of know-how into a know-howish conception of the speaker's communicative intention, and of the situation as the speaker conceives it. Sentences, in effect, tell you how to cook up a thought, where the thoughts thus cooked up are as different from words as are the cakes and pies from the recipes that tell you how to cook *them* up.

Viewed this way, it is possible—indeed, likely—that language can be used to communicate things it cannot begin to express, something poets and good novelists have always known. You can begin to get a sense of this by looking at the provision that language makes for “plug-ins.” A plug-in, as every web browser knows, is an independent routine that your browser can “call” when needed, e.g., to decompress a downloaded file. Language uses demonstratives to construct docking places for these devices, as illustrated in figure 10.5.

In your head, though, it is *all* plug-ins, a fact that has, I think, been obscured by the exaptation of language, especially written language, for expressive purposes quite foreign to its original biological function of facilitating communication in the service of social coordination. The expressive power of language is impressive, but hardly universal. It is, I think, much better at communicating thoughts than it is at expressing them. Failure to notice the distinction has led to the view that the only thoughts that can be communicated are the ones that can be expressed. When we put this together with Davidson's Conjecture, we

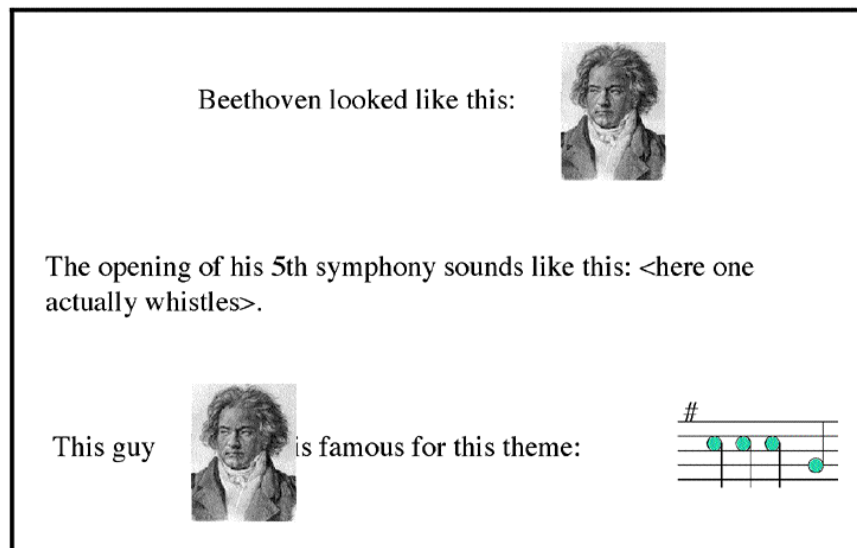


Figure 10.5. Linguistic expressions with plug-ins.

get the result that the only thoughts that can be communicated are those that have truth-conditional contents—propositions, in short. It is a short step from this position to the widespread view that the only thoughts we can have are the propositional attitudes, and hence that there is no thought or mental representation whose content language cannot adequately express. In our hearts we all know this is wrong, but recent philosophical tradition has taught us to live with it or suffer professional extinction.

It is nearly universally assumed that the communicative meanings of linguistic utterances are the same as their representational meanings. The idea goes like this: I have the thought that *p* that I wish to communicate to you. I construct a sentence that means (representationally) that *p*, and you decode it—i.e., you figure out what its representational meaning is, and conclude that that is what I meant to tell you. This story could be right. But it is important that we not just assume it. To see that it isn't inevitable, imagine a communicative system that works like this: There are instructions that tell you how to assemble nonlinguistic representations—pictures, say—from elements—pixels—you have available. In this system, the instructions and the messages communicated need have no meanings in common. Language *might* work like that. Sentences might be recipes for assembling thoughts, or even images, in the heads of others. If so, then the truth-conditions of my sentences, if they have any, will tell us nothing about what I communicate. This is because I can communicate an accurate picture to you without saying anything true about the scene pictured. The truth-

conditions of my sentences yields the limelight to the accuracy of the thoughts or other cognitive states they help to assemble.

To get a clearer view of the sort of possibility I have in mind here, consider the following communication system. You have a set of numbered storage bins. In these are little model houses, trees, ponds, lawns, roadways, signs, street lights, etc. You also have a table with a grid marked on it, with rows numbered and columns lettered. You get instructions like this:

- Put an item from bin 23 on 7A
- Center an item from bin 14 on 8C
- Put an item from bin 12 on 8D–8H

The result will be a model village. You assemble this representation on the basis of instructions that are built from a vocabulary that is utterly incapable of expressing any of the things represented by the model. The signal system and the representations it helps to assemble are representationally disjoint.

This sort of example demonstrates the possibility of a communication system in which the meanings the communicative symbols communicate are not the meanings they have. Could this be true of natural language? We are, I think, already in a position to see that it very likely is true of natural language. The words ‘house’, ‘tree’, ‘pond’, and so on, do not express the knowledge that constitutes your understanding of houses, trees, and ponds. They are signals that activate that knowledge, bring it on line, and, somehow, orchestrate its assembly into a more or less unified conception.

BEYOND THE PROPOSITIONAL ATTITUDES

I used to think (Cummins 1996) that nonlinguistic schemes could express propositions. For example, I thought we could take pictures to express propositions by following Stalnaker (1984) in thinking of a proposition as a set of possible worlds. Since a picture will “hold” in some possible worlds and not others, it partitions the set of possible worlds, and hence expresses a proposition. I now think, however, that Haugeland (1990) was right: sentences and propositions were made for each other, and so we must look elsewhere for the contents of nonlinguistic representations.

The striking thing about maps, diagrams, partitioned activations spaces, pictures, graphs, and other nonlinguistic representations is that they are not true or false, but more or less accurate. A sentence either hits its propositional target, or it fails. Nonpropositional representations, however, are better evaluated in terms of a graded notion of accuracy. Moreover, such representations are typically multidimensional. Pictures, for example, represent (relative) size, shape, color, and (relative) location simultaneously. The possibility thus arises that two pictures might be incomparable in overall accuracy, since one might do better

on some dimensions—size and shape, say—while the other does better on others—color and location.⁹ The concepts of truth and falsehood, and the Tarskian combinatorial semantics we have come to associate with them, will be no help at all in understanding how these nonpropositional representations fit or fail to fit their targets. Representational meaning for nonpropositional representations will have to be understood in different terms, as will their semantic structures.

A consequence of the graded and multidimensional nature of many non-linguistic representations is that they do not partition up the set of possible worlds in any neat way. What we get instead is a kind of shading along a number of interdependent dimensions. Since I cannot think of a more catholic notion of propositions than the one Stalnaker endorses, I have to conclude that most, perhaps all, nonlinguistic representations do not express propositions and are not true or false.¹⁰ But they evidently do represent. They represent how their targets are, with greater or less accuracy, along various dimensions. If we really want to understand meaning, we need to understand not only the representation of propositions, but the graded and multidimensional representation of nonpropositional contents as well. And if we want to understand the kind of meaning that is involved in mental representation, and hence in language understanding, we had best understand the kind of representation effected by the sort of dynamic partitioning of neuronal activation spaces that our synapses learn to effect. It would amaze me if truth-conditional semantics had anything significant to offer to this crucial research problem.

NOTES

1. There is a missing step here: Gricean stories provide only propositional contents, hence provide meanings for nothing smaller than a sentence. The Tarskian combinatorics, however, require satisfaction conditions for terms. See Cummins 1996 for a proposal about how to get the Tarskian combinatorics into a Gricean picture.
2. I am going to use ‘understanding’ as short hand for ‘meaning and understanding’ or ‘using and understanding.’ The idea is to have a single word for whatever you need to be either party—speaker or hearer—in successful linguistic communication. Passive mastery and active mastery of language differ, with the former outrunning the latter, especially in young children, and this suggests that there is more to speaking the language than there is to understanding it. Still, you have to understand it to speak it, and it is at least plausible that whatever you have to add to understanding (passive mastery) to get active mastery, it isn’t more *semantics*.
3. It ought to be darkly suspicious, too, since it is a license to do experimental cognitive psychology from the armchair. We begin by asking after the truth-conditions of propositional attitude sentences, and wind up with conclusions about the structure and contents of psychological states. For more on this theme, see Cummins 1991.
4. This need not be the case for artificial languages, I suppose, since these need not be primarily in the communication business. They may be primarily in the business of expressing truths, and rely for whatever communicative efficacy they have on their connections with natural languages.
5. For the picky: Of course, you need to be awake, and to be smarter than a post. What we want is what you have to add to the mind to enable understanding of some particular expression not previously understood.
6. I’m not sure what the referent of a theory would be. If you thought a theory was a set of

- sentences, which I do not, then, perhaps, the referent of a theory would be a proposition, viz., the proposition expressed by a conjunction of sentences used to express the theory.
7. Psychologists, of course, have suggested a number of theories about the form our concepts take. (The classic review is Smith and Medin 1981. For a recent review of the literature, see Gelman 1996.) They all have in common, however, the idea that a concept of *X* is stored knowledge about *X* that mediates recognition of and reasoning about *X*s. The dispute is over how that knowledge is stored and deployed, e.g., as a prototype or exemplar that is compared to instances in recognition and used to generate premises in inference, or as a frame, script, or semantic net. What you do *not* find in the psychological literature is the idea that concepts are terms in Mentalese that are satisfied by the instances of the concept in question. You do not find this because it wouldn't work, as we will see.
 8. Mathematically, we could reduce this to weights alone, dealing with connectivity by setting the weights between disconnected nodes to zero. But it is more intuitive to think in terms of what is connected to what, and how those connections are weighted. This allows us to think of a number of more or less independent nets that are only sparsely connected to each other.
 9. It seems likely that high accuracy on one dimension will often have to be paid for in lower accuracy in others, given limited resources. The eye, for example, gains considerable resolution and color information via foveation, but loses light sensitivity in the process. A map that shows all the streets of London on one page will be either too big to use in the car, or viewable only with magnification.
 10. Vagueness in language introduces problems that appear similar on the surface. Whether they are genuinely related to the kind of multidimensionality and gradedness we find in pictures, models, and graphs is not at all clear.

REFERENCES

- Alston, W. 1964. *Philosophy of Language*. Englewood Cliffs, N.J.: Prentice Hall.
- Austin, J.L. 1962. *How to Do Things with Words*. Oxford: Clarendon Press.
- Bach, K. and R.M. Harnish. 1979. *Linguistic Communication and Speech Acts*. Cambridge, Mass.: MIT Press.
- Bennett, J. 1976. *Linguistic Behavior*. Cambridge: Cambridge University Press.
- . 1973. "The Meaning-Nominalist Strategy." *Foundations of Language* 10: 141–68.
- Churchland, P. M. 1998. "Conceptual Similarity Across Sensory and Neural Diversity: The Fodor/LePore Challenge Answered." *Journal of Philosophy* 95: 5–32.
- Cottrell, G. 1991a. "Extracting Features from Faces Using Compression Networks: Face, Identity, Emotions and Gender Recognition Using Holons." In D. Touretsky, J. Elman, T. Sejnowski, and G. Hinton (eds.), *Connectionist Models: Proceedings of the 1990 Summer School*. San Mateo, Calif.: Morgan Kaufmann.
- Cottrell, G. and J. Metcalfe. 1991b. "EMPATH: Face, Emotion and Gender Recognition Using Holons." In R. Lippman, J. Moody, and D. Touretsky (eds.), *Advances in Neural Information Processing Systems*, vol. 3. San Mateo, Calif.: Morgan Kaufmann.
- Cummins, R. 1996. *Representations, Targets and Attitudes*. Cambridge, Mass.: MIT Press.
- . 1991. "Methodological Reflections on Belief." In R. Bogdan (ed.), *Mind and Common Sense*. Cambridge: Cambridge University Press.
- . 1979. "Intention, Meaning and Truth Conditions." *Philosophical Studies* 35: 345–60.
- Davidson, D. 1967. "Truth and Meaning." *Synthese* 17: 304–23.
- . 1965. "Theories of Meaning and Learnable Languages." *Proceedings of the 1964 International Congress of Logic, Mathematics and Philosophy of Science*, 383–94. Amsterdam: North Holland.
- Elman, J. 1992. "Grammatical Structure and Distributed Representations." In S. Davies (ed.), *Connectionism: Theory and Practice (Vancouver Studies in Cognitive Science, vol. 3)*. New York: Oxford University Press.
- Fodor, J. 1990. "Psychosemantics, or Where do Truth Conditions Come From." In W. Lycan (ed.), *Mind and Cognition*. Oxford: Basil Blackwell.

- . 1987. *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge, Mass.: MIT Press.
- . 1975. *The Language of Thought*. New York: Crowell.
- Fodor, J. and B. McLaughlin. 1990. "Connectionism and the Problem of Systematicity: Why Smolensky's Solution Doesn't Work." *Cognition* 35: 183–205.
- Fodor, J.A., and Z.W. Pylyshyn. 1988. "Connectionism and Cognitive Architecture." *Cognition* 28: 3–71.
- Gelman, S. 1996. "Concepts and Theories." In R. Gelman and T.K. Au (eds.), *Perceptual and Cognitive Development. Handbook of Perception and Cognition*, 2d ed. San Diego: Academic Press.
- Grice, P. 1957. "Meaning." *Philosophical Review* 66: 377–88.
- Haugeland, J. 1990. "Representational Genera." In W. Ramsey, S. Stich, and D. Rumelhart (eds.), *Philosophy and Connectionist Theory*. Englewood Cliffs, N.J.: Lawrence Erlbaum Associates.
- Laasko, A., and G. Cottrel. 2000. "Qualia and Cluster Analysis: Assessing Representational Similarity Between Neural Systems." *Philosophical Psychology* 13: 46–76.
- Lewis, D. 1969. *Convention*. Cambridge, Mass.: Harvard University Press.
- McClelland, J., and D. Rumelhart. 1988. *Explorations in Parallel Distributed Processing: A Handbook of Models, Programs, and Exercises*. Cambridge, Mass.: MIT Press.
- Schiffer, S. 1982. "Intention Based Semantics." *Notre Dame Journal of Formal Logic* 23: 119–59.
- . 1981. "Truth and the Theory of Content." In H. Parret and J. Bouraressse (eds.), *Meaning and Understanding*. Berlin: Walter de Gruyter.
- Sejnowski, T., and C. Rosenberg. 1987. "Parallel Networks that Learn to Pronounce English Text." *Complex Systems* 1: 145–68.
- Smith, E., and D. Medin. 1981. *Categories and Concepts*. Cambridge, Mass.: Harvard University Press.
- Smolensky, P., G. LeGendre, and Y. Miyata. 1992. "Principles for an Integrated Connectionist/Symbolic Theory of Higher Cognition." Technology Report 92–08. Institute of Cognitive Science, University of Colorado.
- Stalnaker, Robert. 1984. *Inquiry*. Cambridge, Mass.: MIT Press.
- Tarski, A. 1956. "The Concept of Truth in Formalized Languages." In *Logic, Semantics and Metamathematics*. Oxford: Oxford University Press.

BLANK

194

PART III
Aspects of Linguistic Meaning

BLANK
196

In insensitive Quantifiers

Herman Cappelen, *Vassar College*
and
Ernie Lepore, *Rutgers University*

CONTEXT-INSENSITIVE SEMANTICS FOR QUANTIFIERS (CIS)

ON THEIR FACE, the three explicitly quantified noun phrases in (1)–(3) do not seem to exhibit context sensitivity.

- (1) Most Norwegians like fish.
- (2) All ducks have soft beaks.
- (3) Exactly one person at Rutgers is a linguist.

It seems unobjectionable that a semantic theory should have the result that (1)–(3) express the general propositions represented by (1s)–(3s).

- (1s) [Most x : Norwegian x](x likes fish)
- (2s) [All x : Duck x](x has a soft beak)
- (3s) [Exactly one x : person x & At(x , Rutgers)](x is a linguist)

Compare (1)–(3) with (4)–(5):

- (4) That's nice.
- (5) I've eaten.

Unlike the former, the latter include the context-sensitive nouns 'that' and 'I' respectively. It is a fact about the meaning of such nouns that their referents are fixed by certain features of the context of utterance, and competent speakers know how their referents are determined in context. No component of (1)–(3) seems to be an indexical or demonstrative expression, and we don't have the intuition that there is some hidden indicator whose semantic-value needs to be fixed in the con-

text of utterance. Call the view that the semantics for quantifiers should be insensitive to context ‘CIS’ (short for Context-Insensitive Semantics for quantifiers).

INCOMPLETE QUANTIFIERS AND CONTEXT-SENSITIVE SEMANTICS FOR QUANTIFIERS

Most philosophers think CIS is mistaken. According to an alternative view, all quantifiers are context sensitive, much like the nouns in (4)–(5). Call this view CSS (short for Context-Sensitive Semantics for Quantifiers). According to CSS, the propositional component contributed by a quantified noun phrase change from context to context. Examples such as the following motivate this view.¹

- (6) Nina took the car.
- (7) Every computer was stolen.

Imagine an utterance of (6) in a context where it is clear that what the speaker means to convey by ‘the car’ is something like ‘the car we own’, and an utterance of (7) where it is clear that what the speaker means to convey by ‘every computer’ is ‘every computer in room 401 of Rockefeller Hall’. According to CIS, the semantic content of these can be represented in a first order language with restricted quantifiers as (6s) and (7s).

- (6s) [The x: car x] (took(Nina, x))
- (7s) [Every x: computer x] (x is stolen)

Here’s the problem: (6s) is true only if (and if) there’s just one car in the world; (7s) is true just in case absolutely every computer in the world was stolen. But these conditions are unlikely ever to obtain. Many authors take this to be a *reductio* of the view that (6s) and (7s) provide correct truth conditions for (6) and (7). It’s taken to be a *reductio* because it appears to be incompatible with some widely held intuitions about typical utterances of (6) and (7). Take a particular utterance of (6). A speaker will typically not intend to *assert, say, or claim* that there is just one car in the world. Nor will a typical audience so interpret any such utterance. A typical speaker of (7) will not intend to say, assert, or claim that all computers in the entire world have been stolen, and she is typically not at risk of being so interpreted.

Since one of the aims of this paper is to show that none of these intuitions provide genuine evidence against CIS, it is important to be accurate in the description of the relevant intuitions. We’ll quote directly how opponents of CIS describe their intuitions.

Stanley and Williamson (1995, 291) have an intuition to the effect that an utterance of a sentence containing an incomplete quantifier can be true.

Since there are clearly true utterances of

- (I) Every bottle has been put on the shelf

in ordinary contexts, it follows that utterances of sentences containing quantified expressions are evaluated with respect to contextually restricted domains.²

Discussing an utterance of the sentence ‘They took everything’, Francois Recanati writes that he has:

a *feeling* that ‘everything’ ranges over the domain of valuable objects in the house—not everything in the world (Recanati 1996, 445, emphasis ours)

Scott Soames (1986, 276) says we can use a sentence of the form ‘The F is G’ ‘to make true statements’ even though more than one F inhabits the world. Schiffer talks about what is stated, saying:

. . . it is clear that in uttering ‘The dog is chewing your hat’ the literal speaker is not *stating* something that entails that there is exactly one dog in the universe. (Schiffer 1995, emphasis ours)

Neale talks about what a speaker is committing himself to and what is asserted:

If I say to you right now

- (I) The table is covered with books

I would not normally be understood as *committing* myself to the existence of one and only one table. (Neale 1990, 93–94, emphasis ours)

Suppose I had a dinner party last night. In response to a question as to how it went, I say to you:

- (2) Everyone was sick

Clearly I do not mean to be *asserting* that everyone in existence was sick, just that everyone at the dinner party I had last night was (Neale 1990, 95, emphasis ours)

At the beginning of Marga Reimer’s (1998, 96) paper on the topic, she writes that in uttering the sentence, ‘The carpenter is not getting along with the electrician’ “the proposition thereby expressed might well be true” even though there is more than one carpenter in the world (and more than one electrician, for that matter).

These quotations are a representative selection of intuitions. Each falls into one of three categories:

- (INT_I) The intuition that for many utterances of sentences containing quantifiers it seems correct to say that what was asserted/committed

to/stated /said is a contextually enriched proposition (Soames, Schiffer, Neale, Reimer).

- (INT₂) The intuition that many utterances of sentences containing quantifiers are true even though the proposition semantically expressed (according to CIS) is false (Stanley and Williamson).
- (INT₃) The intuition that the propositions expressed by many utterances of sentences containing quantifiers are true even though the proposition semantically expressed (according to CIS) is false (Reimer).

Just so that we have all the relevant intuitions labeled and on the table, we'll call the intuition that we share about the relevant quantifier sentences INT₄:

- (INT₄) The intuition that many utterances of sentences containing quantifiers don't require contextual completion or enrichment. (Their domains are not restricted in a context of utterance.)

In response to INT₁–INT₃, a number of philosophers and linguists have suggested that the semantics for quantifiers should be rendered sensitive to context. There are a variety of ways of implementing this strategy formally, but we'll focus in what follows on only one such strategy. Just the same, we claim, our objections extend to every such suggestion of which we are aware.³

Standard semantic accounts of quantification assume that a domain for a quantifier expression is determined by its nominal alone, say, students for the noun phrase 'every student'. INT₁–INT₃ applied to utterances of, say, (7), is alleged to show that the quantifier 'every' doesn't quantify over every computer, only discourse-relevant ones. Context is supposed to determine just how broadly or narrowly a domain of quantification is to be restricted. Accordingly, a narrower quantifier domain might get selected, say, a contextually salient set that intersects with the extension of 'computer' to provide the domain of discourse for 'every'. It might be restricted to every computer in a certain room, say, in room 401, or every computer owned by a certain person, say, Bill Gates. Truth conditions for (7) might be explicitly represented by (7*).

- (7*) Every computer (i) was stolen.

where 'i' indexes a set that, in context, restricts the domain of 'every computer'.

Truth conditions for (7) are represented only slightly more rigorously by (7**).

- (7**) [Every x: x is a computer & x is in i](x was stolen)

In short, in some contexts, a contextually determined domain of quantification might be the set of computers in room 401; in another, it might be com-

puters in the entire building. In uttering (7), a speaker succeeds in expressing something true, because her utterance succeeds in indexing a set of objects that serves to restrict the complex quantifier expression ‘every computer’ to a narrower contextually defined domain. How a speaker conceptualizes this domain may or may not be relevant in determining which set gets selected in much the same way as it is, on standard semantic accounts of indexicals, irrelevant how a speaker conceptualizes himself with a use of ‘I’.

A bonus of appealing to context sensitivity in accounting for domain selection is its explanation for how distinct tokens of the same quantifier expression-type in a single sentence might range over distinct domains (Stanley and Williamson 1995). Suppose there are two groups of sailors, one on deck and one on shore, and all the sailors on deck waved to all the sailors on shore. In such circumstances, one might conclude that (8) can be used to express something true.

(8) Every sailor waved to every sailor.

But relativizing quantification over a fixed set of sailors, an utterance of (8) asserts that every sailor in this set waved to every sailor in this same set, including each to him or herself. Instead, an utterance of the two quantifiers in (8) might reasonably be taken as restricted to distinct domains, expressing in its context of utterance what (9) would in that same context of utterance.

(9) Every sailor (here) waved to every sailor (there).

Accordingly, truth conditions for (9) might be better represented as (9*),

(9*) Every sailor (i) waved to every sailor (k),

where ‘i’ and ‘k’ can index disjoint sets, so that ‘i’ together with the first token of ‘sailor’ can pick out the group of sailors on deck, and ‘k’ together with the second token of ‘sailor’ can pick out the group on shore, thus, rendering (9) to express a truth.

The rest of this chapter has three parts. First, we raise an objection to CSS. We argue that it makes predictions about utterances of, for instance, (7), not supported by the evidence. Second, we show that the evidence for CSS, i.e., INT₁–INT₃, fails to support it. It is simply not the case that these intuitions provide support for CSS. Finally, we use some of the data appealed to in the discussion of INT₁–INT₃ to raise a second objection to CSS.

FIRST OBJECTION TO CSS: NONEXISTENT INTERPRETATIONS

According to CSS an utterance of (7) can be used to express a truth, because its context of utterance can effect a completion of its incomplete quantifier: for a

use of (7) context restricts the range of its quantifier ‘every’. Suppose that in using (7) a speaker restricts the domain to computers in a salient location, say room 401, at time *t*, and thereby expresses what she would have expressed had she instead uttered (7’).

(7’) Every computer in room 401 at time *t* was stolen.

What, then, shall we conclude about (10)?

(10) Every computer is in room 401 at time *t*.

(10), with its incomplete quantifier, is as likely to be used to express a truth as (7). Since (7) can express what (7’) *can* express in the same context, why can’t we infer that (10) can express in a context what (10’) would?

(10’) Every computer in room 401 at time *t* is in room 401 at time *t*.

Clearly, (10) cannot express the necessary truth expressed by (10’). The challenge to CCS strategies is for them to establish that once appropriately contextually relativized, (10) and (10’) must express distinct propositions. We fail to see how that can be done.⁴

A proponent of CSS might protest that these alleged troublesome interpretations are never available. Perhaps a policy of screening off is in place, thereby excluding any contextually determined supplementation that is explicitly referenced in a predicate (or elsewhere). So, e.g., take a domain *D*. If *D* is explicitly referenced by an utterance of (10), it is screened off as a candidate completer. This screening-off constraint is supposed to preempt contextual completions that would render some seemingly contingent sentences as expressing necessarily true propositions. Any item referenced, or expressed, explicitly is thereby rendered contextually irrelevant, at least for the purposes of supplementing a token of an incomplete quantifier. If this ploy can be made to work, then what is expressed with (10) cannot be what would be expressed with (10’).

Whatever can be said in favor of screening off surely reflects nothing more than handy wisdom about the pragmatics of sound interpretation, and nothing about semantics. For three reasons we fail to see how semantics can prohibit an explicitly referenced domain from also being most contextually salient. First, with uses of potentially complete quantifiers, as in (1)–(3), an explicitly referenced domain is typically also the only contextually salient domain. As far as we can see, the only way for a proponent of CSS to account for (1)–(3) is to say that indexed domain restriction is the same as the explicit domain restriction. Second, we see no reason why a speaker can’t stipulate beforehand that she wants a certain domain to be the most contextually salient aspect of our conversation,

and then merrily proceed with (10). Even in such circumstances, what she expresses is not, in any sense, rendered necessary.

Third, even if some sort of screening-off strategy could be made to work in the simple cases we have been discussing—though we don't see how—it would not help to avoid counterintuitive consequences for more complicated cases. Contexts may or may not be accurately representable as sequences of items that context-sensitive expressions can take as semantic values (<speaker, addressee, time, place, topic of discourse, perceptually salient objects, etc.>),⁵ but we presume it's not controversial that in any given context at most finitely many domains are salient. Let C be an ordering $\langle a_1, a_2, \dots, a_n \rangle$ of every salient domain, and then try to quantify over computers in that domain using (11).

(11) Every computer j (i).

where ' i ' indexes a member of C , and ' j ' specifies the predicate of which ' i ' is a constituent. ' j (i)', e.g., might be 'is in room 401'.

No such envisaged instances of (11) seem to express a necessary truth (or falsehood) in C , yet counterparts of form (12) can.

(12) Every computer (i_1) j_1 (i_2).

(as would 'Every computer in 401 is in 401'). Suppose, because ' i_2 ' occurs in the predicate in (11), it follows that *distinct* contextually salient domains must be indexed by ' i_1 ' and ' i_2 ' in (12). But then consider a new sentence with enough disjoined predicates that each potential domain in C can be picked out by a distinct index, as in (13).

(13) Every computer $j_1(i_2)$ or $j_2(i_3)$ or \dots $j_m(i_n)$.

According to the supplementation proposal under consideration, an utterance of a sentence of form (13) expresses in a context C what an utterance of a sentence of form (13') would.

(13') Every computer (i_1) is $j_1(i_2)$ or $j_2(i_3)$ or \dots $j_m(i_n)$.

But then (12), which expresses a seemingly contingent claim about computers, has been transformed into a sentence that in that same contextual setting has a necessary reading, without a possibility of further screening off. That we can devise such sentences might convince you something is fundamentally wrong with CSS strategies.

By promiscuously forcing the *semantics* to acknowledge contextual aspects not explicitly represented as determinants of what is expressed on an occasion

of use, supplementation strategies compromise a pragmatic, but semantically independent, alliance between context and linguistic meaning. No one can (or should) deny that contextual salience is exploited by a speaker and auditor in their effort to converge on some point or other. But most of us thought that we were being subtle and creative in so doing and not that we were obliged to as a matter of meaning alone.

KRIPKE'S TEST AND INT₁–INT₃

Kripke (1977) suggests that when some phenomenon P is alleged to show that a semantic theory T is mistaken, then imagine a language L, which is just like English with the (possible) exception that T is stipulated to be true of L. If P still occurs in L, then P is not evidence against the view that L = English.

A proponent of CIS can apply the test as follows. INT₁–INT₃ are alleged to show that CIS is wrong. To test that claim, imagine a language, call it CIS-E, which is like English in all respects, with the (possible) exception that CIS is stipulated to be true in CIS-E. If speakers of CIS-E would have intuitions INT₁–INT₃, then these intuitions do not count as evidence against the hypothesis that CIS-E = English.

We shall argue that speakers of CIS-E would have such intuitions. Our hypothesis is that these intuitions (for us and for CIS-E speakers) are generated by our indirect reports. In order to run Kripke's test, we first have to say a bit about how the practice of indirect reporting functions in English (and in CIS-E).

The Connection between Semantic Content and Reports of What Was Said/Asserted/Claimed/Stated/Committed to

In earlier work (Cappelen and Lepore 1997) we pointed out that reports of the form 'S said/asserted/stated that p' can be true even if S never uttered a sentence that expressed the proposition that p. If a speaker, S, utters (14) in a sarcastic tone of voice,

(14) You're a good friend.

it would typically not be correct to report that speaker as having asserted, claimed, stated, or committed himself to the addressee being a good friend. It would however, be correct to report S as having asserted/claimed/stated and committed himself to the addressee not being a good friend. This is a rather uncontroversial case of a true report of the form 'S claimed /asserted/committed himself to/stated p,' where p is a proposition pragmatically expressed by the reported utterance. (To simplify, we will in what follows focus on 'said that' reports, but our examples extend equally to reports of the form 'asserted that', 'claimed that', etc.) This is not a peculiar case. Close examination of the practice of indirect reporting reveals that this case is typical.

For example, Al, convinced that Stanley is Smith's murderer, says, looking at Stanley, 'Smith's murderer didn't comb his hair today'. Bill wants to report Al's utterance to Harriet, whom Bill knows is convinced of Stanley's innocence. Since Bill knows Harriet is unaware both of Al's contrary belief and of the context of Al's utterance (staring at Stanley), he might report Al to Harriet as having said that *Stanley didn't comb his hair today*. His so doing is appropriate, but is it correct? Even a philosopher whose deepest conviction is that definite descriptions are not ambiguous between referential and attributive uses can accept Bill's report of Al's utterance as true.⁶ Or, suppose the domain of computers under discussion includes only those in the speaker's father's office, but that the speaker is unaware of this fact about the contextually salient domain. (Say, he's not aware of the fact that room 401 is his father's office, or, that his father got the office after the relevant utterance of (7).) It is easy to see that the speaker still can be reported as having said that all the computers in his father's office were stolen.

One reaction to this sort of case is that, though someone might *report* an utterance of (7) as saying that all the computers in room 401 were stolen or that all the computers in the speaker's office were stolen, since these reports disagree about what was said, at least one of them *must* be wrong. The soundness of this reaction requires holding that an utterance of an (unambiguous) sentence can be used to say no more than one thing. So anyone who reacts this way to our alleged data must hold that:

The relationship between an utterance of a sentence and what's said by that utterance prohibits a speaker from saying both that p and that q with a single utterance, if p and q are distinct propositions.

But, then, how are we to explain that an utterance of 'Rudy loves New York and New Jersey' can say that Rudy loves New York, as well as that he loves New Jersey, as well as that he loves New York and New Jersey?⁷

Amending this restriction on what can be said with a single utterance with "unless the proposition ('immediately') follows from whatever is said" won't explain how Clinton, in detailing a new economic program, says that he will not cut taxes—even though none of his words express or imply this proposition. Is that what Clinton really said? Why, yes. But did he literally say it, or strictly speaking, is this what he said? Are you asking me for a direct quote? If not, then that's what he literally or strictly speaking said. Can you imagine his denying it, or the press recanting, "Well, yes, he sort of said it, but he didn't literally say it?"

Or take Francois who, in uttering 'Amethyst is Maria's favorite color', not only says that amethyst is Maria's favorite color, but also that *the color of that stone* is Maria's favorite color. The latter clearly does not follow from what his words expressed. But if he has not said both, why, then, is it acceptable to re-

port him as such? Indeed, it is *inappropriate* to report him as saying that amethyst is Maria's favorite color, if you know your audience is unfamiliar with the color word 'amethyst'.

To continue, consider a professor who, when asked whether Alice failed her exam, replies, "I failed no one." Has he said that Alice passed? If not, why is it correct to report him as such, in which case what is said makes reference to Alice, though his words do not? Imagine telling poor Alice that her professor didn't say whether she failed or not. Or, that, strictly speaking, or literally, or actually he didn't say. That would be a lie.

A skeptic might conclude that actual reporting practices are irrelevant in determining what is said. But how could anyone reach this conclusion without letting a theory override our practices? Still, one might protest, there are limits to what can be said with an utterance. Who would deny that? These various examples establish no more than the fact that delimiting *a priori* what those limits are is not only difficult, but inadvisable. For example, suppose Frank utters (7) in a context in which only the computers in room 401 are salient. Later, when another set of computers have become contextually salient (say, the computers in some other room), a question arises about what Frank said with his earlier utterance. Anyone who reports Frank in this context as having said that every computer was stolen might grossly misrepresent what he said, but anyone who reports him as having said that every computer in the room right under this was stolen has gotten him just right. If you disagree, how would you in the context described usefully and correctly answer the question "What did Frank say?" Would you conclude that you cannot?

Some rather big issues are at stake here. According to Frege (or at least some Fregeans), what is said by a single utterance is a single proposition. On this picture, a speaker says that *p* just in case he assertively utters a sentence that expresses the proposition that *p*. Accordingly, a speaker's words make reference to an object or a domain just in case what he says with those words does as well. Therefore, any indirect report of his utterance according to which what he says makes reference to an object or a domain that neither he nor his words does, *misrepresents* what was said. However, what our various examples are supposed to have established is that correctly determining what is said by an utterance often requires attending to *noninterpretive*, *nonsemantic* considerations. When we try to represent or articulate what is said by an utterance we aim to characterize a speaker's act (that utterance). In so doing, our interests often are not in systematicity or generality, but rather our aim is to determine something about a particular act in a particular context *C* in order to pass it along to a particular audience situated in a (perhaps a very) different (sort of) context *C**. In effect, our practice of reporting others treats what is said as a four-place relation between a sentence and its context of utterance and a reporting sentence

and its context of utterance. In determining what is said, we obviously draw upon information about specific intentions, knowledge, and history of the speaker in *C* and, not so obviously, we can also draw upon *like features of C**, the context in which we report what is said. Our reporting practices clarify that semantics should not *a priori* constrain what can and cannot be said by an utterance. Competent speakers make such judgments all the time, often relying on information that exceeds anything expressed or meant. This competence consists, in part, in a capacity to judge whether a report about what is said is accurate or misleading. Theorists who try to systematically incorporate contextual cues into semantic accounts of what is said seek to theorize about a practice that does not admit of it.⁸ There is no reason to believe that determining what is said will be simpler or more systematic than determining whether two items are similar.⁹

Kripke's Test Revisited

CIS-E, remember, is like English in all respects, with the (possible) exception that CIS is stipulated to be true in CIS-E. In particular, indirect reports function as in English, i.e., as outlined above. If speakers of CIS-E would have intuitions as those described in INT₁–INT₃, then the fact that English speakers have these intuitions does not count as evidence against the hypothesis that CIS-E = English.

It is intrinsically difficult to give a general account of how a people end up having certain intuitions. In general, we philosophers know far too little about this process, and it is an area where more empirical work is needed. But absent such work, here is an educated conjecture about how INT₁–INT₃ can be generated by utterance of, for instance, (7) for speakers of CIS-E.

Take a typical utterance, *u*, of (7) in CIS-E. Let us suppose that this utterance occurs in a context where it is clear to all participants that only a certain limited domain of computers are under discussion, say the computers in room 401. The proposition expressed by *u* will be false because there is a computer, somewhere, that was not stolen. Not only will *u* express an obviously false proposition, it will also be an entirely irrelevant proposition because it quantifies over computers no participants in the conversation have an interest in. That of course raises the question of why any speaker of CIS-E would utter (7) under such circumstances. The answer is fairly obvious. The speaker knows, and the audience knows, that only a certain set of computers is relevant. So the audience (and those reporting on the utterance from another context [like we theorists]) infers that what the speaker said, asserted, claimed, etc., is something other than the proposition expressed. They could for instance say things, e.g., such as (7r), (7r'), or (7r").

- (7r) She said that all the computers in room 401 were stolen.
 (7r') She asserted that all the computers in room 401 were stolen.
 (7r'') She claimed that all the computers in room 401 were stolen.

This is why speakers of CIS-E would have intuitions like those described in INT₁–INT₃. The intuition that someone who uttered (7) asserted, claimed, stated, committed to something true, is explained by the fact that (7r)–(7r'') are true and that the complement clauses of those reports are true.

So, our suggestion is that the intuitions on which proponents of CSS put so much weight are generated by true indirect reports. Where they go wrong is in not recognizing that such reports do not reflect the semantic content of the reported sentence. Stanley and Williamson have the intuition that an utterance of (7) can be true. That is an intuition you are likely to have if you think that by uttering (7) the speaker asserted and said something true. Reimer's (1998) intuition that the proposition expressed is true is, we think, a result of being steeped in a philosophical tradition that consistently conflates the proposition expressed with what was said. As a result of this theoretical bias, we philosophers are particularly prone to misconstrue (or misinterpret) intuitions generated by indirect reports.

So speakers of CIS-E are as likely to have intuitions INT₁–INT₃ as speakers of English are (especially CIS-E speakers who happen to be philosophers brought up the sort of Fregean tradition described above). In other words, the alleged evidence against CIS is impotent.

SECOND OBJECTION TO CSS: INDETERMINACY

Indeterminacy and Wettstein's Observation

We now turn to our final objection to CIS. Wettstein (1981) made a much-discussed observation about various attempts to account for how domain restrictions are determined. Take some context C, in which an utterance of (6) intuitively says that Nina took the car we own. Wettstein points out that there is often no unique domain restriction the speaker has in mind, and other features of context are often incapable of determining one enrichment as the correct one. There are typically any number of ways in which to enrich the utterance, all equally compatible with all contextual features. In C, for instance, the car might be 'more fully described in any number of ways, by the use of any number of non-synonymous, uniquely denoting descriptions' (Wettstein 1981, 262). Suppose that when asked 'Which car?' the speaker would consider (6a)–(6c) equally good answers.

- (6a) The car that was parked outside a few minutes ago.
 (6b) The car we bought last week.
 (6c) The car we own.

The speaker thinks these pick out one and the same car, and neither one of them was more prominent in her mind at the time of utterance. If asked, she would consider them equally good answers to the question, ‘What do you mean by ‘the car’? Which car?’ According to Wettstein, it is “implausible in the extreme” to assume that either the speaker’s intention or the context of utterance could select one of these as the correct completion.

The same sort of indeterminacy can be found in connection with uses of (7). When asked “What do you mean by ‘every computer?’” the speaker might consider (7a)–(7c) equally good answers, and, again, appealing to the speaker’s intentions (or other relevant aspects of the context of utterance) will not provide a unique completion.

- (7a) Every computer in Professor Smith’s office
- (7b) Every computer owned by the philosophy department
- (7c) Every computer in the office to the left of the speaker’s father’s office

In general terms, Wettstein’s Observation (WO) is this:

(WO) Appeal to speaker’s intentions and other aspects of the context of utterance will typically not suffice to yield a unique completion of an incomplete quantifier; there are typically equally good alternatives (some of which are not even co-extensional).¹⁰

It is fascinating to see how proponents of various versions of CSS attempt to deal with this issue. Stanley and Szabo (2000) serve as an illustration in this respect. Stanley and Szabo use an argument by elimination in favor of their version of CSS. They first argue that what they call the syntactic ellipsis theory and the pragmatic theory fail to account for the data. This, they claim, leaves their theory as the only tenable alternative. For the purposes of this illustration, the crucial point is that their only serious argument against the ellipsis theory is what they call the *under-determination problem* (and what we call the *indeterminacy problem*). The ellipsis theory is the view that a sentence such as (7) is elliptical for a sentence of the form,

- (7) Every computer which is F was stolen,

where ‘F’ is some unarticulated predicate, determined in context. Their central objection to this view goes as follows:

If context has to provide a specific predicate whose extension will contribute to the determination of the domain, a solution to the foundational problem involves specifying the relevant features of the

context which selects the predicate F among other candidates. And it is exceedingly hard to see what feature of the context could do that.¹¹ (Stanley and Szabo 2000, 37)

That is, WO provides the basis for their objection to the syntactic ellipsis account. Since they argue by elimination, it is also a central argument for their own positive view. The fascinating part is this: According to their view, the logical form of (7) is (and we simplify slightly) (7**),

(7**) [Every x : x is a computer & x is in i](x was stolen)

where the value of ‘ i ’ is a property, and the domain is restricted to the intersection of *computer* and i . Since the under-determination problem blows the syntactic ellipsis account out of the water, one should expect that Stanley and Szabo would tell us how their theory deals with it. If it is “exceedingly hard” to see how context can choose among different completing predicates, isn’t it at *a bit difficult* to see how contexts can choose among different completing properties (especially since these, as we have seen, do not even need to be coextensive)?

The surprising part is that this issue is not even addressed in their paper. This illustrates the attitude many proponents of CSS have toward the indeterminacy problem. They consider it a serious objection to certain versions of CSS, but fail to tell us how it can be solved with respect to their preferred version.¹²

We realize that this kind of *ad hominem* argument does not show that proponents of CSS *cannot* deal with WO. We will briefly outline why we think the prospects are bleak. A proponent of CSS has two options. She can deny Wettstein’s Observation and claim that there is something about the context that singles out a unique domain restriction. Call this view “Unique Completionism.” The alternative strategy is to say that there is no need to choose between different domain restrictions. They all get expressed. Call this strategy “Multiple Completionism.”

Both of these options are problematic. The problem with Unique Completionism is that no one has a clue as to how a unique completion can be selected. Even those who have suggestions, such as Soames, recognize that they fail in general.¹³ Inevitably, proponents of Unique Completionism just leave open exactly how to respond to WO. In effect, there is not one single suggestion in the entire literature on this subject for how to find a unique completion. For that reason, we refuse to discuss that option further until someone says something useful about this indeterminacy problem. We recommend others follow our lead.

Schiffer presents (but does not ultimately defend) a version of Multiple Completionism. He says:

You did not definitely mean any general proposition in uttering ‘the guy is drunk’ but you sort of meant, or vaguely meant, several general propositions, one for each definite description that could be used to sharpen what you vaguely meant. And your indeterminate statement might reasonably be held to be true just in case it is true under every admissible sharpening of what you meant, false, just in case it is false under every such admissible sharpening, and neither true nor false if it is true under some admissible sharpenings while false under others. (Schiffer 1995, 377)

Here is a counterexample to this suggestion. Consider Schiffer’s examples of a clearly intoxicated speaker approaching the podium. A member of the audience says:

(15) I’ll be damned! The guy is drunk.

According to Schiffer, (15.1)–(15.4) are all equally good completions of the underspecified quantifier ‘the guy’.

(15.1) The author of *Smells and Tickers*

(15.2) The only man within sight wearing a yellow jacket and red golf pants

(15.3) The man we are waiting to hear

(15.4) The man now staggering up to the podium

On the Multiple Completionism approach, an utterance of (15) is true only if every sentence in which ‘the guy’ is completed by (15.1)–(15.4) is true. But, ask yourself whether the utterance (15) would lack a truth value just because it turned out that the man, famous as the author of *Smells and Tickers*, actually stole the manuscript from an unknown German, von Trickers. The fact that the speaker had a false belief about the man does not imply that he did not assert something true. It does not imply that there is no true report of the form ‘He asserted that p and p is true’. So, truth of all acceptable completions (where the core of these are determined by what the speaker would accept as completions) is not required for the speaker to have asserted something true.

How Indeterminacy Is Predicted by CIS

So WO is a problem for CSS. Both solutions suggested in the literature are problematic. No one has a clue as to how to develop Unique Completionism, and there are limitless counterexamples to Multiple Completionism. These are all reasons for not endorsing CSS.

CIS, on the other hand, has no difficulty at all dealing with WO. Not

only does CIS not have a problem dealing with WO, it predicts WO. Our practice of indirect reporting is, as pointed out above, not restricted to reporting on the semantic content of utterances. Reports of what a speaker asserted/is committed to/said are sensitive to pragmatic aspects of the original utterance and to various aspects of the context of the report itself. As a result, there is typically not just one true report of the form ‘S asserted/said that . . .’ of a particular utterance. So in any particular context, we should expect that several reports will be equally correct. It should not be surprising, for example, that in some particular context, all of (7.1)–(7.3) are correct reports of an utterance of (7) by A.

- (7.1) A said that every computer in Professor Smith’s office was stolen.
- (7.2) A said that every computer owned by the philosophy department was stolen.
- (7.3) A said that every computer in the office to the left of the speaker’s father’s office was stolen.

This feature of indirect reporting is the source of WO.

NOTES

1. In the context of this paper we will ignore questions about tense.
2. Throughout this paper, we assume that descriptions are quantifiers, in accordance with Russell’s theory (Russell 1905, 1919).
3. See also Stanley and Szabo 2000 and Westerstahl 1985.
4. For a more developed defense of this charge, see Lepore 1999 and Cappelen and Lepore 1997.
5. In the same way, CSS predicts necessarily false propositions where none are forthcoming.
6. See Montague 1974, Kaplan 1989, and Lewis 1970.
7. Though we will not argue for it here, it is easy to see that we can use a description attributively to report someone who used it referentially, and *vice versa*.
8. Notice that if Rudy loves New Jersey but not New York, then, though the original utterance is false, at least one report of it attributes a truth—namely, Rudy said that he loves New Jersey.
9. For further discussion of this point, see M. Richard 1998, and Cappelen and Lepore 1998.
10. Salmon (1991, 88–89) seems to endorse a similar view, though his use of ‘literally saying’ vs. ‘the loose or popular sense [of ‘say’]’ is *incompatible* with the facts. There is nothing loose about reporting what was said by an utterance of (1) as described above; indeed, each may be a literal report. Furthermore, in correctly reporting an utterance with complement that does not express the proposition expressed, Salmon arbitrarily constrains what is acceptable, suggesting, wrongly we believe, that such departures are disguised *de re* reports (see p. 88). See also Cappelen and Lepore 1997.
11. That the alternatives do not need to be extensional should be obvious. The speaker might believe that all F’s are the same as all G’s, and hence be indifferent between the two completions, even though some G’s are not F’s.
12. For further illustrations, see Soames 1986, 309n7; Neale 1990, n55; and Reimer 1992, 62.
13. See Soames 1986, 279, and then the retraction (1998, 301n7). For more illustrations, see the other references in note 12.

REFERENCES

- Blackburn, W. 1988. "Wettstein on Definite Descriptions." *Philosophical Studies* 53: 263–78.
- Cappelen, H., and E. Lepore 1997. "On an Alleged Connection between Indirect Quotation and Semantic Theory." *Mind and Language* 12: 278–96.
- . 1998. "Reply to Richard and Reimer." *Mind and Language* 13: 617–21.
- Kaplan, D. 1989. "Demonstratives." In J. Almog, J. Perry, and H. Wettstein (eds.), *Themes from Kaplan*. New York: Oxford University Press.
- Kripke, S. 1977. "Speaker Reference and Semantic Reference." *Midwest Studies in Philosophy* 2: 255–76.
- Lepore, E. 1999. "Is There a Problem with Incomplete Definite Descriptions?" Manuscript.
- Lewis, D. 1970. "General Semantics." *Synthese* 22: 18–67.
- Montague, R. 1974. "Pragmatics." In *Formal Philosophy*. New Haven, Conn.: Yale University Press.
- Neale, S. 1990. *Descriptions*. Cambridge, Mass.: MIT Press.
- Recanati, F. 1996. "Domains of Discourse." *Linguistics and Philosophy* 19: 445–75.
- Reimer, M. 1998. "Quantification and Context." *Linguistics and Philosophy* 21: 95–115.
- . 1992. "Incomplete Descriptions." *Erkenntnis* 37: 347–63.
- Richard, M. 1998. "Semantic Theory and Indirect Speech." *Mind and Language* 13: 605–16.
- Russell, B. 1919. *Introduction to Mathematical Philosophy*. London: Allen and Unwin.
- . 1905. "On Denoting." *Mind* 14: 479–93.
- Salmon, N. 1991. "The Pragmatic Fallacy." *Philosophical Studies* 63: 83–97.
- . 1982. "Assertion and Incomplete Definite Descriptions." *Philosophical Studies* 42: 37–45.
- Schiffer, S. 1995. "Descriptions, Indexicals and Belief Reports: Some Dilemmas." *Mind* 104: 107–31.
- Soames, S. 1989. "Direct Reference, Propositional Attitudes, and Semantic Content." *Philosophical Topics* 15: 47–87.
- . 1986. "Incomplete Definite Descriptions." *Notre Dame Journal of Formal Logic* 27, 349–75. Reprinted in G. Ostertag (ed.), *Definite Descriptions: A Reader*. Cambridge, Mass.: MIT Press (1998). Page numbers cited in this article are from the 1998 volume.
- Stanley, J., and Z. Szabo. 2000. "On Quantifier Domain Restriction." *Mind and Language* 15: 219–61.
- Stanley, J., and T. Williamson. 1995. "Quantifiers and Context-Dependence." *Analysis* 55: 291–95.
- Westerstahl, D. 1985. "Determiners and Context Sets." In J. van Benthem and A. ter Meulen (eds.), *Quantifiers in Natural Language*. Dordrecht: Foris.
- Wettstein, H. 1981. "Demonstrative Reference and Definite Descriptions." Reprinted in H. Wettstein, *Has Semantics Rested on a Mistake? And other Essays*. Stanford, Calif.: Stanford University Press, 1991.

Deferred Demonstratives

Emma Borg
University of Reading

A BRIEF SURVEY of utterances containing demonstrative expressions in natural language reveals a perhaps surprising range of acceptable cases. For instance,

- (1) “That’s mine” said while pointing at a toy.
- (2) “You can have that one, I’ll have this one” said indicating first one bed, then another.
- (3) “This is a great composer” said while holding up a recording of Beethoven’s *Moonlight Sonata*.
- (4) “That’s a bear” said while indicating a paw print.

However, when philosophers think about demonstrative utterances, they tend to focus immediately on cases like (1–2)—which I will call ‘perceptual’ cases—where an object is currently seen (or otherwise perceived) by all interlocutors, pointed at (or otherwise demonstrated by the speaker), and thereby becomes the subject of the utterance. Those working in the tradition of truth-conditional semantics and concentrating on perceptual cases often dismiss other types of occurrences, like (3–4) (which I will group together under the heading of ‘deferred’ expressions) as in some way deviant or parasitic on the chosen use of the referential expressions in (1–2).¹ For instance, Kaplan talks of them as ‘contextually appropriate, though deviant’ uses of demonstratives; while Evans, considering a case like (4), is clear that they do not meet his standards for Russellian referring terms, as perceptual demonstratives do.²

In this chapter I explore the phenomenon of deferred demonstratives, examining the apparent difference between these and perceptual occurrences, and seeing how such a difference might best be accommodated within a semantic theory for these expressions. The first and, I will suggest, most immediately appealing move is to see some kind of parallel between deferred expressions and descriptions; what is said by an utterance like (3) is thought to be in some way synonymous

with an utterance like ‘The composer of this (piece of music) is great’. This intuitive idea will be later spelled out in two different ways; however, both these accounts will be seen to face difficulties. Further, I suggest that these problems indicate a deeper mistake, for, contrary to our initial hypothesis, deferred expressions do *not* behave in a way at all similar to descriptive phrases; rather, in all the relevant contexts, they behave just as ordinary referring terms do. In the final discussion I sketch an account of these expressions that handles them as semantically akin to ordinary, perceptual demonstratives and suggest that this understanding of deferred expressions sheds some light on the kinds of features that should be taken as constitutive of referential status for expressions in natural language.

DEFERRED DEMONSTRATIVES: EXPLORING THE PHENOMENON

In deferred cases, although there is an ostensive gesture, what seems to matter is not primarily the indicated object but some further object lying in a conventionally recognized relation to this first object. That is to say, in deferred cases we seem to be picking out an object to be talked about just in case it lies in some appropriate relation, or satisfies an appropriate relational predicate, one place of which is filled by the object being pointed at. This can be seen more clearly by looking at an example. For instance, consider the following scenario: you and I are walking around the Hermitage testing our knowledge of famous painters. Rounding a corner and seeing a large seascape painting entitled ‘Harbour View’, I point to it and say authoritatively, “That’s Turner”; but you, with a greater knowledge of classical French painters, correct me by saying, “No, that’s Claude Lorrain”. Now clearly, what neither of us wants to claim is that the painting being pointed at *is* the named individual; rather it seems that we intend to be taken as claiming that *the person who painted that picture* is the named individual.³ We seem to be picking someone out by description. Furthermore, it seems that we need know very little about this further person in order to talk about them using a deferred demonstrative. For instance, I might reply to you: “Well, that’s a great artist, whoever it is”; yet the possible inclusion of this kind of interpolation (i.e., ‘who/which-ever’), and the lack of any need for perceptual or ‘acquaintance’ knowledge of the object talked about, have often been taken as characteristic of quantificational, as opposed to referential, expressions.⁴

It seems an intuitively appealing move, then, to treat these deferred occurrences of demonstratives as in some way akin to descriptive phrases. And indeed we can find theorists in the literature apparently adopting this kind of proposal, at least for what we might call ‘deferred indexicals’. For instance, with respect to a similar case to (4) above, save that it involves the use of a pronoun rather than a demonstrative, Schiffer (1995, 123) writes,

Indexicals arguably do have attributive uses. For example, upon encountering a huge footprint in the sand, you might exclaim, “He must

be a giant!”, and arguably what you would mean is that the man whose foot made the print, whoever he is, must be a giant.

Recanati (1993, 230) also endorses the intuitive force of the move to treat deferred expressions as in some way akin to descriptions.⁵ However, as is well known from the literature on definite descriptions, there are divergent ways to take this kind of suggestion, for it may be thought to recognize a semantic or a pragmatic level phenomenon. That is to say, if we think that demonstratives have an attributive (i.e., descriptive) occurrence, then we might take this as a semantic feature or merely an aspect of their use. If it is a semantic phenomenon, then the literal meaning of (3) would be a descriptive phrase such as:

(3') The composer of this (piece of music) is great.

If it is a pragmatic phenomenon, however, we might retain an ordinary (referential) semantic interpretation for the demonstrative, but maintain that some alternative, descriptive proposition is also conveyed in the deferred cases.

The implications of adopting the first, semantic level, Descriptivist approach to deferred demonstratives should not, however, be underestimated, for if it were to prove correct, we would then be forced to relinquish a background assumption about the way in which semantic categories and surface forms hook up (at least as far as object words are concerned), which many theorists have seemed keen to preserve. The assumption is that, given some 'sub-NP' (sub-noun phrase) category, such as definite description or demonstrative, which can be recognized on the basis of quite superficial (say orthographic or phonetic) features alone, this category will map as a whole to a given semantic category. Of course, such an assumption is most familiar from the debate about definite descriptions where it is often held correct to treat all expressions of the form 'the F' as members of a single semantic category (although, of course, *which* single semantic category this is to be remains contentious).⁶ However, if we now embrace an account that treats some occurrences of 'this' and 'that' as referring terms and some as disguised descriptive (and hence quantified) phrases, then this would clearly undermine the general assumption that we can map from such sub-NP categories as a whole to semantic kinds. If it turns out that we have to treat demonstratives as semantically ambiguous in this way, one might well be led to expect a similar approach for other sub-NP categories, like definite descriptions. On this approach deferred demonstratives belong with other occurrences of expressions that lend support to the idea that surface form is a merely defeasible guide to semantic category, e.g., 'referential' and 'incomplete' descriptions, and anaphoric demonstratives. Thus the repercussions of adopting a semantically Descriptive approach to deferred demonstratives resonates beyond merely our understanding of 'this' and 'that'.

Yet accommodating the Descriptive approach to deferred cases at a seman-

tic level is, as noted above, just one option here. Alternatively we might retain an ordinary referential semantic analysis for deferred utterances, but hold that a descriptive proposition like (3') is the one *pragmatically* conveyed by such utterances (perhaps because the literal proposition flouts some Gricean principle of good communication). On this approach, although the speaker uses an expression which is understood to be semantically referential, whenever communication is successful she *always* conveys an alternative, descriptive proposition as the one meant. The pragmatic approach is clearly the weaker of the two proposals as it does not have the same implications for any background assumption about how surface forms and semantic kinds might tie up; yet it does still accommodate our initial assumption that deferred demonstratives are in some way akin to descriptions. So let us now turn to consider each of these proposals in more detail and see if either of them offer a satisfactory understanding of deferred expressions.

DEFERRED DEMONSTRATIVES ARE SEMANTICALLY AKIN TO DESCRIPTIONS

Initially, the idea is that deferred demonstratives should be taken to literally go proxy for some descriptive expression: when a speaker utters 'this' or 'that' intending to speak not about the indicated object but about some related object, then we should understand her as having said something semantically descriptive.⁷ However, it seems that this semantic level proposal quickly runs into difficulties. The problem is that there are multiple, semantically nonequivalent descriptions that *could* play the role of the deferred demonstrative in any context, and we need to know *which one* the deferred expression is supposed to mean. Yet the advocate of the semantically descriptive view can give us no answer to this question. For instance, in (3) both 'the composer of this piece of music' and 'the person who wrote this sonata' might be plausible replacements for the demonstrative, and so too might be 'the musician responsible for this' or 'the individual who composed this beautiful melody'; but how we are to choose the *one* such description that gives the literal, semantic value of the demonstrative seems quite unclear.⁸ Furthermore, it seems, on this proposal, that the meaning of 'that' must become 'contextually shifty': on one occasion, the semantic value of the expression-type will be given by one description, while on another exactly the same expression-type will mean something completely different, i.e., have its semantic value be given by a quite different description. Yet this kind of contextual shiftiness seems to be something we want to avoid.⁹

This problem of selecting among the plethora of fitting descriptions also becomes pressing once we realize (as was evident in the above quote from Schiffer) that the same kind of phenomenon occurs with pronouns, e.g., 'He has big feet' said while pointing to a pair of boots. Here, since there is even less descriptive material vocalized than in (3) and (4), the question of with which description we choose to replace the deferred expression seems even more underdetermined;

e.g., ‘the (male) owner of those boots’, ‘the man who wears these’, ‘the individual whose feet fit these shoes’ or ‘the person who bought this footwear’. All these seem to be entirely possible descriptive replacements for the original occurrence of ‘he’, and it seems that the Descriptivist owes us an account of how we choose among them. The most obvious option for the semantic Descriptivist at this juncture might be to appeal to the description the speaker has in mind. Yet such a move seems dubious as we have no guarantee that the speaker who chooses to use a deferred demonstrative is thinking of the object under some one particular description, nor, indeed, that the speaker has *any* description in mind prior to her utterance. Furthermore, even if she did have one such description in mind, we lack absolutely any way of guaranteeing that the speaker and audience will *coincide* in their selection for the privileged description. Yet without this kind of guarantee it is quite unclear how communication using deferred demonstratives is supposed to take place.

Finally, I suggest, we should seriously question any kind of semantic approach (like that now to hand) that assumes that we can simply ignore the apparent surface contents of a sentence, in order to posit a range of complex, unvocalized material, for which there is no good evidence. Though we must tolerate some, very limited, violations of the idea that all semantically present material makes it to the surface form of an utterance (i.e., that our semantic theory has a certain degree of freedom to posit meaningful elements that do not appear in the vocal or orthographic form of the expression), the kind of wholesale disregarding of surface form envisaged here is unwarranted. A semantic theory that tells us that part of the *literal content* of ‘That is a great composer’ includes ‘composed’, or ‘responsible’, or ‘sonata’, or so forth, simply diverges too far from surface constraints.¹⁰

So it seems that there are serious difficulties with the semantic Descriptivist position; despite the initial plausibility of the move to treat deferred demonstratives as akin to descriptions, it seems that the existence of multiple fitting descriptions, the lack of any guaranteed agreement between speaker and hearer on choice of description (and the subsequent threat to communication this entails), and the violation of surface form constraints, all mean that we cannot treat them as semantically analogous. However, recall that we have a second alternative way to accommodate our initial assumption, for we might instead treat all deferred utterances as *pragmatically* akin to descriptions.

DEFERRED DEMONSTRATIVES ARE PRAGMATICALLY AKIN TO DESCRIPTIONS

On this approach, whenever a speaker utters a demonstrative expression intending to speak not about the indicated object but about some related thing, though she uses a semantically referential expression, she *conveys*, as an act of speaker-meaning, a suitable descriptive proposition. So the speaker who utters (3) expresses a referential proposition, but conveys a descriptive proposition

along the lines of (3'). For this proposal to be feasible, however, it seems that we need some account of the mechanism by which the speaker succeeds in conveying the alternative descriptive proposition (in addition to the literal proposition expressed) in these cases.

Yet the advocates of the pragmatic Descriptivist approach have just such an explanation readily to hand, for it seems that they can claim all deferred demonstratives, taken literally, are trivially false. For instance, in (3) I indicate a record and say of it that it is a composer, while in (4) I point at a paw print and say, of that very impression in the ground, that it is a bear—but records are not artists and marks on the ground are not bears. So it seems in both (3) and (4) I make some kind of category mistake, but, as an apparently competent user of the language, it seems, as Grice (1967) has told us, that my utterance of trivial or obvious falsehoods should not be taken at face value. Instead, since I violate an obvious principle of good communication (roughly, do not say things which appear obviously false), I should be taken as pragmatically conveying some alternative, more suitable proposition (in these cases, a suitable descriptive proposition).

So, does this pragmatic Descriptivist proposal give us a good account of deferred demonstratives? Again, I think there are problems to be faced in adopting this kind of approach, for it seems the picture it gives us of sentences containing deferred demonstratives simply does not fit the way we expect these expressions actually to behave in natural language. Take first the putative parallel between deferred demonstratives and other cases where the speaker is supposed to utter obvious falsehoods in the hope of conveying some alternative (possibly true) proposition, e.g., occasions where the speaker is being ironic. In the latter case it is clear that the speaker intends to be taken nonliterally (she means to be taken as saying the opposite of that which she literally expresses); but in the former case things seem very different, for it seems that the speaker using a deferred demonstrative need have no such nonliteral intent. Indeed, the speaker who utters (3) is most probably not trying to be funny or ironic, but is simply trying to express the perfectly ordinary proposition that the composer she is talking about is good. Yet, in this case, it is hard to see why we should expect the speaker to reject all the (true) descriptive propositions she could have used to communicate literally what she wanted to say. We simply have no explanation of why a competent speaker in this kind of situation would opt for a trivial falsehood, merely in the hope of pragmatically conveying her intended proposition; nor have we any explanation of the speaker's own intuitions in these cases that she has not made any kind of semantic mistake. This leads us on to the second problem with the pragmatically Descriptive view; for the claim that all sentences involving deferred demonstratives are literally false entails that, in certain domains, we are constrained never to utter literal truths—a consequence that seems far too strong.

For instance, it seems that we can use demonstrative utterances to express what certainly appear to be literal truths about abstract objects, like numbers. I

can point first at an inscription of ‘2’ in the Arabic counting system and then at an inscription of ‘ii’ in the Roman numeral counting system, and say ‘That is the same as that’. Here it seems that I say something true, indeed learning this truth would be very useful to someone trying to acquire one or other counting system (given she already possessed the other). However, on the current approach we are forced, very much contrary to intuition, I would suggest, to hold that what I have said is literally false: I can *never* succeed in talking demonstratively about numbers, only in talking about their inscriptions. Thus any demonstrative utterance I make concerning numbers must be literally false and said only in the hope of pragmatically conveying some appropriate descriptive proposition to you. Yet what I said certainly seemed true enough—it certainly does not seem semantically on par with an utterance of ‘that is the same as that’ said while pointing first at ‘2’ and then at ‘iii’.

The claim that all occurrences of deferred demonstratives must give rise to literally false propositions seems both ad hoc and too strong. It leaves us forced to treat a great swath of natural language sentences as literally false, but with no explanation of why speakers are supposed to opt for obvious falsehoods, nor our own strong intuitions that at least some such utterances are literally true. So it seems that both the semantic and pragmatic renditions of the Descriptivist approach to deferred demonstratives quickly run into difficulties. However, rather than try to resuscitate these approaches in the face of such problems, I pause now to reconsider the basic assumption that has been driving the debate thus far. The assumption we made at the outset was that deferred demonstratives were in some way akin to descriptions; but the question we must face now is whether in fact this was a good assumption to make, for the behavior of these expressions in certain contexts seems to indicate it was misconceived.

THE BEHAVIOR OF DEFERRED DEMONSTRATIVES

The behavior that matters for us here is that displayed within those linguistic contexts where we expect genuine referring terms and quantified, descriptive phrases to come apart; these contexts include modal, empty, and scope contexts. The contention of this section will be that, in each of these contexts, deferred demonstratives behave like ordinary referring terms and not like descriptive phrases; thus our initial assumption—that deferred demonstratives demand treatment as in some sense analogous to descriptions—was mistaken. We must instead aim to treat these terms as akin to ordinary referential expressions.

To turn to modal considerations first: as Kripke has stressed, when we embed a description under a modal operator it (usually) behaves very differently to a referring term, like a name, in the same environment.¹¹ Descriptions can pick out different objects at different possible worlds (selecting whichever object at that world satisfies the relevant predicates), but a referring term picks out the same object at every world (or at least every world where the referent exists).¹²

It is because of this feature of the two kinds of expressions that (5) expresses a truth, while (6) expresses a falsehood, despite the fact that ‘the first president of America’ and ‘George Washington’ pick out the *same* object relative to the actual world:

- (5) John Hanson might have been the first president of America.¹³
- (6) John Hanson might have been George Washington.

Sentence (5) is true because there is a possible world in which, say, the Declaration of Independence was signed during Hanson’s term of office and he won the subsequent election to the post of president; however, there is no possible world in which the unique individual picked out by the name ‘John Hanson’ is the same as the (distinct) individual picked out by the name ‘George Washington’ (of course, John Hanson might have been called ‘George Washington’, but, as Kripke pointed out, this would not suffice to make him George Washington).

What then should we say about modal claims involving deferred demonstratives, such as

- (7) ‘That might have been my favorite composer’, said while pointing to a recording of Beethoven’s *Moonlight Sonata*.

If (7) contains a covert descriptive phrase such as ‘the composer of that’, as suggested in the previous section, then we should expect the sentence to pick out different objects at other worlds, i.e., to pattern alongside (5); whereas if (7) contains a genuine referring term we should expect it to behave more like (6), sticking with the same object across different possible worlds.

On the first (Descriptivist) proposal the worlds that would make an utterance of (7) true would be worlds where whichever object has the first property mentioned (being the composer of that) also has the second (being my favorite composer). So, let us say my favorite composer is Schoenberg, but I think that, had I played some musical instrument, I might have preferred a more harmonious and ‘playable’ composer; so I make the claim in (7). However, if we understand the deferred demonstrative in (7) along the Descriptivist lines suggested in the previous section, then it seems (7) will be made true by a world in which my actual favorite composer (Schoenberg) wrote *The Moonlight Sonata*. But intuitively this is *not* what (7) claims, rather (7) will be made true by a world in which my favorite composer is Beethoven and not Schoenberg.¹⁴ Yet this is just to say that the deferred demonstrative in (7) does *not* behave like a descriptive phrase, picking out whichever object satisfies some predicative expression, but rather that it behaves like a genuine referring term, picking out the actual referent of the expression and sticking with it across all possible worlds.¹⁵

The same phenomenon can also be seen with the so-called attributive or deferred use of pronouns that we noted earlier. Say Bill, on receiving a bad grade for his logic test, storms furiously from the room, slamming the door behind him. Indicating the door, it seems I can now say

- (8) ‘He’s upset, but he might not have been’.

Again, if ‘he’ here is understood as going proxy for some description like ‘the person who just slammed the door’, then (8) will be made true by any world where the door-slammer is not upset—say a world where happy Phil accidentally lets the door slip out of his hands and slam shut. Yet, intuitively, when I utter (8) what I say is not concerned with anyone but Bill: what I say is made true by a world in which Bill is not upset (say because he does not really care about his logic result, or where he worked harder and got a better grade), regardless of his other properties at that world, specifically regardless of whether or not Bill slammed the door in that world. So, once again, we have evidence that the deferred expression is correctly treated not as some kind of disguised description, but as an expression that succeeds in genuinely referring to the object in question.

Next, let us examine how these expressions behave in ‘empty’ contexts, i.e., contexts where there is no object answering to the use of the noun phrase. It seems that descriptions that fail to secure an extension still retain their usual content (since the claims they make concern the extension of predicates and not specific individuals that instantiate those predicates); but empty uses of referring terms are vacuous. As both Kaplan (1977) and Evans (1982) have emphasized, when a referring expression is used where there is no referent (say, the speaker is hallucinating) we simply cannot assign a truth-value to the utterance.¹⁶ Yet, once again here, deferred cases seem to pattern with referring terms and not with descriptions. Consider an utterance of ‘She’s gone’, pointing to an empty chair, but imagine that our speaker has been hallucinating the presence of a woman, whom he now believes to have left the room. In this scenario, I would suggest, the utterance seems devoid of all content, it is vacuous in exactly the same way that pointing at an empty region of space (while hallucinating) and saying ‘She’s happy’ is vacuous; in neither case does the speaker succeed in expressing a proposition. The mere use of an intermediary object to try and pick out the hallucinated object does *not* result in any content for the utterance, specifically it does not result in interlocutors taking a descriptive proposition like ‘the previous occupant of that chair has left’ as the literal content of what has been said.¹⁷

The last test I would like to appeal to in order to help us differentiate referring terms and descriptions turns on the recognition of scope ambiguities. As is well known, descriptive phrases offer scope ambiguities in certain contexts (e.g., those with other quantifiers, operators, and intentional verbs). A sentence such as

(9) Every girl desires the boy who won

is ambiguous between the following two interpretations:

- (9') [the y : boy who won y] [every x : girl x] desires $\langle x, y \rangle$
 (9'') [every x : girl x] desires [the y : boy who won y] $\langle x, y \rangle$

On (9') there is some particular boy whom every girl wants, whereas on (9'') every girl desires merely whoever it happens to be who won the race (race-winning figures as a part of the girls' intentional content). Compare this to a similar sentence containing an ordinary perceptual pronoun (and thus a paradigm referring term):

(10) Every girl desires him

where 'him' is accompanied by a pointing to a perceptually present individual. Here there is no ambiguity, only one reading is possible, viz., that every girl wants that particular individual.

So, now, what of deferred utterances such as:

(11) Every girl desires him

where this time the pronoun 'him' is accompanied by a pointing to a (now empty) chair. If 'him' here is read as equivalent to a description such as 'the person who was sitting in that chair', then we should get two readings for (11), viz.:

- (11') [the y : person who was in that chair, y] [every x : girl x] desires $\langle x, y \rangle$
 (11'') [every x : girl x] desires [the y : person who was in that chair, y] $\langle x, y \rangle$

But (11'') is not an admissible interpretation of (11). (11) does not tell us that every girl wants just whoever happens to have the property of having sat in the indicated chair (where this could be a different individual in different possible worlds). Instead it describes an intentional state of the girls directed toward a specific individual: they desire the person who, a few moments earlier, they could have picked out using a perceptual demonstrative, but who now can only be referred to through the use of an intermediary object. This failure of deferred expressions to give rise to different scope readings in appropriate contexts reinforces the results from modal and empty contexts, viz., that deferred expressions are simply not descriptive phrases.

So, then, it seems that if we look at the behavior of these expressions, we find evidence that the correct kind of analysis for deferred demonstratives will treat them as ordinary referring terms and not, as we initially supposed, as dis-

guised descriptions. Assuming that this is correct, the final task before us is now to deliver a semantic analysis that does actually treat these expressions in the desired way. In conclusion I would like to sketch one way in which we might achieve this end.

A REFERENTIAL ANALYSIS OF DEFERRED DEMONSTRATIVES

Recall what seems to be going on in a deferred demonstrative occurrence: there is an ostensive gesture (a pointing, a looking, a head tilt, etc.) to a perceptually present object, but this object does not thereby become the subject of the utterance. Instead some related object, not available for direct indication, is picked out. The idea so far has been that we could incorporate within our semantic rule the way in which the indicated object and the ultimate subject of conversation are related (for instance, by alluding to the relational property within a description). However, a question now emerges about whether we were right to do this, for on reflection the kind of information required to decide on the referent of an act of deferred ostension seems to go far beyond what we in general hope to capture within a theory of meaning for object words like ‘this’ or ‘that’.

For it seems that whether or not we can use ostension to one object to pick out some other object depends not on a predetermined list of ways in which objects can relate (such as we could give within a semantic theory), but rather on highly contextual and variable features, like social conventions, the common knowledge between interlocutors, their conceptual framework, and the salient features of objects. We can always tell a story that makes an apparently outlandish act of deferred ostension possible, even where two objects are not related in any way that previously appears on our putative list. Yet, prior to such a convention growing up, how were we to know whether the relationship it appeals to could underpin an act of deferred ostension? These considerations reveal that how we get from one object to another in deferred cases goes far beyond the meager resources of linguistic meaning (where this is understood as the kind of context-independent rules for understanding language envisaged by formal semantics).

The sorts of features that govern the move from indicated object to referent are simply not the sorts of things we can hope to make our semantic theory sensitive to, but nor should we want to. Understanding how objects relate to one another is a complicated matter, involving a great deal of world knowledge, which goes far beyond the merely semantic. Knowing that pointing at a picture is a good or acceptable way of making the painter of that picture salient, or that the referent of a deictic use of the word ‘he’ can sometimes be called to attention by indicating a set of footprints, are not things we should expect an analysis of the linguistic meaning of these words to give us. So if we need to treat deferred expressions as genuine referring terms, yet we cannot hope to capture the complex and context-bound relationship between indicated object and referent

within a formal semantic theory, where does this leave us? Well, perhaps surprisingly, it leaves us able to account for deferred utterances without straying very far from the now quite standard direct reference analysis of demonstratives—as long as we are careful about how we spell out the technical notions that this theory utilizes.

To implement this proposal, all we need to do is to separate the notions of ostensive gesture, e.g., pointings or head tilts, from the notion of demonstration, i.e., allowing that one object may be pointed at, while *another* is demonstrated. Liberating demonstration from an equivalence claim with ostensive indication allows us to adopt a quite standard rule for perceptual demonstratives as fitting for deferred demonstratives as well: for *all* demonstrative expressions the rule would simply be that the object demonstrated must satisfy the predicate mentioned in order for the utterance to be true. Thus:

- (D) In any context c , $d[\delta]$ is a directly referential term that designates the demonstratum, if any, of δ in c , and that otherwise designates nothing.¹⁸

Simply, ‘That’s a great composer’ is true just in case the object demonstrated is a great composer; what changes here is that the object demonstrated may *not* be the object pointed at (though, of course, there is nothing to stop it being this object). A rule like (D), because it extends to both deferred and nondeferred uses of demonstratives, satisfies constraints like simplicity and learnability: we do not have to posit any new rule that is acquired by the language learner and that suddenly accounts for their ability to use deferred expressions. Rather, the speaker who is competent with perceptual demonstratives *already* possesses all the semantic information they need to use deferred expressions (what they may be *lacking* is sophisticated enough world knowledge to put to use the semantic tools they have at their disposal). This kind of rule also passes what some have called a ‘Kripke-test’ for explaining the data: given a language just like English, save for the fact that we *specify* that the correct analysis for context-dependent expressions divorces demonstration and indication, we could expect the range of uses of demonstratives to mirror exactly the pattern of use we see in natural languages.¹⁹

One possible objection to the current proposal might be that it is in conflict with some sort of pretheoretical notion of demonstration we have, whereby it is simply identical with ostensive gesture. However, even were we to accept the role of pretheoretic intuitions in constraining formal semantics, it is quite unclear that we do have such a notion anyway. Talk of ‘demonstration’ in natural language usually involves doing proofs or illustrating actions (e.g., ‘he demonstrated how to open the milk carton’, ‘she demonstrated the incompleteness of mathematics’) and *not* ostensive gestures. The notion of demonstration being deployed with respect to demonstratives is already a semitechnical one,

which abstracts from the locutions of ordinary speakers.²⁰ Taking into account the vast range of contexts in which demonstratives can successfully be used illustrates that this semitechnical notion of demonstration must be broad enough to incorporate a vast range of ways in which an object may be demonstrated. These ways include such things as objects that are demonstrated through parts of them being pointed at; objects that are demonstrated through being talked about or raised to salience in discussion; and objects that are demonstrated or raised to salience through the indication of (nonmereologically) related objects. There are *lots* of ways to draw an object to attention prior to a referential act, and pointing to a related object is just one of these ways.

In retrospect, it seems clear that the sort of sharp distinction between deferred and nondeferred cases, assumed by the previous accounts, never really existed. Once we recognize that I can refer to you by pointing at your arm, or at just that part of you that is visible through a door that is slightly ajar, or at the tail of your coat as you leave the room, or at your image in a mirror, or at your photograph, or at your shadow, etc., the idea of drawing a semantic distinction at any point on this scale comes to seem quite hopeless. It is thus a strength of the current account that it does not need one, and a failing of other approaches, which assume that there is some delimitable class of deferred cases that mean something different to perceptual cases. Finally, however, at this juncture we might wonder what has become of the apparently strong intuitions, with which this chapter opened, which told us that there *was* a clear difference between cases like (1–2) and cases like (3–4); isn't the current proposal in tension with these intuitions?

The answer to this question is No. To claim that there is no semantic distinction is not to claim that there is no relevant distinction at all. Speakers *can* differentiate between deferred and nondeferred uses of demonstratives, but when they do this what they are sensitive to are nonsemantic features. For instance, they may be sensitive to the fact that in cases like (3–4) the relationship between the object pointed at and the object referred to is more oblique, and that more world-knowledge is required to identify the referent than is required by utterances like (1–2). Or again, they may be sensitive to the fact that, since no perceptual identification of the referent is presumed in deferred cases, they are less informative than other uses of demonstratives, where such 'up close and personal' knowledge of the referent is presumed. Yet, however identification of the referent is supposed to proceed, and however much 'world knowledge' versus perceptual knowledge it invokes, what understanding deferred uses of expressions drives home is that this is simply not an issue for a semantic theory. A theory of linguistic meaning must tell us what utterances of 'this' and 'that' mean, what contribution they make to larger sentences in which they occur, and widening our attention, from merely perceptual cases, to consider the vast range of ways in which demonstratives can be used to talk about objects, makes it clear

that object identification (in any substantial, nondescriptive sense) can be no part of such linguistic meaning.

This in turn supports the idea that the kind of epistemic condition sometimes taken to be constitutive of referential status, e.g., by Russell, is misplaced—if there are genuine referring terms (deferred demonstratives being a case in point) that flout any such epistemic constraint between speaker and referent, then clearly it cannot be constitutive of referential status that such an epistemic condition be fulfilled. If it is right, as I have tried to argue here, that deferred uses of demonstratives should be analyzed as ordinary referring terms then it cannot be the case that ordinary referring terms require nondescriptive object identification, for when I pick out an object via some related object I may clearly lack any such nondescriptive information. When I point at a record and say ‘That’s a great composer’ it may well be that *all* I know about the referent of my utterance is that they are the composer of the indicated record.

We have seen that deferred demonstratives should not, perhaps contrary to initial intuitions, be handled as semantically or pragmatically akin to descriptions; rather, I have argued, their behavior in modal, vacuous, and scope contexts reveals that they demand analysis as ordinary referring terms. I have suggested that the best way to spell out this insight is to treat deferred expressions as entirely semantically synonymous with ordinary perceptual demonstratives by divorcing the notions of ostensive gesture and demonstration. There is thus no semantic distinction between deferred and nondeferred expressions; rather, our intuitions about the difference in these roles for demonstrative expressions should be accommodated outside the semantic realm, in considerations like the relative cognitive usefulness of the utterance. As long as we are prepared to divorce demonstration and indication, the standard mechanisms of direct reference are opened up to deferred expressions. We can then have a semantic theory that offers us a univocal analysis of all such context-dependent referring terms, without giving up our strong intuitions that deferred cases are different to perceptual cases.

NOTES

Thanks are due to Kent Bach, Maite Ezcurdia, Kirk Ludwig, Michael O’Rourke, Nathan Salmon, and audiences at University of Florida, Autonomous University of Mexico (UNAM), University of Nottingham, and, of course, the Inland Northwest Philosophy Conference. Particular thanks must go to Peter Ludlow, whose excellent commentary at the INPC much improved this version of the paper, and to Ernie Lepore, whose help much improved innumerable earlier versions. Research for the paper was, in part, made possible by a grant from the University of Reading Research Endowment Trust Fund.

1. Cases like (3) and (4) were perhaps first brought clearly to light in Quine (1968), where he labelled the phenomenon ‘deferred ostension’; qualms about the suitability of this title, which will not be discussed here, lead me to adopt the truncated label ‘deferred demonstratives’.
2. Kaplan 1977, 490; Evans 1982, 199. See also Grimberg 1996, 424: “Pointing to a photograph to refer to a boy just isn’t the same thing as pointing to the boy himself, and

- models of demonstrative reference tend to come unstuck when such examples are introduced.”
3. Some may feel that this example, perhaps even together with (3), is really an instance of metonymy (i.e., the replacement of the name of an attribute for the thing meant, e.g., ‘the Crown’ for the King), rather than a genuine deferred demonstrative—what is meant is rather that *that is a Claude Lorrain*. However the example does, I believe, illustrate the intuitive force of a move from a demonstrative to a description, while it is clear that in general the phenomenon of deferred demonstratives cannot be simply collapsed into that of metonymy. For instance, pointing at a scarf and saying ‘this girl’, or at a picture of an elephant and saying ‘that animal’, cannot be treated as instances of metonymy (ordinarily understood).
 4. For instance, see Donnellan (1966), who took the interpolation test as one of the hallmarks of attributive uses of definite descriptions.
 5. We should note, however, that Recanati also explicitly resists the idea that deferred demonstratives can be understood as semantically akin to descriptions—the position to be explored first below.
 6. For instance, Frege (1879) and Russell (1905) tacitly agree on this assumption, though they disagreed on the semantic type all definite descriptions should map to; while the assumption forms an explicit commitment for many contemporary theorists, like Neale (1990).
 7. The suggestion that all demonstratives be understood as covert descriptions has been put forward by Neale (1993), though he considers only (complex) nondeferred expressions.
 8. This problem for the semantic Descriptivist should not, however, be taken as a positive point in favor of the view (to be advocated below) that deferred demonstratives are referring terms, since similar difficulties arise for some attributive uses of definite descriptions and no one wishes to argue that these are referential terms. I am grateful to a referee for this volume for raising this point.
 9. At the very least, adopting this kind of approach would require that we see the primary bearer of meaning and truth as the *utterance* of a deferred expression, and not the *sentence type* containing the demonstrative term.
 10. A possible objection to this line of argument (pointed out by Peter Ludlow) concerns the parallel between deferred demonstratives and E-type pronouns. If we accept the notion of E-type pronouns then *precisely* what we allow is that superficially simple expressions like ‘he’ and ‘it’ can be replaced at a semantic level by the complex descriptive material they go proxy for (and, *mutatis mutandis*, the argument goes, ‘this’ and ‘that’). Two remarks on this: first, the correct theory of anaphora itself remains a tricky business and simply because one theory of these expressions (a theory that itself faces problems) allows such semantic revisionism, this may not be enough to license similar violations of surface constraints in the common cases of deferred demonstratives. Second, the classic examples of E-type pronouns all involve recovering the descriptive material that the pronoun abbreviates from the surrounding *linguistic* environment, whereas, in the case of many deferred demonstratives, the required descriptive material will not have been vocalized at all. Thus, in the latter case, unlike the former, we cannot even hope for any kind of formal rule linking the demonstrative or pronoun to a verbal antecedent; rather the appropriate descriptive information, which is supposed to figure as part of the explicit semantic content of the utterance, must somehow be conjured out of the wider, nonlinguistic context, and this may serve to make the parallel with E-type pronouns seem somewhat strained. Ludlow is however right to point out that, pending a clearer understanding of the principles underlying warranted violation of surface form constraints, this may not be a point the opponent of semantic Descriptivism wants to rest too much weight on.
 11. The qualification ‘usually’ is necessary to take account of descriptions with ‘specialized’ predicates that maintain a stable extension across worlds, e.g., ‘the actual F’ or ‘the smallest prime’.
 12. See the debate between Kaplan and Kripke in Kaplan 1977, 492–93.

13. John Hanson was the first President of the Council of Confederate States, the forerunner of today's U.S. Congress.
14. For instance, we might imagine the utterer of (7) continuing with: 'But since I don't play any instrument, Schoenberg is my favorite'; but in so saying they explicitly state the fact that they expect their utterance of 'that' to pick out a composer *other than* Schoenberg.
15. As Peter Ludlow and others have pointed out, it may be that some deferred uses of demonstratives do seem to license a narrow scope descriptive reading; for instance, consider a conversation about the affairs of American presidents, where a speaker points to a podium bearing the presidential seal and says 'Interns always want to have an affair with him'. Here we *might* take the proposition the speaker conveys to be that *interns always want to have an affair with whoever speaks from that podium*. Personally, I find this interpretation somewhat strained; however, even if it is a licit reading in some cases this is not enough to show deferred expressions are nonreferential. For, even if the reading is possible in *some* cases, we might then explain this pragmatically (so that it is a kind of 'loose talk' that the charitable hearer is able to understand, despite its divergence from the semantic); whereas, if the reading is genuinely absent in some cases (as I think is the case with (7)) then this *demands* semantic level explanation.
16. This difference between descriptions and referring terms was especially emphasized by Evans (1982, 52 and 70–73).
17. It should be pointed out that intuitions about empty cases vary, with some theorists finding they do not expect a parallel between this kind of case and those where the demonstrative is used to talk 'directly' about the hallucinated object. So, even if one gets the vacuous reading in the latter case, in an empty deferred case one can still recover a proposition expressed, though a false one. For instance, in cases like (4): 'That's a bear' said while pointing at what looks like a paw print. If we here imagine that the print was made not by a bear but by a freak wind then, the suggestion is, we hear the utterance as expressing a perfectly acceptable proposition (e.g., 'the maker of that was a bear') which is, in this case, false. However, I would suggest that those cases where it seems possible to recover a (false) proposition are those where there *is* an object to be recovered as the referent of the deferred demonstrative, though this is one that does not possess the property being ascribed to it. For instance, in this use of (4) we get a proposition expressed, but one that refers to the (existent) thing responsible for the object ostended to (in this case, the wind) and says of it (the wind) that it is a bear. I would argue that, in the *radically* empty cases (like the one in the text), where there is genuinely *nothing* to play the role of the referent, it is in fact very hard to hear the speaker as expressing any proposition at all. For further discussion of the range of 'empty' cases possible, see Salmon (1998).
18. Kaplan 1977, 527, where '*d*' is a directly referential term and '*δ*' is a related demonstration.
19. See Kripke 1977; a similar kind of strategy is also employed in Bach 1975.
20. A point that Kaplan himself was clear on from the outset, cf. (1977, 490): "My notion of demonstration is a theoretical concept." We might also note that his talk of a 'salience platform' fits very well with the view of demonstration advocated above; although his repeated remarks that a demonstration gives a 'perceptual presentation' of an object seem to indicate that a more traditional view of demonstration is in play.

REFERENCES

- Bach, K. 1975. "Performatives are Statements Too." *Philosophical Studies* 28: 229–36.
- Donnellan, K. 1966. "Reference and Definite Descriptions." *Philosophical Review* 77: 281–304.
- Evans, G. 1982. *The Varieties of Reference*. Oxford: Oxford University Press.
- Frege, G. 1879. *Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*. Halle: Niemeyer.
- Grice, P. 1967. "Logic and Conversation" (William James Lectures). In P. Cole and J. Morgan (eds.), *Syntax and Semantics*, vol. 3. New York: New York Academic Press.

- Grimberg, M. 1996. "Pragmatically Determined Aspects of What Is Said: A Reply to Bezuidenhout." *Mind and Language* 11: 415–26.
- Kaplan, D. 1977. "Demonstratives." In J. Almog, J. Perry and H. Wettstein (eds.), *Themes from Kaplan*. Oxford: Oxford University Press.
- Kripke, S. 1977. "Speaker Reference and Semantic Reference." *Midwest Studies in Philosophy* 2: 255–76.
- Neale, S. 1990. *Descriptions*. Cambridge, Mass.: MIT Press.
- . 1993. "Term Limits." In J. Tomberlin (ed.), *Philosophical Perspectives 7: Language and Logic*. Atascadero, Calif.: Ridgeview.
- Nunberg, G. 1993. "Indexicality and Deixis." *Linguistics and Philosophy* 16: 1–44.
- Russell, B. 1905. "On Denoting." *Mind* 14: 479–93.
- Quine, W. V. 1968. "Ontological Relativity." *Journal of Philosophy* 65: 185–212.
- Recanati, F. 1993. *Direct Reference: From Language to Thought*. Oxford: Blackwell.
- Salmon, N. 1998. "Nonexistence." *Noûs* 32/3: 277–319.
- Schiffer, S. 1995. "Descriptions, Indexicals, and Belief Reports: Some Dilemmas (but Not the Ones you Expect)." *Mind* 104: 107–31.

What Unarticulated Constituents Could Not Be

Lenny Clapp
Illinois Wesleyan University

INTRODUCTION

IN THIS CHAPTER I clarify and take a side on an issue that currently divides theorists working in semantics and the philosophy of language. On one side of the issue are theorists who defend the traditional theoretical frameworks of Davidson's or Montague's semantic programs¹ and on the other side are those who favor rejecting these traditional frameworks in favor of some form of *dynamic semantics*, or *truth conditional pragmatics*.² My allegiance lies with the latter camp. I think there are many *semantic* phenomena that cannot be adequately explained within the more limited constraints of Davidson's and/or Montague's semantic programs. More precisely, I think these semantic programs cannot adequately account for the truth conditions of all assertions. Moreover, the demise of these more traditional semantic frameworks has significant consequences for issues of much interest to philosophers—issues concerning the nature of intentional states, meaning, and communication. But, for reasons that will be made clear in what follows, there is no way to demonstrate *directly* that there are assertions whose truth conditions cannot be accounted for by some semantic theory that falls within the traditional semantic framework; i.e., I cannot provide a counterexample. So here I take a somewhat indirect approach: I argue that the situation of the traditional semantic theories is analogous to the situation of logicism/reductionism and the problem of *multiple reductions* in the philosophy of mathematics. That is, I argue that *if* we agree with Benacerraf (1965) that numbers cannot be sets of various sorts—because there are too many sorts of sets that would do the trick and no principled means of choosing between them—then we should concede that there are semantic phenomena that cannot be adequately explained within the constraints of the traditional semantic framework. Moreover, though I will not support this claim here, one ought to accept Benacerraf's argument, and thus the traditional semantic framework should be rejected.

TRUTH CONDITIONAL COMPOSITIONALITY AND TRADITIONAL SEMANTIC THEORIES

What is at issue in the disagreement between the traditional semanticists and the truth conditional pragmaticists? The essence of the disagreement concerns the following general principle,

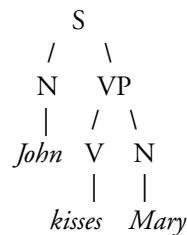
Truth Conditional Compositionality: The truth conditions of an utterance are a function of (i) the logical form of the utterance (i.e., the structure of the LF of the utterance), and (ii) the meanings of the words in the utterance (i.e., the *semantic values* of the terminal nodes of the LF of the utterance).

Traditional semantic theories presuppose Truth Conditional Compositionality. Indeed, Truth Conditional Compositionality is the central motivating idea of traditional semantic theories—the principle is a slightly more precise rendering of the slogan, often attributed to Frege, that “the meaning of a sentence must be a function of the meaning of the words in the sentence.”

Consider the sentence

(1) John kisses Mary.

How would a traditional semantic theory explain the meaning, or truth conditions, of an utterance of (1) in keeping with Truth Conditional Compositionality? A traditional semantic theory specifies a function that takes as inputs the LF of an utterance, and the semantic values of the words in the utterance, and has as its output the truth conditions of the utterance. Following most contemporary theorists working within the traditional semantic framework, I here assume the “Revised Extended Standard Theory” of syntax, and thus I assume that the entities interpreted by a semantic theory are phrase structure markers at the level of LF, or simply “LFs”. (The syntactic details do not matter for my purposes; all that matters is the assumption that some sort of syntactic representations of sentences play the role of LFs as specified by Truth Conditional Compositionality.) Suppose then that the LF of (1) is something like this:



The following is a very simple, and partial, traditional semantic theory that can account for the truth conditions of an utterance of (1) in keeping with Truth Conditional Compositionality:³

Lexical Rules

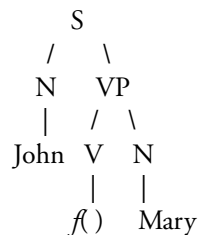
1. $SV(John) = John$
2. $SV(Mary) = Mary$.
3. $SV(kisses) = f: D \Rightarrow \{g: g \text{ is a function from } D \text{ to } \{true, false\}\}$
For all $x, y \in D$, $f(y)(x) = true$ iff x kisses y .

Combinatorial Rules

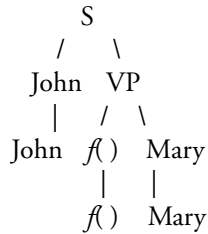
1. If α has the form $\begin{array}{c} VP \\ / \quad \backslash \\ \beta \quad \gamma \\ S \end{array}$, then $SV(\alpha) = \beta(\gamma)$.
2. If α has the form $\begin{array}{c} / \quad \backslash \\ \beta \quad \gamma \\ \alpha \end{array}$, then $SV(\alpha) = \gamma(\beta)$.
3. If α has the form, $\begin{array}{c} | \\ \beta \end{array}$ then $SV(\alpha) = SV(\beta)$.

This fragment of a simple traditional semantic theory determines the truth conditions of an utterance of (1), in keeping with Truth Conditional Compositionality in the following way. First, the lexical rules are applied to determine the semantic values of the lowermost nodes in the LF for (1): 'John' is assigned the semantic value John, 'Mary' is assigned the semantic value Mary, and 'kisses' is assigned a particular function from individuals to functions, these latter functions being functions from individuals to truth values.

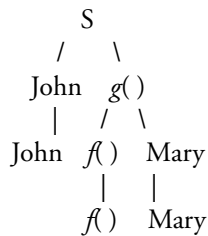
Once these semantic values are assigned to the lowermost nodes, the combinatorial rules, which are directed by the structure of the LF, can be applied to determine the semantic values of the non terminal nodes. That is, we can think of the combinatorial rules applying to an LF structure where the words (or morphological information) are replaced by the corresponding semantic values:



The Combinatorial Rules are now applied to determine the semantic values of the non terminal nodes. Where there is no “branching,” Combinatorial Rule 3 is applied. Thus the semantic values of the intermediate N and V nodes are assigned as follows:



Combinatorial Rule 1 is now applied to determine the semantic value of the VP node. Hence the function $f()$ from individuals to functions that is the semantic value of ‘kisses’ is applied to the argument Mary. The value of this function applied to this argument is another function, a function $g()$ from individuals to truth values which is such that $g(x) = \text{true}$ iff x kisses Mary. In this way the non-terminal node VP is assigned a semantic value in such a way that the semantic value assigned to it, viz., $g(x)$, is a function of the semantic values assigned to the nodes which VP immediately dominates.



Finally Combinatorial Rule 2 is applied to determine the semantic value of the top S node, or rather to determine the conditions under which this node is assigned *true* as its semantic value. More specifically, the semantic value of VP, function $g()$ from individuals to truth values, is applied to John, the semantic value of ‘John’. Hence the top S node is appropriately determined to be true iff John kisses Mary.

This simple example illustrates how traditional semantic theories respect Truth Conditional Compositionality. The semantic theory determines the truth conditions of an utterance, or more specifically the LF of an utterance, by first assigning semantic values to the terminal nodes in the LF via the Lexical Rules. The truth conditions of the nonterminal nodes, including the top S node (or IP node, or whatever) are then determined by the semantic values of the terminal

nodes and the structure of the LF via the combinatorial rules. In this way a traditional semantic theory, which is composed of such lexical and combinatorial rules, illustrates how the truth conditions of an utterance are a function of (i) the *semantic values* of the terminal nodes of the utterance's LF, and (ii) the structure of the utterance's LF.

It should be noted that utterances of context-sensitive sentences such as

(2) I am upset

do *not* constitute even *prima facie* counterexamples to Truth Conditional Compositionality. The truth conditions of utterances of (2) of course vary depending upon relevant aspects of the contexts of utterance: If Monica utters the sentence on Tuesday, November 23, 1998, at 6:30 P.M., *her* utterance is true if and only if *she* is upset at *that* time. And if Bill utters the sentence on Wednesday November 24, 1998, at 7:00 A.M., then *his* utterance is true if and only if *he* is upset at *this other* time. The truth conditions of sentences such as (2) can be adequately explained by traditional semantic theories which deviate only slightly from the sort of theory sketched above. This is because though (2) is context sensitive, its context sensitivity can be traced to the presence of two phonetically realized indexical words and morphological features, viz., 'I', and the tensed form 'am' of 'to be'. Thus, to explain the truth conditions of sentences such as (2) in keeping with Truth Conditional Compositionality one need only allow that *some* of the elements occupying the terminal nodes of LFs will not have a *constant* semantic value. The semantic values of such *indexical elements* will not be provided by fixed lexical rules like lexical rules 1–3 above, but will instead be determined by context-sensitive lexical entries (or *characters*, as Kaplan and his followers call them) together with relevant information provided by the particular context of utterance. For example, the *character* of 'I' is something along the lines of

SV('I') in context *c* = the agent of *c*.

Such context-sensitive lexical rules allow traditional semantic theories to adequately explain the truth conditions of context-sensitive expressions such as (2) in keeping with Truth Conditional Compositionality. The semantic values of the terminal nodes in the LF for (2) are allowed to vary from context to context, but once these semantic values are fixed the truth conditions of an utterance of (2) are *still* a function of (i) the semantic values of the terminal nodes of the sentence's LF, and (ii) the structure of the sentence's LF.⁴

PROBLEMS FOR TRADITIONAL SEMANTIC THEORIES: PRIMA FACIE COUNTEREXAMPLES TO TRUTH CONDITIONAL COMPOSITIONALITY

Though utterances of sentences such as (2) are not counterexamples to Truth Conditional Compositionality, consideration of such sentences does suggest a

strategy for finding counterexamples. Suppose an expression *S* (which may or may not be a complete sentence) can be used to make assertions, and further suppose that *S* has the following two properties:

- (a) *S* is context sensitive so that its truth conditions vary from context to context.
- (b) The LF of *S* contains no relevant context sensitive words and/or features.

Such an expression would constitute a counterexample to Truth Conditional Compositionality. This principle claims that the truth conditions of *every utterance of S* are a function of (i) the *semantic values* of the terminal nodes of *S*'s LF, and (ii) the structure of *S*'s LF. But it does not even make sense to think of the LF of an expression *S* changing from utterance to utterance; if *S* and *S** have different LFs, then, in the relevant sense of 'expression', *S* and *S** are different expressions. Consequently, if Truth Conditional Compositionality is to be preserved, no expression that can be used to make an assertion can possess both property (a) and property (b). Therefore, for any context-sensitive expression *S* that can be used to make an assertion, Truth Conditional Compositionality requires that the LF of *S* contain some context-sensitive element so that the variance in truth conditions across contexts can be explained by a variance in the semantic values assigned to this context sensitive element across contexts. The upshot is that a potentially assertion making sentence, or mere phrase, that had properties (a) and (b) would constitute a counterexample to Truth Conditional Compositionality, and thus if such sentences and/or mere phrases can be found, then this principle ought to be rejected.

As Bach (1994), Sperber and Wilson (1986), and other advocates of truth conditional pragmatics have demonstrated, there are many expressions that at least *seem* to have properties (a) and (b) and thus there are many expressions that constitute *prima facie* counterexamples to Semantic Compositionality. Here I present only four sorts of *prima facie* counterexamples.

Quantifier Domain Restriction

One sort of *prima facie* counterexample concerns quantifier domain restriction. Sentences such as

- (3) Every student came to Anna's party

clearly have property (a), and they seem to have property (b). A typical utterance of (3) does not state that every student in the universe came to Anna's party; rather a typical utterance of (3) states merely that every student relevant to the people engaged in the discourse came to the party. That is, the domain of quantification is not restricted to merely the set of *students* in the universe, but is fur-

ther restricted to a proper subset of relevant students. Moreover, the proper subset of students that serves as the restriction varies from context to context. For example, one utterance of (3) might be true if and only if every philosophy graduate student attending MIT in 1999 came to Anna's party. But a different utterance of (3) might be true if and only if every philosophy or linguistics graduate student attending MIT in 1998 came to Anna's party.⁵ So if Truth Conditional Compositionality is to be upheld, the LF of (3) must contain a context-sensitive feature that, relative to a context of utterance, serves to further constrain the domain of quantification. But there is no overt, phonetically realized, term or feature that could plausibly serve this purpose. So sentences such as (3) seem to possess properties (a) and (b), and thus there are at least *prima facie* counterexamples to Truth Conditional Compositionality.⁶

Comparative Adjectives

Another sort of *prima facie* counterexample concerns relative adjectives. Sentences such as

- (4) Bradley is tall

also seem to possess both (a) and (b). Again, (4) is clearly context sensitive. A typical utterance of (4) does not state that Bradley is tall *simpliciter* (whatever that might amount to), but rather that he is tall relative to some contextually salient contrast class. For example, an utterance of (4) that occurred in a discussion concerning the physical characteristics of presidential candidates would be true if and only if Bradley is tall for a presidential candidate, while an utterance of (4) that occurred in a discourse concerning great centers in the NBA would be true if and only if Bradley is tall for a great center in the NBA. Thus, the truth conditions of (4) depend upon what contrast class is invoked by the utterance. But, again, there is no overt, phonetically realized, word or feature that might have the relevant contrast class as its semantic value. So sentences such as (4) also seem to possess properties (a) and (b), and thus they also are at least *prima facie* counterexamples to Truth Conditional Compositionality.⁷

Propositional Attitude Ascriptions

Yet another, and much more widely appreciated, problem for Truth Conditional Compositionality is posed by attitude ascriptions. It is now widely appreciated that attitude ascriptions are context sensitive. Consider the ascription

- (5) Jerry believes that Marie baked the cookies.

In many contexts an occurrence of (5) is true only if Jerry thinks of Marie as the referent of 'Marie'; i.e., in many contexts an utterance of (5) is true only if Jerry

utilizes a woman-named-‘Marie’ “mode of presentation” of Marie. For example, suppose we are at a cookie baking contest, where there are plates of cookies set out with name tags, where the name tags state who baked that particular plate of cookies. We are observing Jerry who is tasting the cookies from a plate with a tag that reads ‘Ms. O’Connor’. Further suppose that we know that Jerry does not realize that the baker of the cookies, Ms. O’Connor, just *is* his acquaintance Marie, and that we are discussing Jerry’s plight. I thus say, “Ha! Poor Jerry does not know that ‘O’Connor’ is Marie’s last name, so he doesn’t know that those are Marie’s cookies!” If you were to utter (5) immediately following my statement, your utterance would be false. Moreover, it at least seems that it would be false because Jerry does not think of the baker of the cookies as the referent of ‘Marie’—he does not utilize the referent-of-‘Marie’ mode of presentation. But in other contexts (5) is more transparent, and in these contexts it is not the case that (5) is true only if Jerry utilizes a woman-named-‘Marie’ mode of presentation of Marie. Again suppose that we are at a cookie baking contest, and suppose that Jerry, whom we know to have no prior acquaintance with Ms. Marie O’Connor, observes her at a distance placing her cookies on a plate. Seeing Jerry observe Marie putting her cookies on a plate, I utter (5) to you. In this context my utterance of (5) is (probably) true, even though Jerry does not think of Marie under a referent-of-‘Marie’ mode of presentation. So (5) is clearly context sensitive in that some (“opaque” and/or *de dicto*) utterances of (5) are true only if Jerry thinks of Marie as the referent of ‘Marie’ (under a mode of presentation associated with the name ‘Marie’) while other (“transparent” and/or *de re*) utterances of (5) do not require that Jerry think of Marie in such a way (or under such a mode of presentation). So, if Truth Conditional Compositionality is to be preserved, there must be some indexical element in the LF of (5) that is assigned different semantic values in different contexts. But again, there are not any phonetically realized words or feature in (5) that are indexical in the relevant way. Consequently, attitude ascriptions such as (5) also constitute *prima facie* counterexamples to Truth Conditional Compositionality.⁸

Nonsentential Assertions

And finally, there are cases of what Stainton (1994, 1995) calls “nonsentential assertion.”⁹ Suppose that Ben and Melia are at a party, and across the room Melia observes a woman who is surrounded by a large group of people who are listening attentively to her. Melia turns to Ben and quizzically raises her eyebrows while nodding toward the woman. Ben then utters the phrase

(6) A world famous topologist.

In this context Ben’s utterance of a mere noun phrase has truth conditions and thus is an *assertion*: Ben’s utterance is true if and only if the observed woman is a world famous topologist. But clearly (6) is context sensitive, for in most con-

texts the utterance of a mere noun phrase does not constitute an assertion at all, and thus has no truth conditions whatsoever. Moreover, (6) does not have any phonetically realized indexical words or features that might have different people as semantic values in different contexts. Consequently nonsentential assertions such as (6) also constitute *prima facie* counterexamples to Truth Conditional Compositionality.

One might claim that the above example involving Ben's utterance of (6) does not constitute even a *prima facie* counterexample to Semantic Compositionality on the grounds that this example involves the phenomenon of *syntactic ellipsis*. Consider the following segment of discourse:

Jeremy: "Who believes in God?"
Anders: "James."

In this brief dialogue it is plausible to suppose that Anders' utterance of the noun phrase "James" is *syntactically elliptical* for the complete sentence, 'James believes in God.' That is, it is at least plausible to suppose that the verb phrase "believes in God" is somehow "copied" as a phonetically unrealized element into the LF of Anders' utterance, so that the LF of Anders' utterance is the LF of a complete sentence. Thus, the explanation runs, in some contexts syntactic ellipsis occurs, and an utterance of (6) has a complete sentence for its LF, while other utterances of (6) do not involve syntactic ellipsis, and in these utterances (6) does not have a complete sentence for its LF. So cases such as Ben's utterance of (6) do not constitute counterexamples to Truth Conditional Compositionality. This explanation, however, is inadequate. For, as Stainton points out, syntactic ellipsis requires a syntactic antecedent that can be copied into the LF of a later utterance. Note, however, that in the case at hand Ben's utterance of (6) occurs in *discourse initial position* and thus there is no syntactic antecedent that can be copied as an unarticulated constituent into the LF of (6). More specifically, there is no previously occurring noun phrase referring to the relevant woman that could be copied into the LF of Ben's utterance. Thus it seems that Ben's assertion of (6) cannot plausibly be construed as an instance of syntactic ellipsis.

***Two Strategies for Rescuing Truth Conditional Compositionality:
Pragmatic Ellipsis and Hidden Indexicals***

If the defender of traditional semantics is to defend Truth Conditional Compositionality, she must show that each of the above sentences and/or phrases, despite appearances, does not really have property (a), or does not really have property (b). That is, she must show that each of the *prima facie* counterexamples is either not really context sensitive, or really does contain a (relevant) context-sensitive lexical item. She can at least attempt to do this by arguing that the LFs of these sentences are richer than they seem to be. More specifically, the de-

fender of traditional semantics must argue that the LFs of these sentences and/or phrases contain phonetically unrealized elements that can explain, in keeping with Truth Conditional Compositionality, the truth conditions of the assertions.

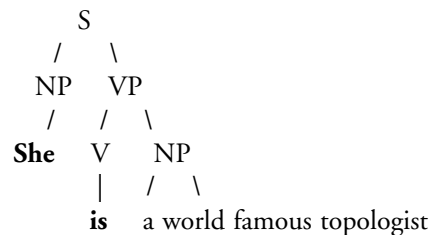
There are two principal ways in which this general strategy of response can be fleshed out. The first way, which I shall refer to as the “pragmatic ellipsis” substrategy, involves positing *familiar* lexical items, i.e., lexical items that are normally phonetically realized, as phonetically unrealized elements in LFs. These phonetically unrealized yet familiar lexical elements are then assigned semantic values by the semantic theory in the usual way. Thus, in proposition speak, the proposition expressed by an LF that contains such phonetically unrealized lexical items contains “unarticulated constituents,” i.e., semantic values that are not the semantic value of any phonetically realized word or feature. The second way, which I shall refer to as the “hidden indexical” substrategy, involves positing a new, *unfamiliar*, sort of phonetically unrealized indexical element in LFs and claiming that these “hidden indexicals” are assigned different semantic values in different contexts. The semantic values so assigned are again *unarticulated constituents*, as they are not the semantic values of phonetically realized words or features. The “hidden indexical” substrategy thus attempts to rescue Truth Conditional Compositionality from *prima facie* counterexamples by claiming that the assertion in question really does contain context sensitive elements.

Both substrategies have been utilized in attempts to rescue the Principle of Truth Conditional Compositionality from the *prima facie* counterexamples discussed above.¹⁰

Stanley (2000) utilizes the “pragmatic ellipsis” substrategy to explain away the *prima facie* counterexamples involving nonsentential assertions. Stanley proposes that the LF of Ben’s assertion making utterance of

(6) A world famous topologist

is



where both the familiar noun phrase ‘She’ and the familiar main verb ‘is’ are phonetically unrealized. Stanley claims that even cases of nonsentential assertion that occur in discourse initial position are a special case of ellipsis. Stanley explains,

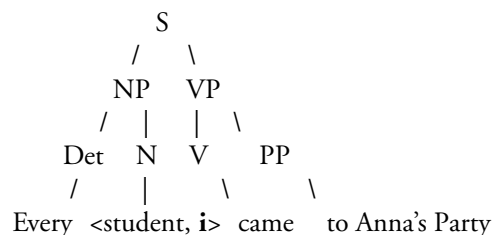
It is true that syntactically elliptical sentences cannot felicitously occur in the absence of a linguistic antecedent. But explicitly providing a linguistic antecedent by mentioning it is only the simplest way to provide it. There are other methods of raising linguistic expressions to salience in a conversation with explicitly using them. (2000, 21)

Thus, in the case of Ben's felicitous and assertion making utterance of (6), Stanley claims that the context of utterance somehow makes the lexical items 'She' and 'is' salient, and in virtue of this salience, the LF of the utterance contains these lexical items as phonetically unrealized elements. These phonetically unrealized elements are then assigned semantic values in the usual way, and thus the semantic values so assigned are unarticulated constituents. *If* the LF of an utterance of (6) is elliptical in this way and the requisite sort of semantic values are assigned as unarticulated constituents, then the truth conditions of utterances of (6) can be explained in keeping with Truth Conditional Compositionality. Stanley's pragmatic ellipsis proposal also accounts for the apparent context sensitivity of (6). In some contexts appropriate lexical items are made salient, by pragmatic processes such as the raising of eyebrows. In such contexts utterances of (6) are elliptical and thus correspond to full blown sentential LFs, and they thereby have truth conditions. But in other contexts appropriate lexical items are not made salient. In these deficient contexts the nonsentential utterances are not elliptical for full blown sentences. But, Stanley claims, such nonelliptical utterances lack illocutionary force and thus have no truth conditions.

Stanley and Gendler Szabo (2000) utilize the "hidden indexical" substrategy in an attempt to explain away the *prima facie* counterexamples concerning quantifier domain restriction. Consider again sentences such as

- (3) Every student came to Anna's party.

Stanley and Gendler Szabo propose that this sentence be analyzed as containing at the level of LF a "hidden indexical" that takes on different semantic values in different contexts. More specifically, they propose the LF of (3) is something like this

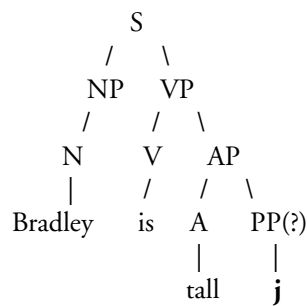


According to Stanley and Gendler Szabo's analysis, the terminal node corresponding to the phonetically realized noun 'student' is syntactically complex: It is an ordered pair, the first member of which is the phonetically realized familiar lexical item 'student' and the second member of which is a new sort of phonetically unrealized indexical element *i*.¹¹ This phonetically unrealized indexical element is assigned, relative to a context, a semantic value. Thus there is a semantic value invoked by an utterance of (3) that is not the semantic value of any phonetically realized, or articulated, word or feature in (3); the semantic value of *i* is then an "unarticulated constituent" of the proposition expressed by an utterance of (3). Moreover, since *i* is an indexical element, it is assigned different semantic values, i.e., different unarticulated constituents, in different contexts. In terms of the previous example involving different utterances of (3), in some contexts *i* is assigned as its semantic value the set of all philosophy graduate students attending MIT in 1999, while in other contexts *i* is assigned as its semantic value the set of all the linguistics and philosophy graduate students attending MIT in 1998. *If* there is such a "hidden indexical" in the LF of sentences such as (3), then such sentences do possess a context sensitive lexical element and thus they would not constitute counterexamples to Truth Conditional Compositionality. That is, positing such a "hidden indexical" element that has different "unarticulated constituents" as its semantic value in different contexts explains the context sensitivity of (3) in keeping with Truth Conditional Compositionality.

Ludlow (1989) has proposed using the hidden indexical substrategy to cope with the *prima facie* counterexamples posed by comparative adjectives. That is, Ludlow has proposed that the LF of

(4) Bradley is tall.

is something like



The phonetically unrealized indexical element *j* can be assigned different semantic values, different unarticulated constituents, in different contexts. In

items, and thus such sentences do not constitute counterexamples to Truth Conditional Compositionality.¹⁵

PRAGMATIC ELLIPSIS, HIDDEN INDEXICALS AND BENACERRAF'S ARGUMENT

I claim that the recently rehearsed attempts to explain away the *prima facie* counterexamples to Truth Conditional Compositionality are unsuccessful, and they are all unsuccessful for essentially the same reason that the reduction of the natural numbers to sets cannot succeed. Let us revisit Benacerraf's (1965) argument against such reductions. According to one proposed reduction of the natural numbers, the natural number sequence is *really* the following sequence of sets:

$$\{\emptyset\}, \{\emptyset\{\emptyset\}\}, \{\emptyset\{\emptyset\}\{\emptyset\{\emptyset\}\}\}, \{\emptyset\{\emptyset\}\{\emptyset\{\emptyset\}\}\{\emptyset\{\emptyset\}\{\emptyset\{\emptyset\}\}\}\}, \dots$$

But according to another proposed reduction, the sequence of natural numbers is really the following distinct sequence of sets:

$$\{\emptyset\}, \{\{\emptyset\}\}, \{\{\{\emptyset\}\}\}, \{\{\{\{\emptyset\}\}\}\}, \dots$$

Benacerraf argues that neither sequence of sets can be identified with the sequence of natural numbers. His argument proceeds from two key premises. First, the two proposed reductions are incompatible. For example, according to the first proposed reduction one is a member of three, but according to the second proposed reduction one is not a member of three. Since one cannot both be and not be a member of three, *both* sequences of sets cannot be the sequence of natural numbers. And second, there is nothing that could recommend one proposed reduction over the other; there is no possible evidence that would make it rational to prefer one proposal over the other. Benacerraf concludes, rightly, that neither proposed reduction is correct and thus numbers are not sets:

If numbers are sets, then they must be particular sets, for each set is some particular set. But if the number 3 is really one set rather than another, it must be possible to give some cogent reason for thinking so; for the position that this is an unknowable truth is hardly tenable. But there seems to be little to choose among the accounts. Relative to our purposes in giving an account of these matters, one will do as well as another, stylistic preferences aside. There is no way connected to the reference of number words that will allow us to choose among them, *for the accounts differ at places where there is no connection whatever between features of the accounts and our uses of the words in question*. If all the above is cogent,

then there is little to conclude except that any feature of an account that identifies 3 with a set is a superfluous one—and that therefore 3, and its fellow numbers, could not be sets at all. (1965, 62)

The same considerations that led Benacerraf to reject the reduction of natural numbers to sets compel us to reject the above attempts to rescue Truth Conditional Compositionality. In what follows I will demonstrate how Benacerraf's argument refutes Stanley's pragmatic ellipsis analysis of nonsentential assertions and Ludlow's hidden indexical analysis of comparative adjectives. I intend my examination of these cases to illustrate that any analysis utilizing either the pragmatic ellipsis substrategy or the hidden indexical substrategy will fall to Benacerraf's argument.

Consider again Stanley's proposed pragmatic ellipsis analysis of nonsentential assertions. According to the pragmatic ellipsis analysis, Ben's assertion making utterance of

(6) A world famous topologist

is elliptical for

(6*) **She is** a world famous topologist

even though there is no appropriate linguistic antecedent such as "Who is she?" present in the context. The problem is that there are equally plausible candidates for what the elided material could be other than 'She is'. Here are three plausible alternatives:

- i. *That woman*
- ii. *That person*
- iii. *The loud mathematician*

First, these candidate-unrealized lexical items are incompatible; only one of them can appear as the phonetically unrealized noun phrase in the LF for a felicitous and assertion-making utterance of (6). But, second, there is no possible evidence that would recommend one candidate over the other. In many contexts the speaker will have no discernible intentions discriminating enough to recommend one candidate over the other, and neither will there be salient features of the context that recommend one over the other. But these are the only kinds of admissible evidence; any other facts to which one might appeal to support one candidate over the others is, to use Benacerraf's term, *superfluous*. So Benacerraf's reasoning compels us to deny that any phrase such as (i), (ii), or (iii) occurs phonetically unrealized in the LF for an assertion making utterance

of (6).¹⁶ Therefore, Stanley's proposed analysis fails, and we must conclude that nonsentential assertions such as (6) constitute counterexamples to Truth Conditional Compositionality.

A defender of Truth Conditional Compositionality might respond by suggesting that the threatened indeterminacy can be resolved by appeal to *simplicity*.¹⁷ Perhaps the candidate for the elided material that is to be selected is the simplest possible candidate. An obvious problem for this proposal is that it is not all clear what the notion of simplicity amounts to here. One possible dimension along which simplicity might be judged concerns the semantic content of the candidates for the elided material. Judged along this dimension, one might suppose that candidate (ii) is to be preferred over (i), (iii), and even Stanley's proposed 'she is', as its semantic content is intuitively simpler: All women and loud mathematicians are persons, but not all persons are women, nor are all persons loud mathematicians. But if this sort of simplicity is invoked, then the even less informative 'that is' is to be preferred over candidate (ii), because every person can be referred to using 'that', but not everything that can be referred to using 'that' is a person.

There are, however, serious problems with this appeal to simplicity of semantic content. First, the proposal is at odds with the generally accepted principle of communication that requires that speakers be interpreted as being maximally informative (e.g., Grice's maxim of quantity). But more importantly, there is no reason to believe that the proposal will resolve the indeterminacy. For example, 'it is' and 'that is' seem to have equally simple semantic content. One might attempt to bolster the simplicity of semantic content by invoking simplicity along another dimension; perhaps the simplest candidate is also to be judged along dimensions of *syntactic* and/or *lexical* simplicity. But then which to take precedence, simplicity of semantic content, or simplicity of syntactic structure? Stanley's 'She is' is simpler syntactically than 'That person is', but the latter seems to have simpler semantic content. Moreover, regardless of which dimension is to take precedence, there is still no reason to think that the indeterminacy can be resolved, as 'it is' and 'that is' seem to be equally simple along both dimensions. It is apparent that no intuitive and straightforward notion of simplicity will determine a unique candidate for the elided material.¹⁸

In addition, it should be noted that candidate (iii) would not suffice as a complete specification of the elided material, as it contains a quantifier phrase (an "incomplete definite description") and a comparative adjective. So, for example, Ben's utterance of (6) cannot be elliptical merely for the sentence 'The loud mathematician is a world famous topologist', for, like (6), this expression also seems to have properties (a) and (b). And thus it also constitutes a *prima facie* counterexample to Truth Conditional Compositionality. A *complete* specification of the elided material would have to specify some lexical element whose semantic value was the relevant contrast class for the comparative adjective

'loud'. And it would also have to specify some lexical items whose semantic values were the relevant quantifier domain restrictions for the quantifier phrase 'the loud mathematician'. If one were to apply the pragmatic ellipsis strategy "all the way down," one would have to avoid positing incomplete definite descriptions and comparative adjectives as phonetically unrealized elements, for these elements would themselves be in need of further analysis; one would have to "bottom out" with an analysis that posited no such problematic elements. This is reminiscent of a familiar problem with Russell's descriptive analysis of referring terms: For example, one cannot maintain that 'Plato' is *really* an abbreviation for merely 'the teacher of Aristotle', for 'Aristotle' is itself a referring term, and thus in need of further analysis. Hence Russell's doomed search for "logically proper names," i.e., expressions in no need of further analysis. The advocate of the pragmatic ellipsis substrategy is committed to a very similar, and equally implausible, search.

The defender of Truth Conditional Compositionality might object that I have failed to establish the second key premise on the grounds that discernible intentions of the speaker and salient features of the context do *not* exhaust the admissible evidence. After all, if LFs have some sort of psychological reality, and are somehow represented in people's brains, then all sorts of psychological and/or neurological facts about Ben could be brought to bear on the question of what the LF of Ben's utterance *really* was. And of course Ben need not have explicit knowledge of, or be conscious of, any of these facts.

This objection, however, confuses psychology and semantics. If LFs are instantiated in people's brains somehow, then all sorts of psychological and neurological evidence is relevant to determining what LF is instantiated in Ben's brain. But most of these psychological facts are irrelevant to the *semantics* of Ben's utterance, because semantics is concerned with communication and interpretation. Stanley and Gendler Szabo (2000, 11) assume that the task of semantics is to explain the interpretation of "typical assertions," and they maintain that such interpretation "is successful just in case the hearer can identify the proposition the speaker intends to communicate."¹⁹ The model of interpretation assumed by Stanley and Gendler Szabo is paradigmatic of the model presupposed by traditional semantic theories generally. According to this model, interpretation is a two-step process whereby a hearer identifies the proposition the speaker intends to communicate, or equivalently determines the truth conditions of an assertion. In the first step the hearer uses her syntactic and phonological knowledge, together with whatever clues she can garner from the context of utterance, to determine the LF of the assertion. Stanley and Gendler Szabo (2000, 13) use the equation, "what is articulated + context = what is uttered" to describe this first step, where "what is articulated" is a "phonological sentence," and "what is uttered" is a "grammatical sentence," i.e., an LF. In the second step the hearer uses her semantic knowledge, together with whatever clues she can garner from

context, to determine the proposition expressed, or equivalently the truth conditions of the utterance. Stanley and Gendler Szabo (2000, 15) use the equation “what is uttered + linguistic meaning + context = what is said” to describe the second step, where “what is said” is the proposition expressed, or the truth conditions of the utterance.²⁰

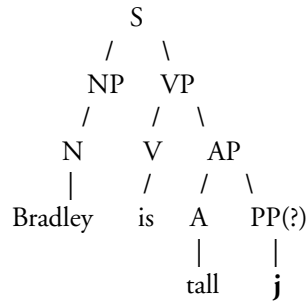
The general problem with the pragmatic ellipsis substrategy is that it renders the first step in the process of interpretation impossible. If the truth of the equation “what is articulated + context = what is uttered” is to be preserved, the hearer must be able to determine what is uttered, i.e., the LF, from discernible features of the context of utterance. So even if facts about the speaker’s psychological state determine what LF is instantiated in his brain, these facts are not discernible by the hearer. Therefore such indiscernible psychological facts are irrelevant to interpretation, and so are irrelevant to the semantics of his utterance.²¹ Again in terms of Benacerraf’s argument, such indiscernible psychological facts are *superfluous*. (Consequently, *if* the level of syntactic representation known as LF is not in the relevant sense discernible in a context of utterance, then LF is irrelevant to semantics, where *semantics* is concerned with *interpretation* and *communication*.)

The defender of Truth Conditional Compositionality might accept the above objections against the pragmatic ellipsis substrategy yet still endorse the hidden indexical substrategy, for the hidden indexical substrategy seems well suited to avoid the problems encountered by the pragmatic ellipsis substrategy. In particular, because the hidden indexical substrategy eschews positing *familiar* phonetically unrealized elements in LFs and instead posits specially designed *unfamiliar* (i.e., never phonetically realized) context sensitive items, there is not an overabundance of candidate LFs. Consider again an utterance of

(4) Bradley is tall.

If one were to attempt to explain away this *prima facie* counterexample utilizing the pragmatic ellipsis substrategy, one would be faced with an overabundance of suitable proposals for the elided material. That is, an utterance of (4) might be elliptical for ‘Bradley is tall for a middle aged American male in 2000’ or it might be elliptical for ‘Bradley is tall for a presidential candidate’. Thus there are many equally plausible proposals as to what the LF of the utterance is, and no reason to prefer one proposal over the others. According to Ludlow’s hidden indexical analysis, however, there is only one plausible proposal as to what the LF of the utterance is, for one can maintain that it is built into the grammar that in sentences such as (4) comparative adjectives occur only with the appropriate sort of hidden indexical. So on Ludlow’s hidden indexical

analysis there is only one plausible candidate for the LF of (4), and it is something like



Hence it appears that the hidden indexical substrategy is much more plausible than the pragmatic ellipsis substrategy.

This appearance, however, is illusory, for the hidden indexical substrategy succeeds only in relocating the fundamental problem. According to the model of interpretation presupposed by traditional semantic theories, successful interpretation requires that the hearer identify the proposition expressed by an utterance, and this identification proceeds by way of the two-step process described above. The problem with the pragmatic ellipsis substrategy is that it renders this first step impossible, for it allows for an overabundance of proposals concerning what the LF for an utterance is, and no means of choosing between them. But notice that on the pragmatic ellipsis substrategy, *if* a hearer somehow managed to succeed in taking the first step and thereby identified the correct LF, then there is nothing especially problematic precluding her from successfully completing the second step of interpretation. This because the pragmatic ellipsis strategy posits only *familiar* elements as phonetically unrealized elements, and consequently no special problem is posed regarding the assignment of semantic values to such phonetically unrealized elements. Hence the pragmatic ellipsis strategy poses no special problem for the second step of interpretation, but does so only because it renders the first step impossible.

The hidden indexical substrategy faces the complimentary problem. The hidden indexical analysis poses no special problem for the first step of interpretation, for according to it a *hearer's linguistic knowledge alone* would enable her to determine the LF of an utterance of (4); this because it is simply built into the grammar that in sentences such as (4) comparative adjectives occur only with the appropriate sort of unfamiliar hidden indexical. The problem for the hidden indexical substrategy arises for the second step of interpretation, *viz.*, going from the LF of the utterance to identifying the proposition expressed by the utterance, or equivalently determining the truth conditions of the utterance. The

problem now is that there are too many equally plausible proposals as to what the semantic value of the hidden indexical is, and no admissible evidence that would justify a hearer in choosing one proposal over the others.

Consider Ludlow's hidden indexical analysis of an utterance of (4). According to this analysis the LF for an utterance of (4) contains a hidden indexical, *j*, that has as its semantic value the appropriate contrast class. The problem is that there is an overabundance of plausible proposals as to which contrast class is the semantic value of the hidden indexical. Here are two, among many, plausible contrast classes which might serve as the semantic value of *j* relative to a particular utterance of (4):

- (i) {x: x is a current presidential candidate}
- (ii) {x: x is a current or past presidential candidate}

Again, these candidates are *incompatible*. The sets in question are not identical, and therefore only one of them can be assigned as *the* semantic value of *j* relative to a particular utterance of (4). Moreover, there is no possible evidence that would make it rational to prefer one candidate over the other. In most circumstances a speaker who utters a perfectly felicitous utterance of (4) has no discernible intentions that would determine which, if either, of (i) or (ii) was the contrast class he *really* "tacitly referred to."²² And it can simply be stipulated that there are no discernible features of the context that recommend one proposal over the other. Moreover, appeals to simplicity again cannot be invoked to decide the matter because, first, it is not at all clear what it is for one set to be simpler than another, and second, there is no reason to believe that such a notion of simplicity, even if it could be made precise, would determine a unique set. And finally, for reasons given above, indiscernible facts concerning the speaker's psychological state cannot be appealed to as evidence to support one candidate over the others; such facts are again superfluous. So we are in the same position with regard to the question of which set is the semantic value of the posited hidden indexical *j* as we are with regard to the question of which sequence of sets is the natural numbers. Hence we ought to conclude that no set is the semantic value of such a hidden indexical; there is no such unarticulated constituent. And consequently Ludlow's proposed hidden indexical analysis cannot explain context sensitivity of sentences such as (4) in keeping with Truth Conditional Compositionality. Sentences such as (4) involving comparative adjectives constitute counterexamples to this principle after all.²³

The defender of Truth Conditional Compositionality might respond by pointing out that even familiar phonetically realized elements suffer from an indeterminacy of semantic value. Consider a typical utterance of 'Now it is time to go.' Precisely what span of time is to be assigned as the semantic value of the occurrence of 'now'? Is it a two-second span, a five-minute span, a ten-minute

span, or what exactly? Again, it is quite likely that the speaker has no intentions, discernible or otherwise, that would discriminate between these candidate semantic values. Or consider again a typical utterance of

- (3) Every student came to Anna's party.

Precisely what set (or intension) is to be assigned as the semantic value of the occurrence of 'student'? Is it the set of all full-time students, or does it include people who take an occasional night course? If the former, precisely what constitutes being a "full-time" student? Once again, it is likely that the speaker has no discernible intentions, nor is there anything in the context, that would provide answers to these questions, and thus there is nothing that would discriminate between a number of candidate semantic values. So, the response concludes, there is nothing especially problematic concerning the assignment of semantic values to hidden indexicals; such indeterminacy is no more problematic for the posited unfamiliar hidden indexicals than it is for familiar phonetically realized elements. And consequently such indeterminacy ought not prevent us from positing hidden indexicals to rescue Truth Conditional Compositionality.²⁴

This response gives rise to a number of perplexing issues, but I think it is relatively clear that it fails to justify positing hidden indexicals. The responder is correct to acknowledge that the indeterminacy of semantic value is commonplace. The meanings of words and speaker's intentions, discernible or otherwise, cannot decide for every possible case whether or not the word applies to that case. (This, I think, is a major theme in Wittgenstein.) But it seems to me that the defender of Truth Conditional Compositionality cannot be so sanguine about the indeterminacy of semantic values. For far from supporting the positing of hidden indexicals, acknowledgement of how commonplace this sort of indeterminacy is seems to undermine the need for positing such hidden indexicals in the first place. If there is no precisely defined set (or intension) that is the semantic value of, e.g., 'student', then what need is there for machinery that would further constrain the (indeterminate) domain of students? Why posit machinery to fine tune that which is indeterminate? Consider again the case quantifier domain restriction involving a typical utterance of (3). One might respond to such an utterance in a number of ways. One might *accommodate* the utterance, that is, accept it as true and move on: 'Yeah, what a blast! The faculty left early, but no student left before two!' Or one might refuse to accommodate by rejecting the assertion: 'No, not *every* student was there; for example several students with nonresident status were not there'. Or, one might refuse to accommodate by requesting clarification: 'By "student" do you mean to include students with nonresident status?' If one accepts the indeterminacy of semantic value, then one can and should view this phenomenon of accommodation, or

refusal to accommodate, as discourse participants to some extent deciding and/or stipulating as they go how expressions are to be used and interpreted in their discourse. There is no reason to posit a precise and fixed proposition that is exactly what the utterance of (3) expressed when it was uttered; i.e., there is no reason to suppose that there is a precise content that can be precisely stated in other terms. But if there is no reason to posit such a precise and fixed proposition, then there is no reason to posit machinery that would explain how such a precise and fixed proposition could, in keeping with Truth Functional Compositionality, be expressed by such an utterance. So, far from justifying hidden indexicals, the appeal to the widespread phenomenon of semantic value indeterminacy undermines the need to posit such entities. These brief remarks do not resolve these perplexing issues, but they do demonstrate that the defender of Truth Conditional Compositionality cannot unproblematically appeal to the widespread phenomenon of semantic value indeterminacy in defense of the hidden indexical substrategy.

CONCLUDING REMARKS

If we accept Benacerraf's argument against reducing numbers to sets, then we ought to reject both of the substrategies for rendering the many *prima facie* counterexamples compatible with Truth Conditional Compositionality. Moreover, we ought to accept Benacerraf's argument, and there are no other plausible strategies for rescuing the principle that will not fall to Benacerraf's argument. Therefore Truth Conditional Compositionality ought to be rejected. Rejection of the principle has obvious consequences for semantic theory. If the principle is rejected, then the traditional theoretical frameworks of Davidson and Montague must be rejected. Moreover, the traditional Gricean distinction between semantics and pragmatics breaks down. That is, if the domain of semantics is truth conditions and "what is said," then semantics cannot be concerned only with LFs and the semantic values of terminal nodes of LFs; rather the domain of semantics must be expanded to include features of utterances and discourses that were relegated to pragmatics under Grice's way of drawing the distinction. It does *not* follow that Frege's fundamental insight that the meaning of an utterance is a *function* of the logical structure of the sentence uttered together with antecedently given information; Frege's fundamental insight that determining meaning is a matter of computing functions need not be rejected. What does follow is that the truth conditions of an utterance cannot be a function of *only* the logical structure of *the utterance itself*, and the meanings of the words and semantically relevant features of *the utterance itself*. It is still eminently plausible that the meaning of an utterance is a function of logical structure, the meanings of the words and semantically relevant features of the utterance, and *other relevant information provided by the context of utterance*. Such other information might include more global discourse properties such as *topic* and *focus*, as well as the previous utter-

ances of the discourse and their structure. Many of the projects and proposals of truth conditional pragmatics, including the project of Discourse Representation Theory, attempt to formally model and thereby explain how such *extrasentential* information contributes to the meaning of an utterance.

NOTES

This chapter benefited as a result of comments and criticism from Andrew Botterell, Jason Stanley, Robert Stainton, Michael Glanzberg, Mark Richard, Michael O'Rourke, and those who attended the session at the 2000 INPC where an earlier version was read.

1. Heim and Kratzer 1998 and Larson and Segal 1995 are excellent texts introducing semantic theory within the traditional framework.
2. Kamp and Reyle (1993), Recanati (1996), Bach (1994), Carston (1991), Sperber and Wilson (1986), and Travis (1885) are good examples of theorists developing dynamic semantics and/or truth conditional pragmatics.
3. I will here give a simple extensional semantics, but my remarks apply *mutatis mutandis* to intensional, Montague inspired, semantic theories as well. The simple theory fragment is inspired by the semantic theory developed in Heim and Kratzer 1998.
4. Similar remarks apply to sentences containing demonstratives such as 'this' and 'that', even though such demonstratives do not seem to have anything like a *character*.
5. I here ignore other obvious ways in which (3) is context sensitive: tense, the referents of 'Anna' and 'Anna's party'.
6. The problem posed by quantifier domain restrictions was, I believe, first invoked by Strawson (1950) as an objection to Russell's quantificational analysis of definite descriptions.
7. The problem posed by comparative adjectives is discussed by Bach (1994).
8. Bach 1997 contains a cogent discussion of the problem attitude ascriptions pose for the traditional semantic framework.
9. Stainton's work on nonsentential assertion builds upon Barton 1990. Moreover, an anonymous referee informed me that some of Barton's criticisms of elliptical analyses of nonsentential assertions are similar to my objection against Stanley's pragmatic ellipsis strategy. See Barton 1990, chapter 2.
10. These two substrategies do not exhaust the possible ways of rescuing Truth Conditional Compositionality. Another strategy would be to claim that, despite appearances, one of the phonetically realized elements of a *prima facie* counterexample is actually context sensitive. Richard (1990) applies this strategy to *prima facie* counterexamples involving attitude ascriptions; Richard claims that, despite appearances, propositional attitude verbs are really context sensitive. And Heim and Kratzer (1998) suggest applying this strategy to *prima facie* counterexamples involving comparative adjectives. Heim and Kratzer (1998, 71) suggest that the lexical rule for 'small' could be

$$SV(\textit{small}) = \lambda x \in D_c.$$

x's size is below *c*, where *c* is the size standard made salient by the utterance context.

The effect of this lexical entry is to make the semantic value of 'small' vary from context to context. I do not consider these other substrategies here for they run afoul of the same sorts of difficulties as does the "hidden indexical" substrategy.

11. Actually, Stanley and Gendler Szabo's proposal is more complicated, and much less plausible, than this. For reasons that go beyond the scope of this paper, they suggest that the second member of the ordered pair consists of a combination of an indexical *f*() that has as its semantic value, relative to a context, a function from individuals to sets (or properties), and an indexical *i* that has as its semantic value, relative to a context, an individual. The set that is determined by applying the function "provided by context" (!) to the individual provided by context serves to further restrict the domain of quantification. Not surprisingly, Stanley and Gendler Szabo are wholly silent as to how such functions might be provided by context.

12. Ludlow defends a more sophisticated analysis for cases in which the subject NP is headed by a noun whose semantic value would serve as an appropriate contrast class. For example, on Ludlow's analysis, 'That man is tall' is true iff that *man* is tall *for a man*. Generally, on Ludlow's analysis a sentence of the form 'That N is A_C', where N is a noun and A_C is a comparative adjective, is true iff that N is A_C for an N̄. (Please forgive the use/mention sloppiness here.) This analysis cannot be correct, however, as it clearly fails in many, if not most, contexts in which comparative adjectives are used. Suppose I am trying to break the ice on my pond, and I say "I need something hard to break the ice." John, picking up a large rock, replies, "This rock is hard." Clearly John has not said that this rock is hard for a rock. In fact, Ludlow's analysis seems to hold only in very minimal contexts in which it is not all clear why one might be uttering a sentence of the form 'that N is A_C'.
13. I say "one way" because Crimmins (1992) does not commit himself to the existence of a phonetically unrealized element at LF. In my view his analysis is wholly unmotivated if he does not make this commitment. If the purpose of the "unarticulated constituent" analysis is not to preserve (at least something like) the Truth Conditional Compositionality, then what is the purpose? Why posit "modes of presentation" ("notions," "ideas," whatever) as semantic values at all?
14. The phonetically unrealized elements **n₂** and **n₃** are needed to explain opacity allegedly arising as a result of there being various modes of presentation (ideas or notions) of baking and the cookies, respectively.
15. An anonymous referee suggested to me that the context sensitivity of (5) could be accounted for by a sort of structural ambiguity. That is, one could maintain that some ("de re") utterances of (5) have an LF that more closely resembles
- (5*) Jerry believes of Marie that she baked the cookies
- while other ("de dicto") utterances of (5) have an LF that more closely resembles the phonetic form of (5). Thus the context sensitivity of (5) is likened to the structural ambiguity of, e.g., 'Flying planes can be dangerous.' Moreover, this "structural ambiguity" proposal has little in common with either of the two substrategies considered in this paper, and consequently it will not fall to the objections presented against these substrategies. So, with regard to the problem posed by attitude ascriptions there is an alternative strategy of response available to the defender of traditional semantics.
16. This point was made by Howard Wettstein (1981) with regard to an elliptical analysis of "incomplete" definite descriptions, a special case of quantifier domain restriction. Definite descriptions such as 'the murderer' are incomplete, because there is, unfortunately, not a unique murderer in the universe. But Wettstein points out that concerning such incomplete definite descriptions
- there will be any number of ways to fill out the description so as to yield a [complete] Russellian description (e.g., 'Harry Smith's murderer', 'the murderer of Joan Smith's husband', 'the murderer of the junior senator from New Jersey in 1975') and in many cases there will be nothing about the circumstances of utterance or the intentions of the speaker which would indicate that any of these [complete] Russellian descriptions is the correct one. (1981, 250–51)
17. This response was suggested by an anonymous referee.
18. An anonymous referee suggested that evidence in support of Stanley's proposal is provided by the fact that after Ben's utterance of (6), Melia might felicitously utter
- (6a) No, She's not. She's a world famous topologically challenged linguist.
- The suggestion is that the occurrences of 'she' in (6a) must be anaphoric on a previous phonetically unrealized occurrence of 'she'. In my view the alleged phenomenon of intersentential anaphora involving phonetically unrealized items is mysterious enough that little weight should be placed on such evidence. But if such evidence is to be considered, the fact that (6a) would be a felicitous utterance all by itself provides no evidence in support of Stanley's proposal. For note that all of the following would also be felicitous utterances:

- (6b) No, *he's* not. *He's* a world famous topologically challenged linguist.
 (6c) No *it's* not. *It's* a world famous topologically challenged linguist.
 (6c) *That woman* is *not* a world famous topologist! She's a poser!
 (6d) *That person* is not a world famous topologist! *He's* a poser!
 (6e) No way. Just another mathematician who talks too loudly.

Moreover, note that the following would be *infelicitous* following Ben's utterance of (6), though it seems that they would be felicitous if 'She is' were present, but phonetically unrealized, in Ben's utterance:

- (6f) Actually, it's a "he," not a "she."
 (6g) Are you sure that's a "she"?
 (6h) Oh! I thought that was a man!

19. Stanley and Gendler Szabo (2000, 11) define *typical assertions* as follows: "In typical assertions (i) there is a single speaker and a single hearer, (ii) the speaker vocalizes a well-formed, meaningful sentence, and by doing so (iii) the speaker intends to convey a certain proposition."
20. Though Stanley and Gendler Szabo clearly endorse the two-step model of interpretation, they are quick to point out that in practice "interpretation may not be a linear progression from the sentence articulated [i.e., the phonological sentence] through the sentence uttered [i.e., the LF] to the proposition expressed to the proposition communicated" (2000, 17). That is, actual interpretation may involve a complex process of going back and forth over the two steps.
21. Stanley and Gendler Szabo (2000) distinguish between "the foundational problem of context dependence" and the "descriptive problem of context dependence." The descriptive problem of context dependence is solved by determining *what* role context plays in determining the truth conditions of utterances. Hence, solving the descriptive problem is a matter of teasing out the roles played by syntax, linguistic meaning, and context. The foundational problem, on the other hand, concerns *how* context manages to play the role so described. Hence another way of putting my objection to the pragmatic ellipsis substrategy is that it renders the foundational problem unsolvable.
22. Schiffer (1992) raises essentially this problem, which he calls the "meaning-intention problem," for a hidden indexical analysis of attitude ascriptions.
23. One might attempt to avoid these difficulties with the hidden indexical analysis by invoking some sort of *superevaluation* procedure to determine the semantic value of *j*. One might suggest, for example, that the actual semantic value of *j* is the intersection of all the sets that are plausible candidates to serve as the relevant contrast class. Such superevaluation procedures, however, only relocate the problem. For such procedures presuppose a determinate class of sets over which the superevaluation procedure is to be performed. The problem now is that there are many equally plausible, yet incompatible, candidates to serve as this class, and no admissible evidence that would support one over the others.
24. This response, or something like it, has been offered by both Mark Richard and Herman Cappelen.

REFERENCES

- Bach, K. 1994. "Conversational Implicature." *Mind & Language* 9: 124–62.
 ———. 1997. "Do Belief Reports Report Beliefs?" *Pacific Philosophical Quarterly* 78: 215–41.
 Benacerraf, P. 1965. "What Numbers Could Not Be." *Philosophical Review* 74: 47–73.
 Barton, E. 1990. *Nonsentential Constituents*. Philadelphia: John Benjamins.
 Carston, R. 1991. "Implicature, Expliciture, and Truth-Theoretic Semantics." In Stephen Davis (ed.), *Pragmatics*. Oxford: Oxford University Press.
 Crimmins, M. 1992. *Talk About Beliefs*. Cambridge, Mass.: MIT Press.
 Heim, I., and A. Kratzer. 1998. *Semantics in Generative Grammar*. Oxford: Blackwell Publishers.
 Kamp, H., and U. Reyle. 1993. *From Discourse to Logic*. Dordrecht: Kluwer Academic Publishers.
 Larson, R. and G. Segal. 1995. *Knowledge of Meaning*. Cambridge, Mass.: MIT Press.

- Ludlow, P. 1989. "Implicit Comparison Classes." *Linguistics and Philosophy* 12: 519–33.
- Neale, S. 1990. *Descriptions*. Cambridge, Mass.: MIT Press.
- Recanati, F. 1996. "Domains of Discourse." *Linguistics and Philosophy* 19: 445–75.
- Richard, M. 1990. *Propositional Attitudes: An Essay on Thoughts and How We Ascribe Them*. Cambridge: Cambridge University Press.
- Schiffer, S. 1992. "Belief Ascription." *Journal of Philosophy* 89: 499–521.
- Sellars, W. 1954. "Presupposing." *Philosophical Review* 63: 197–215.
- Sperber, D., and D. Wilson. 1986. *Relevance*. Cambridge, Mass.: Harvard University Press.
- Stainton, R. 1995. "Non-Sentential Assertions and Semantic Ellipsis." *Linguistics and Philosophy* 18, 281–96.
- . 1994. "Using Non-Sentences: An Application of Relevance Theory." *Pragmatics and Cognition* 2: 269–84.
- Stanley, J. 2000. "Context and Logical Form." *Linguistics and Philosophy*, 23: 391–434.
- Stanley, J., and Z. Gendler Szabo. 2000. "Quantifier Domain Restriction." *Mind and Language* 15, no. 283.
- Strawson, P. 1950. "On Referring." *Mind* 59: 320–44.
- Travis, C. 1985. "On What is Strictly Speaking True." *Canadian Journal of Philosophy* 15: 187–229.
- Wettstein, H. 1981. "Demonstrative Reference and Definite Descriptions." *Philosophical Studies* 40: 241–57.

Generalized Conversational Implicatures and Default Pragmatic Inferences

Anne Bezuidenhout
University of South Carolina

INTRODUCTION

GRICE DISTINGUISHED BETWEEN generalized and particularized conversational implicatures. The latter are “cases in which an implicature is carried by saying that p on a particular occasion in virtue of special features of the context”. The former are cases in which the “use of a certain form of words . . . would normally (in the absence of special circumstances) carry such-and-such an implicature or type of implicature.” (Grice 1989, 37). Grice himself did not develop this distinction to any great extent. He gave a few examples meant to illustrate the distinction he had in mind. He never indicated that he thought generalized conversational implicatures occupied a separate level, between the level occupied by conventionalized meaning on the one hand and the level occupied by the one-off speaker meanings that correspond to the particularized variety of conversational implicatures on the other. However, some neo-Griceans, especially Levinson (1987b, 1995, 2000), have recently been developing a theory of generalized conversational implicatures (GCIs). Levinson proposes to treat GCIs as (the output of?) “default pragmatic inferences which may be cancelled by specific assumptions, but otherwise go through.” (Levinson 1987a, 723). He has been developing a set of heuristics or default inference rules that he says are used to generate GCIs. These default inferences yield interpretations that represent a level of meaning that he calls utterance-type meaning, which is intermediate between sentence-type meaning and speaker meaning.

In this chapter I argue against the idea of a set of default inference rules that are attached to certain classes of expressions. An account of utterance interpretation that appeals to cognitive strategies that are independent of particular ut-

terance types is to be preferred to an account that appeals to utterance-specific heuristics. I also suggest that the mechanisms of language understanding used in the recovery of the interpretations in question are more flexible and more context-dependent than allowed for by Levinson.

The justification for focusing on Levinson's views is that his theory of GCIs and his defense of their default status have been worked out in detail, especially in his recent book on the topic (Levinson 2000). Thus his is a good account against which to define an alternative view. There is a lot of material in Levinson's book that cannot be discussed here. Also, I am in agreement with Levinson about many of the theses he defends, especially his claim that there are pragmatic intrusions into the truth-conditional contents of utterances.

The focus of this chapter is on certain processing issues. Levinson claims that his theory is not an account of how sentences that give rise to GCIs are processed; i.e., he is not offering a model of speaker-hearer pragmatic performance. Rather, his is an abstract account of the information that must be available to the processor; i.e., he is offering a model of a speaker-hearer's pragmatic competence, or least of some aspects of this competence (Levinson 2000, 187, 397n5). However, it is also clear that processing issues are important to Levinson, and he frequently makes speculations about such issues (Levinson 2000, 5, 162–63, 245, 247, 259). Moreover, he claims that the system of default inferences that he posits has developed in order to overcome the processing bottleneck created by the slow articulation rate of speech. So there is ample justification for looking at Levinson's views from a processing perspective and asking whether his view adequately captures certain important facts about how utterances are processed.

In the next section I lay out Levinson's views. Later I raise some problems for Levinson's account. Finally, I take up the issue of flexible pragmatic processing, and argue for a model of pragmatic processing that is less rigid than the one advocated by Levinson. My views have affinities with those of relevance theorists such as Sperber and Wilson (1986) and Robyn Carston (1997, 1998b).

LEVINSON'S THREE LEVELS OF MEANING AND THEORY OF DEFAULT PRAGMATIC PROCESSING

Levinson argues that GCIs belong to a level of meaning intermediate between sentence-type meaning on the one hand and speaker meaning on the other. He calls this the level of utterance-type meaning, and resists the idea that it can be reduced to either of the other levels. GCIs are preferred but defeasible interpretations invoked by the use of certain expression-types, especially sentence connectives (e.g., 'and', 'if', 'not'), quantifiers (e.g., 'all', 'some'), modal operators (e.g., 'necessarily', 'possibly'), articles (e.g., 'a', 'the') and pronouns (e.g., 'he', 'she'). However, the use of other sorts of expressions can also invoke GCIs. Some examples will be discussed below.

To arrive at these preferred interpretations, we rely on a set of default heuris-

tics, which are “frameworks of assumption that can be taken to amplify the coded content of messages in predictable ways unless there is an indication that they do not apply.” (Levinson 1995, 96). One might ask why such a system of default heuristics has developed. Levinson’s answer is that they have developed to overcome the bottleneck in processing created by the slow articulation rate of speech. Levinson identifies three inferential heuristics used in language production and understanding. Levinson (2000) calls these the Q-, I- and M-Principles.

The Q-Principle

This principle can be summarized in the slogan ‘What is not said is not the case’. From a speaker’s point of view it enjoins one not to make statements informationally weaker than one’s knowledge of the world allows. From the hearer’s point of view it invites one to assume that the speaker has made the strongest statement consistent with what she knows. This heuristic operates under a constraint. It applies only in cases in which there are expressions that form a *contrast set*. So-called Horn or entailment scales would be examples of contrast sets. $\langle S, W \rangle$ forms an entailment scale just in case the ‘stronger’ expression S is such that when it is “substituted in an arbitrary declarative sentence” that sentence “will entail the same sentence with W [the ‘weaker’ member of the pair] substituted for S .” (Levinson 1995, 98).¹ For instance, $\langle \text{all, some} \rangle$ forms an entailment scale. However, Levinson allows that there are other kinds of contrast sets in addition to such entailment scales, as the illustrations below give some hint of.

When the constraint on the Q-Principle is satisfied, i.e., when there is an appropriate contrast set $\{F, G\}$, and $A(_)$ is a sentence-frame of the appropriate sort, then when a speaker asserts that $A(G)$ she implicates that she knows or that she judges that $A(F)$ is false or perhaps more weakly that she does not know that or that she withholds judgment about whether $A(F)$ is true. When Levinson gives examples of such Q-implicatures, he rarely includes any such epistemic modification, but Levinson (2000, 79) makes it clear that it is always to be implicitly understood. Examples are as follows:

- (1) $\langle \text{all, some} \rangle$
 S-sentence: All of the students were in class.
 W-sentence: Some of the students were in class.
 GCI from assertion of W-sentence: (The speaker knows/judges that) Not all of the students were in class.
- (2) $\langle \text{none, not all} \rangle$
 Not all of the students came.
 GCI: (The speaker knows/judges that) Some of the students came.
- (3) $\langle \dots, 3, 2, 1 \rangle$
 We scored two goals.

GCI: (The speaker knows/judges that) We scored no more than two goals.

- (4) <and, or>

You can have either soup or salad.

GCI: (The speaker knows/judges that) You cannot have both soup and salad.

- (5) <hot, warm>

This tea is warm.

GCI: (The speaker knows/judges that) This tea is not hot.

- (6) {white, red, blue,...}

The flag is white.

GCI: (The speaker knows/judges that) The flag is only white, not red, white and blue.

- (7) Contrast set consisting of a superordinate and its subordinates.

I saw an animal in the larder.

GCI: The speaker doesn't know whether it was a mouse, a rat, a squirrel or what sort of animal it was that was in the larder.

The < , > notation indicates that the contrast set is an entailment scale. The { , } indicates that the set of expressions contrasts in some way even though there is not a one-way entailment relation holding between sentences containing a stronger member and a weaker member respectively. Q-implicatures of the sort illustrated in (1)–(5) are called *scalar implicatures* and have been extensively discussed by others besides Levinson, such as Horn (1984, 1989) and Hirschberg (1991).

So-called *clausal implicatures* are also included under the heading of Q-implicatures. Levinson (2000, 76) claims that these are generated when there is a contrast set of expressions {*S*, *W*} and a sentence frame *A*() and a speaker's assertion of *A*(*W*) fails to entail an embedded sentence *Q*, which the assertion of a stronger sentence *A*(*S*) would have entailed. In such a case the assertion of the weaker sentence implicates that the speaker does not know whether or not *Q* obtains. In this analysis of clausal implicatures Levinson is following Gazdar (1979).

- (8) {and, or}

Stronger sentence: John is a poet and a philosopher.

Entails: John is a poet/ John is a philosopher.

Weaker sentence: John is either a poet or a philosopher.

GCI from assertion of the weaker sentence: For all the speaker knows, John is perhaps a poet or perhaps not a poet, perhaps a philosopher or perhaps not a philosopher. (cf. Levinson 2000, 109).

- (9) {knows, believes}

I believe that there is life of Mars.

GCI: For all the speaker knows, it is possible that there is life on Mars and it is possible that there is not life on Mars.

The I-Principle

This principle can be summarized in the slogan ‘What isn’t said is the obvious’ (Levinson 1987b) or, less perspicuously, ‘What is simply described is stereotypically and specifically exemplified’ (Levinson 1995). From the speaker’s point of view, following this heuristic means minimizing what one says when the hearer is able to use contextually accessible information to enrich the informational content of one’s utterance. From the hearer’s point of view following this heuristic means amplifying or enriching the informational content of the speaker’s utterance up to the point that one judges is the speaker’s intended meaning. Once again, there are constraints on the application of the I-Principle. It applies only to “unmarked, minimal expressions” (Levinson 1995, 97). Moreover, this heuristic invokes “world knowledge of stereotypical relations” (ibid.), and it is on the basis of this information about stereotypes that the hearer will engage in the enrichment of the speaker’s minimal expression to arrive at her intended meaning. Examples are as follows:

- (10) John smiled at the *secretary*.
GCI: John smiled at the female secretary.
- (11) *John’s book* is on the table.
GCI: The book John owns/bought/borrowed/read/wrote, etc., is on the table.
- (12) *Conjunction-buttrissing*
Susan turned the key *and* the engine started.
GCI: Susan turned the key and then the engine started.
GCI: Susan turned the key and as a result the engine started.
GCI: Susan turned the key intending it to be the case that the engine started.
- (13) *Conditional perfection*
If you cooperate, there’ll be no trouble.
GCI: If you don’t cooperate there’ll be trouble.
Combination of what is said and GCI: If and only if you cooperate, there’ll be no trouble.
- (14) *Negative strengthening*
I don’t like Alice.
GCI: I positively dislike Alice.
- (15) *Bridging inferences*
Harold bought an old car. *The steering wheel* was loose.
GCI: The steering wheel of the car Harold bought was loose.

The italicized expressions are the “unmarked, minimal expressions” that trigger the I-Principle, thereby giving the hearer access to stereotypical information for enrichment purposes. What exactly the stereotypical information is that is accessed is clear enough in some cases (e.g., we presumably all have something like a mental schema for the typical parts of a car). It is not so clear in other cases (e.g., it is unclear what stereotypical information is used in the inference from ‘If P then Q’ to ‘If not P then not Q’).

It should be noted that enrichments do not always rely on stereotypical information. Levinson concedes this possibility and contrasts example (15) with the following (Levinson 2000, 127):

(16) Hilda climbed into the capsule. The proton thruster was attached to the console.

GCI: The capsule’s proton thruster was attached to the capsule’s console.

It is likely that in order to interpret (15) we activate a mental schema, which tells us that cars have steering wheels. But we have no stereotype, or at least no highly developed one, for capsules and proton thrusters. So this cannot be how we interpret (16). Levinson’s answer is that in this case we make a presumption that the proton thruster is part of the capsule. He says: “It is the presumptive quality of the inference, even in the absence of stereotypical information, that is a hallmark of I-implicature” (Levinson 2000, 127).

It can’t be right that we “just presume” in a case like (16). When a speaker introduces a topic the hearer naturally expects that the speaker’s next reference will be related to the topic in some way. For example, the next reference may be to a property of the thing being talked about, or perhaps to something that is a part of the thing being talked about, or is contiguous to it in space or time, or is related to it in some more abstract way. Thus even though the hearer has only the haziest mental schema for a capsule and has no idea what a proton thruster is, he will at least know that consoles contain instruments for operating machinery. If the proton thruster is on the console it must be part of the equipment needed to operate the capsule. In short, even though the hearer has no stereotype for proton thrusters, the reference is one that is easily accessible, given other, admittedly very general, assumptions operative in the context.

It seems likely that it is this notion of accessibility that is more fundamental than the notion of stereotypicality in the process of informational enrichment/ pragmatic narrowing that Levinson says is encouraged by the I-Principle.

If to say that the inference in (16) is a presumptive one is just to say that it is one that is accessible in the sense just explained, then we can agree with Levinson. There has to be some route via which the reference can be retrieved, which means that there must be certain very general assumptions in play in the context. If there were no such route, or no easily accessible one, then no impli-

cature would be retrieved and the reference would simply fail. But if “presumptive” is here supposed to mean “default” then there is no reason to agree with Levinson, as I will argue in section 3 below.

The M-Principle

This principle can be summarized in the slogan ‘Marked or more prolix expressions warn of an abnormal situation’. From the speaker’s point of view this heuristic directs one to use marked expressions only if one wants to draw the hearer’s attention to something unusual in the situation. From the hearer’s point of view the result of following this heuristic is to arrive at interpretations that are the complement of the ones that would have been induced had the speaker used an unmarked or brief expression. Clearly, the operation of this heuristic is constrained to situations in which a marked or prolix expression is used. It requires a comparison between unmarked and marked expressions. In other words, the hearer must realize that there is a simpler, less marked way of saying the same thing. Examples are as follows:

- (17) (stopped, caused to stop)
James caused the car to stop.
GCI: James caused the car to stop in some non-stereotypical way, e.g.,
by using the hand brake.
- (18) (ate, ate and ate)
He ate and ate.
GCI: He ate more than a normal meal.
- (19) (chair, chairperson)
The chairperson will be arriving soon.
GCI: The female chairperson will be arriving soon.
- (20) (drink, beverage)
He asked for a beverage.
GCI: He asked for a nonalcoholic drink.

The pairs of expressions in parentheses are what Levinson calls “lexical doublets”, with the first member of the pair being the unmarked form and the second the semantically equivalent marked form. Levinson (1995, 104) lists a number of such lexical doublets and the I- and M-implicatures that they would generate.

Of course it is not necessary to use a marked expression in order to convey what is nonstereotypical. Nonstereotypical understandings can arise from the very same sentences that in other contexts yield stereotypical understandings, provided the wider context is set appropriately. For instance, an utterance of ‘John stopped the car’ will be interpreted to mean John stopped it in a non-stereotypical way if the context leading up to this utterance triggers the neces-

sary assumptions. Perhaps the hearer is told that John ran out into the road towards an oncoming car rolling a large tree trunk ahead of him. This is a point that Levinson does not explicitly acknowledge. But it is one he must acknowledge, since it follows from the fact that I-implicatures are cancelable by contrary information available in context.

Interactions between the Principles

The implicatures generated by the Q-, I-, and M-Principles may sometimes conflict. That this might happen should seem especially obvious with respect to the Q- and I-Principles. After all, the Q-Principle invites the hearer to assume that the speaker has said as much as she can, while the I-Principle encourages the hearer to assume that the speaker has said as little as she can get away with. And of course the I- and M-Principles generate incompatible implicatures, since “M inferences are specified as the complement of I inferences.” (Levinson 1995, 105). To resolve potential clashes between GCIs there are *priority rules*: Q inferences take precedence over inconsistent M inferences and both Q and M inferences take precedence over inconsistent I inferences (Levinson 1995, 106). Levinson (1987b, 71) offers a schema for resolving potential clashes between GCIs, and Levinson (2000, 157) adopts essentially the same “resolution schema”:

- (a) Genuine Q-implicatures from tight Horn scales (and their metarule derived negative scales) and similar contrast sets of equally brief, equally lexicalized expressions “about” the same semantic relations, take precedence over I-implicatures;
- (b) In all other cases, the I-principle induces stereotypical specific interpretations, unless:
- (c) A marked expression has been used where an unmarked one could have been employed instead, in which case the M-implicature defeats the relevant I-implicature, by inducing the inference to the complement of the I-implicature that would have arisen from the unmarked expression.

Levinson discusses various potential clashes, for instance between I-implicatures and Q-clausal-implicatures or between clausal implicatures and M-implicatures. He shows how in each case one or another of the implicatures is blocked, thereby averting a conflict (Levinson 2000, 155–64). Psycholinguistically what happens according to Levinson is that “all the potential implicatures are generated and are then filtered out by a process of incremental addition to the context only if consistency is satisfied.” (Levinson 2000, 162).

In fact though, there are fewer potential clashes than one might have thought, due to the restrictive conditions on the formation of contrast sets. For instance, if we were allowed to form the entailment scale <and then, and> the following sentence would generate inconsistent GCIs by the Q- and I-Principles:²

(21) Susan turned the key and the engine started.

I-implicature: Susan turned the key and then the engine started.

Potential Q-implicature: It is not the case that Susan turned the key and then the engine started.

The I-implicature is a case of conjunction buttressing. The Q-implicature would be a case of the negation of a stronger sentence being implicated by the assertion of a weaker sentence. Moreover, according to the priority rules, the Q-implicature would take precedence over the I-implicature, which would give the wrong result because it is the I-implicature that would normally be understood to follow from the assertion of (21).

Denying that there is even a potential Q-implicature here solves this problem. According to the resolution schema, the expressions in a contrast set must be equally brief, equally lexicalized linguistic expressions “about” the same semantic relations. Since ‘and then’ is not lexicalized in the way that ‘and’ is, and the temporal relation signaled by ‘and then’ does not belong to the same semantic domain as the logical relation signaled by ‘and’, there is no scale <and then, and> and hence no Q-implicature.

SOME PROBLEMS FOR LEVINSON’S THEORY OF GCIs

In this section I raise some problems for Levinson’s theory of GCIs. There are two major problems that I see for this account. First, when one considers the mechanisms whereby such GCIs are derived, it becomes less clear in what sense we should say that GCIs are associated with expression- or utterance-*types*. This is especially true of cases falling under the I-Principle, although there are problems too with cases falling under the Q-Principle. But if these GCIs are not tied to expression- or utterance-*types* then it is not clear that we need acknowledge a level of meaning intermediate between sentence meaning and speaker meaning. Second, Levinson’s account of default inferences is supposed to be motivated by the need to overcome a bottleneck in processing. But there appear to be cases in which multiple GCIs are generated by Levinson’s default heuristics, which would exacerbate processing difficulties rather than alleviating them. On the other hand, if Levinson picks one of the possible GCIs as the default, then he faces the problem that there will be many circumstances in which accessing the default leads the hearer down an interpretive dead-end. This once again has negative consequences for Levinson’s account of the pragmatic processing of utterances.

Is There a Level of Utterance-Type Meaning?

Levinson makes it clear that one of his central claims is that there is a level of meaning distinct from either sentence meaning or speaker meaning. He rails against those he calls reductionists, who attempt to assimilate GCIs either to conventional meanings on the one hand or to particularized implicatures on the

other. He thinks that his theory of GCIs shows that there are sets of default inferences that are associated with particular expression- or utterance types, and he claims that these inferences will be drawn unless some other information available in the context cancels them.

For example, the I-Principle is triggered by the use of unmarked, minimal expressions. Once triggered this principle will “invoke and filter” stereotypical information (Levinson 1995, 96). This stereotypical information is used in a process of informational enrichment/pragmatic narrowing in order to yield a GCI. A question thus arises as to what is involved in the accessing of this stereotypical information that leads to informational enrichment. Such information does not have to be accessed via the processing of the minimal form itself, nor from the processing of other expressions belonging to the same utterance. The stereotypical information can come from any source in the context, including nonlinguistic sources. Sometimes it is the processing of the underspecified form itself that triggers the stereotype; e.g., in (10) the word ‘secretary’, which is underspecified for gender, gets informationally enriched to ‘female secretary’ because of the stereotype that most secretaries are female. But in other cases the stereotype need not be associated with the utterance type itself.

Take, for instance, the process that Levinson calls conjunction-buttrressing, illustrated below:

- (22) Susan turned the key *and* the engine started.
 - (a) GCI: Susan turned the key and then the engine started.
 - (b) GCI: Susan turned the key and as a result the engine started.
 - (c) GCI: Susan turned the key with the goal of bringing it about that the engine started.

Levinson gives various possible GCIs that might arise from the use of ‘and’ in a sentence such as (22). (Levinson 1995, 101; 2000, 38, 117). For example, it might give rise to either the temporal sequence implicature given in (22a) or the causal sequence implicature given in (22b) or the teleological interpretation given in (22c). Informational enrichment on the basis of stereotypical information will lead to one or another of these GCIs. But how is the needed stereotypical information accessed? It cannot be retrieved solely by processing the expression ‘and’ or even by processing the expressions in the two component sentences ‘Susan turned the key’ and ‘The engine started’. These elements are common to all of the GCIs (22a)–(22c). Thus hearers will need to rely on the information made accessible in the wider context, such as information from prior discourse context (i.e., the mutual linguistic context), from the mutual physical environment, or from other shared sources of knowledge. But if this is so, this casts doubt on the claim that GCIs are associated with expression- or utterance-types.

If the information utilized in the enrichment process is independent of utterance type, then the GCI itself is in some sense independent of utterance type.

Moreover, as Carston (1993, 1995, 1998a) has pointed out, some of the same sorts of causal and temporal understandings associated with conjunctions are retrievable even when no sentence connective is used. For example, ‘Susan turned the key. The engine started’ can suggest similar implicatures to those suggested by (22). Thus it is not conjunctions *per se* that are associated with default causal or temporal understandings, but at best sentences presented in a certain linear order. Note however that order of presentation is not an infallible clue to temporal ordering. ‘He hit her. She screamed at him’ can be interpreted to mean that the event described in the second sentence occurred temporally prior to the event described in the first sentence, and that it was the reason for the occurrence of the first mentioned event. This backward-pointing understanding is not possible for ‘He hit her and she screamed at him’. So adjoined sentences can give rise to implicatures that are not possible for their conjoined counterparts.

Levinson himself notes that adjoined sentences can generate some of the same GCIs that conjunctive sentences generate. He thinks that this vindicates his view that minimal expressions suggest enriched interpretations, in line with the I-Principle. He says: “Parataxis [the adjunction of sentences] is an important instance of the tendency to find from minimal specifications maximally cohesive, rich interpretations.” (Levinson 2000, 126). However, the fact that adjoined sentences invite some of the same inferences as conjunctions counts against the claim that it is the connective ‘and’ itself that carries the default meanings. And nor can it be that adjunction itself counts as an utterance-type to which these GCIs can be attached. Not all utterances of a sequence of adjoined sentences belong to the same utterance type, as the individual sentences uttered need not belong to any particular utterance type. Whether a GCI is generated by an adjunction and if so which one will depend on the particular content of the adjoined sentences, as well as on what other assumptions are accessible in the utterance context. This in turn counts against the claim that GCIs like (22a)–(22c) belong to a special level of utterance type meaning.

What has been said here about the case of conjunction buttressing could be said also about other cases calling for informational enrichment/ pragmatic narrowing in accordance with the I-Principle. For instance, consider

(23) *Professor White’s book* is on the table.

GCI: The book Professor White owns/bought/borrowed/read/wrote, etc., is on the table.

Different pragmatic narrowings of ‘Professor White’s book’ can be suggested by varying the wider context. Suppose the speaker and hearer are librarians work-

ing in the reserve room of a library, and Professor White has come to pick up a book he has reserved. Then the hearer will understand the speaker to be referring to the book Professor White has reserved. On the other hand, if the speaker and hearer are employees in a university bookstore setting up a display of books by the local faculty, then 'Professor White's book' might refer to a copy of the book authored by Professor White. Moreover, it is important to note that such stereotypical information need not be triggered by the use of any particular expression type. In fact, it may be that the stereotypical information needed for enrichment must be accessed by nonlinguistic means, say by relying on stereotypes associated with some object in the mutual perceptual environment of speaker and hearer.

Thus far I have tried to show that we have reason to think that if stereotypical information is used in the course of the enrichment processes required by the I-Principle, this information is not confined to stereotypical information made accessible through the use of certain types of expressions. Mental scripts or schemas that represent stereotypical scenarios can be accessed through the use of expressions outside the boundaries of an utterance, through background assumptions and even by nonlinguistic means. However, if the stereotypes needed in enrichment are independent of utterance type, this suggests that the presumptive meanings that depend on these stereotypes are themselves not associated with the utterance type. And this threatens the claim that presumptive meanings belong to a special level of utterance type meaning intermediate between sentence-type meaning and speaker meaning. And as go I-implicatures so go M-implicatures, since these are specified to be the contraries of I-implicatures and depend on the same sort of reasoning processes. M-implicatures are also pragmatic narrowings, although narrowings in a direction opposite from the stereotypical.

The problems discussed thus far arise from the fact that the operation of the I-Principle depends on features of the wider discourse context of an utterance, and is not confined to looking at features of particular utterance-types in isolation. It might be thought that Levinson's case for default interpretations associated with utterance- and expression-types is strongest in the case of Q-implicatures, because the operation of the Q-Principle is confined to features of expressions in isolation. So for instance, it may seem that the 'not all' interpretation will be triggered as soon as the word 'some' is encountered, since processing 'some' means accessing the lexical fact that there is an equally lexicalized but stronger expression from the same semantic field, namely 'all'. Recognition of the contrast set <all, some> will trigger the Q-Principle thereby yielding the Q-implicature 'not all'. Similarly, processing 'three' will immediately give rise to the GCI 'no more than three' and thus to the understanding 'three and no more'. However, Levinson recognizes that this cannot be the case. Levinson (2000, 247) says:

If pragmatic processing operated automatically and irrecoverably at the lexical level it would make the wrong predictions: a GCI cannot convert every occurrence of *two* into ‘two and no more’ because immediate collocations may make the decision one way (as with *at least two*) or the other (*exactly two*); . . . But if we retreat to the next level up, say the two-bar constituent level, collocations at the next level up may be immediately relevant (as in *a minimum sentence of two years* vs. *costs two dollars*), and so on, all the way up to the clause, the sentence, and ultimately the discourse.

Levinson is here worrying about the size of constituent that is processed at a time, and of course he does not want the conclusion that we hold off on processing until we have all the information that could possibly be relevant. Presumably he thinks that we need to draw the line somewhere above the lexical level but below the discourse level. In other places he seems to indicate that the heuristic principles will operate very early on in processing. For instance, he says that the heuristics proceed on a word-to-word basis and that therefore “a scalar quantifier like *some* will . . . already invoke default enrichments before the predicate is available” (Levinson 2000, 5). Later he says that the phrase *some of the boys* will trigger the default assumption ‘not all of the boys’ even before the predicate has been processed (Levinson 2000, 259).

However, the main point is not that Levinson is somewhat inconsistent about how much of the context surrounding an expression must be processed before the Q-Principle is triggered. Rather the point is that once one allows that the operation of the Q-Principle is not totally context-independent, it seems that one is on a slippery slope, where it will be hard to draw a principled boundary. If Levinson is going to make a case for GCIs belonging to a level of utterance-type meaning he ought to argue for a cut off roughly at the boundaries of sentences. Obviously we cannot say that there must be a cut-off at the boundaries of sentences on the grounds that GCIs belong to a level of utterance- or expression-type meaning. This would be question begging. But Levinson offers no other arguments in favor of any such cut-off.

When There Are Many Possible Enrichments, Which One Is the Default?

The Q-, I-, and M-Principles are supposed to be default heuristics. The idea that there is a default inference associated with a particular utterance-type suggests that there should be a single interpretation that comes to mind whenever that utterance is processed, unless something in the context warns that things are not normal. That there is a single interpretation may seem most plausible in cases falling under the Q-Principle. For instance, utterances of ‘Some S are P’ will, other things being equal, suggest ‘Not all S are P’. Even so, there are problems about what is the default when there is more than one possibility for a contrast

set. For example, when the speaker utters ‘Some of the students are here’, is the appropriate contrast set <all, some>, <most, some>, <many, some>, or perhaps <all, most, many, some>? If the latter, then are the denials of *all* the sentences stronger than the one with ‘some’ retrieved by default? If there is more than one default GCI, this would add to the bottleneck in processing that Levinson’s theory of default heuristics is meant to relieve. Levinson appears to believe that all these denials will be retrieved, but that the most salient one will be the denial of the head of the scale (Levinson 2000, 77). Perhaps this talk of the salience of one GCI over the others could explain why accessing multiple GCIs does not add to the bottleneck in processing.

Similar remarks could be made about the example ‘The flag is white’. Levinson suggests the GCI ‘The flag is only white, not red, white and blue’. But it looks on the face of it that if the contrast set consists of all the color names, then there will be indefinitely many GCIs. Which is the default? If all possibilities are accessed whenever the expression is processed this would lead to a bottleneck in processing. On the other hand, if which GCI is accessed varies from context to context, then in what sense is it a default? And then again, if only one interpretation counts as the default, there will be the need for error recovery in all cases in which something other than the default is the intended interpretation. None of these three options seems very attractive.

The idea of a single default interpretation seems even less plausible in the case of the I-Principle. Earlier (pp. 266–67) we saw that in some cases there are many informational enrichments possible for an underspecified form. Levinson’s system of default inferences is meant to overcome the processing bottleneck created by the slow rate of articulation, so it would not help matters to have all possible interpretations be accessed by default and then have context select the appropriate one. In this section I assume that in those cases in which multiple enrichments are possible, Levinson intends for only one of these to count as the default. I assume that if we select one of these possible enrichments as the default, then only it can count as a GCI. In those contexts in which another possible enrichment overrides the default one, the overriding interpretation should strictly speaking be called a particularized conversational implicature (PCI).

In this section I explore the plausibility of singling out one of the possible enrichments as a default. Consider:

(24) John’s book is on the table.

GCI: The book John owns/bought/borrowed/read/wrote, etc., is on the table.

Levinson (2000, 37, 207) grants that each of the suggested enrichments ‘book John owns/bought/borrowed/read/wrote’, as well as many more, are possible for ‘John’s book’. Clearly these cannot all be default inferences. So, what is the de-

fault inference that one draws when one hears (24)? Levinson (1995, 100–101) writes:

The construction *X's Y* merely indicates that *some* relation holds between the two noun phrases, and we resolve the relation by pragmatic inference. Thus the phrases *Jupiter's moons*, *John's ideas*, *Anne's address*, *the building's condition*, *the encyclopedia's editor*, *the year's end*, are each understood to involve different relations (gravitational capture, ideational authorship, postal access, etc.). Note that all these phrases seem to have a default interpretation: *John's pens* will naturally be taken to mean the pens belonging to John, unless the context (e.g. talk between pen-designers) warrants another less stereotypical interpretation.

It is true that it is hard to think of contexts in which 'Jupiter's moons' means anything other than *moons in orbit around Jupiter*. So, frequency of use might result in *moons in orbit around Jupiter* becoming the default, or at any rate a dominant, enrichment of 'Jupiter's moons'.³ But why should one of the possible enrichments suggested in (24) be the default interpretation of 'John's book'? And which one would it be? Books *per se* are not things that are stereotypically owned or bought or borrowed or read or written, and so on. Perhaps library books are stereotypically owned, and books on the shelf in a bookstore are stereotypically bought, and books listed on someone's CV are stereotypically authored, and books listed in a bibliography are stereotypically read. This suggests that it is only in the context of some activity that it makes sense to talk of what a book stereotypically is.

In his most recent work Levinson suggests that the phrase *John's book* has two stereotypical readings, namely 'the book John is reading' and 'the book John authored'. Consider the following sentence:

(25) John admires the book he's reading, but John's book is in fact better.

In reference to this example Levinson writes that we "tend to assume that the relation between John and the book in *John's book* is the other stereotypical person-to-book relation, namely authorship" (Levinson 2000, 223). However, for the reasons already given, it is problematic to talk of what is stereotypical about the relation between people and books unless we know more about what sort of situation we are dealing with. For instance, if I know that John in example (25) is one of the students in my beginning logic class it will not enter my head to assume that he has authored a book. So I will search in context for some other interpretation for 'John's book', perhaps coming up with something like 'the book John recommended as a good read'.

In the quotation from Levinson above, he says that a less stereotypical

interpretation of 'John's pen' might be warranted in a conversation among pen designers. But of course, *for pen designers*, the interpretation *pen designed by John* for the phrase 'John's pen' may be highly accessible, and so for them might count as the default interpretation. Similar sorts of remarks could be made with respect to 'John's book'. What the most accessible interpretation is may vary from person to person, and for a single person from one context to the next. What is stereotypical about books will be different for a librarian and for a sales clerk at a bookstore, and will be different for a librarian at work and a librarian when in some other situation.

It is not clear how Levinson thinks these defaults get set up. If it is familiarity due to past usage, or high accessibility due to background circumstances, then it looks as though there will be many defaults, each relative to a different set of background circumstances. But then why call these *default* interpretations, as opposed simply to *easily accessible* interpretations?

A similar point about the great variety of enrichments that are possible can be made with respect to conjunctions. Example (22) above ('Susan turned the key and the engine started') gave three possibilities for conjunction-butchressing that Levinson considers possible, namely the temporal, causal and teleological interpretations (Levinson 2000, 38, 117). But these do not exhaust the possibilities. There are many sorts of relations that can be understood to hold between the events or states described in the two conjuncts of a conjunction. This is a point that has been made forcefully by Carston (1991, 1993, 1995, 1998a, 1998b). Other than the enrichments in which the first event is understood to be the cause of the second, and in which the first event is understood to be temporally prior to the second, we have at least the following other possibilities:

- (26) The second event is understood to be temporally contained in first. For example, 'He went to London and he saw the Queen'.
- (27) The two events are understood to be contemporaneous. For example, 'She likes to ride her bike and to listen to her Walkman'.⁴
- (28) The first event is understood to enable but not to directly cause the second. For example, 'I forgot to hide the cake and the kids ate it'.
- (29) The second event is understood to come into being as the first unfolds. For example, 'I talked to Susan and found I liked her'.
- (30) The first event is the reason for the second. For example, 'His calculator gave the answer "3" and he wrote down "3" as his answer'.

Once again, this great variety of causal, temporal, and justificatory understandings cannot all be default interpretations, in the sense that all these possible enrichments are simultaneously accessed whenever a hearer processes a conjunction. That would defeat the idea of a default interpretation, which presumably

is meant to make processing easier, not to add to the processing load. So if Levinson is committed to the idea of default interpretations, he owes us an account of which of these possible enrichments counts as the default interpretation and why it is the default.

Levinson does at one point suggest that it is the temporal sequence GCI that is the default. He claims that cross-linguistically, and when there are no special prosodic or other markings, this is the favored interpretation. He does grant that “occasionally the very search for some rich connectedness may require the rejection of the ‘and then’ interpretation in favor of a reversed sequence (prosodic clues being important here)” (Levinson 2000, 123). He gives the following example:

(31) He got a Ph.D. and he only did a month’s work.

It is unclear whether in (31) the temporal sequence implicature is derived and then cancelled or whether it is not even considered in the first place. The talk of the “rejection” of the ‘and then’ interpretation suggests that it is derived and then canceled. But then Levinson is committed to a processing account according to which the hearer is led down an interpretive dead end. So the system of default reasoning, far from speeding up processing, actually slows it down. On the other hand, if the temporal sequence implicature is not even considered, because the “search for rich connectedness” yields a better interpretation, then Levinson is once again on a slippery slope. It is going to be hard to draw a principled boundary between those cases in which contextual considerations ought to be weighted more heavily and those in which the system of defaults is going to be allowed to operate.⁵

The enrichment processes triggered by the I-principle appear to be highly context dependent. This speaks against there being any single default enrichment associated with a particular utterance type. With the I-Principle one is not forced to a particular interpretation for a particular utterance type. What may be true is that one is forced to engage in pragmatic narrowing/informational enrichment. But there is no one direction in which one is forced to enrich.⁶ For example, consider cases of bridging inferences, in which the contents of referential expressions must be enriched by connecting them to other items in the wider discourse context. An expression such as ‘the steering wheel’ could mean ‘the steering wheel of the old car Harold bought’ as it does in example (15) above, but in another context it might mean ‘the steering wheel of the truck that crashed on I-26 last night’ and so on indefinitely. Similar remarks could be made about ‘John’s book’, though the options for enrichment may be somewhat more constrained, and somewhat more constrained again for the connective ‘and’. In the latter two cases one might even hope to list all the possible enrichments, or at least to categorize them in some useful way.

FLEXIBLE PRAGMATIC PROCESSING

Levinson at one point calls his theory of GCIs a “generative theory of idiomaticity” (Levinson 1995, 94). Idioms are usually thought to be chunks of meaning that are stored as wholes in the lexicon. So they can be retrieved without being compositionally generated from their parts. Thus the notion of a generative theory of idiomaticity seems somewhat paradoxical. However, I think that the main point Levinson is trying to get across is that his theory of GCIs speeds processing in the same way that the retrieval of ready-made chunks of meaning from the lexicon would. If there is a default interpretation available, this is something like a ready-made chunk of meaning that can be easily retrieved, thereby lightening the hearer/reader’s processing load.

I argued in the previous section that it is far from clear that Levinson’s account has the desirable features that he thinks it has. He faces a trilemma. If there are many possible GCIs for a given sentence-type, this will make processing more difficult rather than alleviating the bottleneck in processing. But if which GCI is communicated varies from context to context, then it is not clear that we can talk about a system of default inferences. And if Levinson insists that one of the many possible GCIs is the default, then he achieves his goal of having a ready-made chunk of meaning available, but it is unclear that this sort lack of flexibility in pragmatic processing is ultimately a good thing. It may speed up processing in some cases, but it could substantially hinder processing in other cases. In every situation in which something other than the default is intended, the default will have to be cancelled/overridden/suppressed. This will require processing resources, thus increasing the hearer/reader’s processing load.

An alternative account, along the lines of the one offered by relevance theorists such as Sperber and Wilson (1986, 1995) and Carston (1997, 1998b) would avoid the problems that I have been discussing. First, I argue that the same sort of pragmatic reasoning is involved in the recovery of GCIs as is used to arrive at PCIs. Given that this is so, it is not an embarrassment to discover that GCIs may depend for their recovery on information made accessible by expressions outside the boundaries of the sentence uttered, and even on information made accessible by nonlinguistic means. Second, although there is a role for mental scripts and schemas and other bits of stereotypical information in relevance theory (RT), the theory also recognizes that nonstereotypical information can sometimes be salient or highly accessible. So RT can account for those cases in which pragmatic enrichment occurs even in the absence of stereotypical information, and for enrichments that go in a direction incompatible with known stereotypes. Third, the theory does not rely on there being default interpretations, and thus the theory predicts that hearers will not have to recover from interpretive errors as frequently as Levinson’s account predicts. This third issue is admittedly one that can only be settled experimentally. I am currently engaged in an experimental project to determine whether default interpretations play a

role in utterance interpretation. Bezuidenhout and Cutting (forthcoming) reports some initial findings in this area.

In what follows I offer a sketch of an alternative view of the pragmatic processing of the sorts of sentences that Levinson claims give rise to GCIs. This alternative account is supposed to be a performance model, but it emphasizes the comprehension side of pragmatic processing, rather than the production side. Here as elsewhere in the paper, when I talk of speakers I mean to include writers and other language producers, and when I talk of hearers, I mean to include readers and other language consumers.

Carston (1997, 1998b) has argued that the literal meaning of a sentence underdetermines the proposition expressed by the utterance of that sentence in some conversational context.⁷ By a process of decoding, the hearer's language system arrives at a representation of the logical form of the speaker's utterance. The lexical concepts that are constituents of the logical form of an utterance then have to be pragmatically processed, undergoing such processes as enrichment (narrowing) or loosening (broadening). These are "local" processes, in the sense that they begin operating before a whole sentence has been processed. Despite the fact that these processes operate over subsentential expressions, they are inferential processes. Lexical concepts in combination with other concepts accessible in the context are used to draw conclusions about the intended interpretations of underspecified forms. That is, the underspecified lexical meanings are used as one clue along with other contextually available information, including nonlinguistic information, to arrive at informationally enriched or loosened interpretations. An example of enrichment is the derivation of *exactly three* from 'three'. An example of loosening is the derivation of *roughly hexagonal* from 'hexagonal'. The results of such local pragmatic processing are *ad hoc* concepts, which become constituents of an overall representation of the proposition expressed by that utterance. This propositional content can then be used in further pragmatic processing to arrive at further, indirectly communicated contents (classical Gricean particularized conversational implicatures). All this pragmatic processing is driven by the search for an interpretation that is optimally relevant, in Sperber and Wilson's sense.

For Sperber and Wilson (1986), the relevance of an utterance is a matter of the balance between the cognitive effects that are generated by processing that utterance and the cognitive effort it takes to process that utterance. According to the Communicative Principle of Relevance, every utterance communicates a presumption of its own *optimal relevance*. According to the presumption of optimal relevance, an utterance communicates (1) that it is relevant enough to be worth the hearer's effort to process it, and (2) that it is the most relevant utterance compatible with the speaker's abilities and preferences. (Sperber and Wilson 1995, 270). As Carston (2000, 90–91) explains, the first clause of this definition sets a lower limit on the cognitive effects the hearer can hope for, namely

sufficient contextual effects. The second clause sets an upper limit on effects. Although an utterance may achieve more than mere adequacy, this is limited by the speaker's abilities (e.g., how much she knows) and by her preferences (e.g., how helpful she wants to be). Similar remarks can be made for the effort side of the effort/effects equation. The first clause guarantees that the hearer will expend no excessive effort, while the second clause promises the least possible effort commensurate with the speaker's abilities (e.g., vocabulary limitations) and her preferences (e.g., her dislike for directness). The comprehension strategy that is warranted by this presumption of relevance, as Carston (1998a, 214) makes clear, is that the hearer should consider possible interpretations in their order of accessibility. The hearer should stop processing when the expected level of relevance—viz., the level compatible with the speaker's abilities and preferences—is achieved or appears unachievable.

What I propose is that the contents that Levinson calls GCIs are in fact pragmatic developments of underspecified forms, and hence are arrived at by local pragmatic processes of the sort just described. Here I am going considerably beyond what relevance theorists have been prepared to claim. Carston (1998b) and in personal communication has made it clear that she thinks only some of what Levinson calls GCIs can be handled in this way. In this chapter I have not systematically investigated many of the implicatures that Levinson includes under the heading of GCIs. I have not discussed clausal implicatures, M-implicatures, and some I-implicatures, such as those arising from negative and conditional strengthening. Thus my sweeping reductive claim is best interpreted as gesturing to a direction for future investigation. What my view requires is that all the forms of expression that Levinson claims give rise to GCIs are semantically underspecified. Although they would not necessarily support the claims being made here, others such as Atlas (1989) and Ruhl (1989) have argued for the pervasive nature of semantic underspecification or semantic generality.

Those who are attracted by Levinson's views ought to find the idea of semantic underspecification unproblematic for cases falling under Levinson's I-Principle, since those are cases in which minimal forms call for pragmatic narrowing in the light of stereotypical information that is invoked in the context. The difference of course is that Levinson would say that the pragmatic narrowings are default interpretations, whereas I would deny that they are, for the reasons given above (pp. 275–76). On my view, the narrowings are simply contextually salient or accessible interpretations.

The greatest difference between my account and Levinson's arises with respect to cases falling under Levinson's Q-Principle. Levinson would claim that expressions that fall under this principle have specific meanings that contrast with other expressions with specific meanings that belong to the same semantic field or domain. The use of one of these contrasting meanings implicates the denial of the other meaning(s). The alternative view I am advocating claims in-

stead that such expressions are semantically underdetermined. They must be specified in context and how they are specified depends on the operation of the sorts of local pragmatic processes described above.

Take, for instance, a cardinal sentence such as:

(32) Jane has three children.

Levinson accepts an 'at least' semantics for cardinals. So, according to him, what is said by (32) is:

(33) Jane has *at least* three children.

The Q-Principle will be triggered by the cardinal expression, thereby yielding the GCI:

(34) Jane has *at most* three children.

What is said together with the GCI entail that:

(35) Jane has *exactly* three children.

This process would be Levinson's explanation for why in many contexts people understand (32) in the 'exactly' sense. The alternative account argues that (32) is semantically underdetermined. A pragmatic process will take the semantically underspecified concept *three children*, and yield a contextually appropriate enrichment. Depending on the assumptions accessible in the context, the proposition expressed could be (33), (34), or (35). In particular, to understand the speaker to have been communicating (35), the hearer will not need to go through a process whereby (33) and (34) are retrieved as well.

One argument in favor of this rival account is that it can deal with cases in which the intended meaning of a cardinal expression appears to be the 'at most' interpretation. Consider:

(36) Sally can eat 2000 calories without gaining weight.

In (36) the speaker should be understood to mean that Sally can eat at most 2000 calories without putting on weight. The underdetermination view claims that the expression '2000 calories' is semantically underdetermined, and that a relevance-driven local process of pragmatic narrowing will yield the interpretation *at most 2000 calories*. Levinson has to say that '2000 calories' means *at least 2000 calories* and that this implicates *at most 2000 calories*. But then by combining what is said with what is implicated there will be no way to block the

inference to *exactly 2000 calories*. The only way to get the *at most* understanding would be to cancel what is said, but this is not something that Levinson indicates he would be willing to sanction. (36) would have to be treated as a case of nonliteral speech, which intuitively it is not.

There are cases that might appear to support Levinson's theory of default GCIs over the semantic underdetermination view. Consider the following:

- (37) In the game of soccer, if each side gets three goals, the game is a draw.
- (38) In the ancient Toltec sacred ball game, if each side got three goals, the game was a draw.

In both these cases it is natural to interpret 'three goals' as *exactly three goals*. In the case of (37) the underdetermination view can claim that the hearer appeals to background knowledge about the rules of the game of soccer in order to informationally enrich 'three goals' to yield *exactly three goals*. But in the case of (38) no such explanation is possible. There is no background knowledge of the rules of Toltec sacred games that the hearer can appeal to. For all the hearer knows, the aim of the Toltec game could have been "to stop the other side getting three goals, after which goals were no longer determinative of victory" (Levinson 2000, 215). Levinson's theory of default GCIs on the other hand can claim that in both cases the Q-Principle is triggered by the use of a scalar item, yielding the interpretation *three and no more goals*, which incorporates the Q-implicature *no more than three*.

However, the proponent of the underdetermination view could say that in the case of (38) the hearer will rely on background knowledge of games with which he is familiar (like soccer) and what it means for a game to be a draw in these cases. The hearer will use this information to enrich the expression 'three goals' to arrive at the interpretation *exactly three goals*. Of course, this strategy will lead to the wrong interpretation if the Toltec sacred ball game was played as Levinson imagines it was. But the strategy of relying on Levinson's Q-Principle also leads to the wrong interpretation. Thus it does not seem that Levinson's account has an edge in this particular case.

Just as my account of the interpretation of cardinal expressions differs from Levinson's account, so does my account of the interpretation of quantificational sentences differ from Levinson's. Consider:

- (39) I ate some of the cookies.

According to Levinson what is said by the utterance of (39) is that the speaker (Anne, let us suppose) ate at least some of the cookies. This has the scalar implicature that Anne did not eat all of the cookies. The alternative account says that the sentence is semantically underdetermined. In some contexts the hearer

will understand the speaker of (39) to have directly expressed the proposition *Anne ate some but not all of the cookies*, whereas in others the hearer will understand the speaker to have expressed the proposition *Anne ate at least some and possibly all of the cookies*. Carston (1995, 1998a) has done a good deal of work to defend this view. Horn (1992) has a response along Gricean lines to some of Carston's arguments.

One other difference between my semantic underdetermination account and Levinson's theory of GCIs is that my view takes in more under the heading of pragmatic developments of underspecified forms than is included under the heading of GCIs. In particular, my account would include quantifier domain restrictions as cases of pragmatic developments of underspecified forms. Consider:

- (40) Everyone is a vegetarian.
- (41) There is nothing to eat in the house.
- (42) Not many children scored an advanced pass.

In each one of these cases the hearer must restrict the quantifier domain in some way in order to understand what proposition the speaker has expressed. For example, in (40) the domain might be understood as restricted to all the guests invited to a particular dinner party. In (41) it might be understood that there is nothing to eat in the house that is appropriate for dinner, so that it would be true even if there happened to be a box of cereal in the house. Similarly, the speaker of (42) might be understood to be talking about all the elementary school children in a certain school district. Moreover, in each case many different pragmatic restrictions of the domain are possible, depending on the wider context. For instance, (40) could be used to talk about all the people living in a certain commune, or all the members of a certain family, and so on indefinitely. Similarly, (41) and (42) can be given multiple interpretations, by varying the wider context.

This makes quantifier domain restriction seem very similar to the cases of informational enrichment that Levinson includes under his I-Principle, such as conjunction buttressing, pragmatic narrowing of possessives, bridging inferences, and so on. Thus it seems arbitrary to exclude quantifier domain restriction from the theory of GCIs. On the other hand, if it is included in the theory under the I-Principle, this only strengthens the points made above (pp. 265–73). Domain restriction depends on stereotypical, or at least accessible, information that is utterance-independent. This suggests that the interpretation that is generated is itself utterance-independent. So the claim that GCIs belong to a special level of utterance-type meaning would be compromised. And as there is no single domain restriction that is associated with a quantifier phrase-type such as 'every person' or 'many children', this would undermine the claim that the operation of the I-principle gives rise to default interpretations.

Thus the semantic underdetermination account seems preferable, as it clas-

sifies cases together that seem intuitively to belong together. Stanley and Szabo (1999) have argued that quantifier domain restriction should not be thought of as the pragmatic narrowing of an underspecified form, but rather as a case of contextually specifying the value of a hidden indexical. They posit a hidden indexical element associated with the nominal in a quantifier phrase. If Levinson accepted that quantifier domain restriction fell under the heading of indexical semantics, then he might have a principled reason for treating this case differently from the cases that fall under his I-Principle. However, it is unclear that this is a direction that Levinson would be willing to go in. He is generally sympathetic to the Radical Pragmatics program, which attempts to reduce the number of phenomena needing to be explained by appeal to semantic principles and attempts instead to account for such phenomena by pragmatic means.

I have suggested that the sorts of processes that Levinson thinks are involved in pragmatic narrowing by means of the I-Principle are no different from those used in cases of quantifier domain restriction. This in turn suggests that the processes whereby GCIs are derived are very similar to those used in the derivation of particularized conversational implicatures (PCIs). From the point of view of an account of language understanding and production, the difference between GCIs and PCIs is not that great—the same sorts of relevance driven inferential processes may be involved in both. What differentiates between GCIs and PCIs is that GCIs are pragmatic developments of semantically encoded meaning, whereas PCIs are independent in some sense from encoded content. Exactly what this sense of independence amounts to is an issue that has been much discussed. I favor the notion of functional independence articulated by Carston (1991). Recanati (1991) criticizes this functional independence criterion, and Vicente (1998) defends it against the sorts of alleged counterexamples that Recanati proposes. I would argue along with Vicente that the prospects for reviving the functional independence criterion are better than some, including Levinson (2000, 196), have argued.

Levinson's criterion for separating out the GCIs from the PCIs is that the former are interpretations that have a default character and that belong to a special level of utterance-type meaning. Levinson says: "I shall presume that we want to define the types of content by the processes that yield them and the important semantical properties they have (e.g., default presumption, defeasability under distinct conditions)" (Levinson 2000, 198). But the arguments given above should make it seem less secure that GCIs have the property of being default presumptions. So Levinson's criterion for separating GCIs from PCIs seems no more secure than the functional independence criterion.

CONCLUDING REMARKS

My main aim in examining Levinson's theory of GCIs and in sketching the alternative semantic underdetermination account was to set up some clear hy-

potheses about the pragmatic processing of utterances that could be experimentally tested. As Levinson (2000, 370) notes: "There is very little psycholinguistic work directly addressed to implicature, and still less of this concerns online processing." Much more could be done to elaborate on and defend the alternative semantic underdetermination view that I have outlined in the above section. My account is not nearly as well worked out as Levinson's theory of GCIs. However, enough detail has been given to generate rival processing predictions that can be tested in on-line reading experiments of the sort that are familiar in psycholinguistics.

For example, one could set up contexts in which a sentence of the relevant sort (i.e., one that Levinson claims is associated with a default GCI) is followed by information that leads to one or another of its possible pragmatic enrichments. If the contextual bias goes against what Levinson would treat as the default GCI, Levinson's theory predicts that the default interpretation will be accessed and then will have to be canceled. The sort of processing difficulty predicted here is akin to the processing difficulties generated by so-called garden-path sentences, where the parser comes up with a syntactic analysis that later has to be rejected in favor of a different analysis. The underdetermination view on the other hand predicts no such garden-pathing in pragmatic processing. Since the target sentence is semantically underspecified and the contextual information needed to specify the interpretation only becomes available downstream in the processing, the processor will hold off on a definite interpretation until the needed information becomes available. The processor does not have to recover from an interpretive error.

Another area in which conflicting predictions can be tested is in regards to the processing of scalar predicates. For instance, as was mentioned in the previous section, the two accounts are associated with rival claims about what information must be accessed in order to interpret cardinal sentences. Levinson's account is committed to the view that the *exactly three X* interpretation of 'three X' requires accessing both the literal meaning (which he believes is specified as *at least three X*) and the GCI *at most three X*. The underdetermination view on the other hand argues that the *exactly three X* interpretation is directly generated, without the need to generate the contents at *least three X* and *at most three X*.

Although Levinson adduces many reasons to favor his theory of GCIs, the semantic underdetermination view seems able to account for much of same data by different means. Obviously not both accounts can be correct, so one way of deciding between the two views is to subject them to experimental tests of the sort gestured at above. I hope in this chapter to have taken one small step in the direction of devising such experiments. Any experimental work must be preceded by an attempt to articulate as clearly as possible what the options are. This critical examination of Levinson's views and the brief sketch of an alternative represent a first attempt to articulate these options.

NOTES

1. Levinson (2000, 79) says that an entailment scale is an ordered n-tuple of expression alternates $\langle x_1, x_2, \dots, x_n \rangle$ such that where S is “an arbitrary simplex sentence-frame” and $x_i > x_j$, $S(x_i)$ unilaterally entails $S(x_j)$.
2. Levinson (1983, 146–47) discusses a very similar example and talks about a clash between the (first) Maxim of Quantity and what he there calls the Principle of Informativeness.
3. It is always possible to indulge in science fiction. One can imagine some other celestial body capturing one of Jupiter’s moons. In such a situation ‘Jupiter’s moon’ might mean *moon that was captured from Jupiter’s orbit*. Barker (1995) distinguishes between lexical possessives, such as ‘John’s child’ and ‘John’s purchase’, and extrinsic possessives, such as ‘John’s cat’ and ‘John’s gift’. According to Barker only the latter sort are open to multiple interpretations that must be pragmatically derived. The interpretation of possessives of the former sort is grammatically constrained. A sign that a possessive is a lexical possessive is that the possessee nominal in such cases is relational, for it entails the existence of another thing.
4. The reduced form contributes to the ease with which this interpretation is retrieved. ‘She likes to ride her bike and she likes to listen to her Walkman’ does not suggest this interpretation as readily.
5. Levinson also admits that there are cases where “local conversational goals” can force an interpretation of a conjunction according to which it is just a list. In such cases the inferences to order and teleology are “not firm” (Levinson 2000, 163–64). For instance, if I ask you what you did today and you reply ‘I went downtown and dealt with some bills’, I might interpret this as just a list of things you did. On the other hand if I ask you where you went today and you reply in this way, then I might infer that you went downtown *and then* dealt with the bills, or that you went downtown *in order to* deal with the bills. It is unclear what Levinson means by saying that in the first case the I-inferences are not firm. Does this mean they are accessed but in such a tentative way that they are easily canceled? Or does he mean that they are not accessed because “local conversational goals” preempt such interpretations? If the latter is intended this once again puts us on a slippery slope. It is difficult to say in a principled way when contextual information should be allowed to drive the interpretive process and when the default heuristics are to be allowed to operate.
6. Levinson (1995, 103) seems to acknowledge this when he says: “Inferences to the stereotypes are thus not ‘generalized’ in the sense that they are independent of shared beliefs . . . but they are ‘generalized’ in the sense that they follow a general principle—restrict the interpretation to what constitutes the stereotypical, central extensions.”
7. Kent Bach (1994a, 1994b, 1999) has also extensively explored the notion of semantic underdetermination. This notion has also been important in the work of certain cognitive and computational linguists, for instance, Pustejovsky (1998) and those whose work is collected by van Deemter and Peters (1996).

REFERENCES

- Atlas, J. D. 1989. *Philosophy Without Ambiguity*. Oxford: Oxford University Press.
- Bach, K. 1999. “The Semantics-Pragmatics Distinction: What It Is and Why It Matters.” In K. Turner (ed.), *The Semantics/Pragmatics Interface from Different Points of View*. Oxford: Elsevier.
- . 1994a. “Conversational Implicature.” *Mind & Language* 9: 124–62.
- . 1994b. “Semantic Slack: What Is Said and More.” In S.L. Tsohatzidis (ed.), *Foundations of Speech Act Theory: Philosophical and Linguistic Perspectives*. London: Routledge.
- Barker, C. 1995. *Possessive Descriptions*. Stanford: CSLI Publications.
- Bezuidenhout, A.L., and J.C. Cutting (forthcoming). “Literal Meaning, Minimal Propositions and Pragmatic Processing.” *Journal of Pragmatics*.
- Carston, R. 2000. “The Relationship between Generative Grammar and (Relevance-Theoretic) Pragmatics.” *Language and Communication* 20: 87–103.
- . 1998a. “Informativeness, Relevance and Scalar Implicature.” In R. Carston and S. Uchida (eds.), *Relevance Theory: Applications and Implications*. Amsterdam: John Benjamins.

- . 1998b. *Pragmatics and the Explicit-Implicit Distinction*. Ph.D. dissertation, University College, London.
- . 1997. "Enrichment and Loosening: Complementary Processes in Deriving the Proposition Expressed?" *Linguistische Berichte* 103–27.
- . 1995. "Quantity Maxims and Generalized Implicature." *Lingua* 96: 213–44.
- . 1993. "Conjunction, Explanation and Relevance." *Lingua* 90: 27–48.
- . 1991. "Implicature, Explicature, and Truth-theoretic Semantics." In S. Davis (ed.), *Pragmatics: A Reader*. New York: Oxford University Press.
- Gazdar, G. 1979. *Pragmatics: Implicature, Presupposition, and Logical Form*. New York: Academic Press.
- Grice, P. 1989. *Studies in the Way of Words*. Cambridge, Mass.: Harvard University Press.
- Hirschberg, J. 1991. *A Theory of Scalar Implicature*. New York: Garland Publishing, Inc.
- Horn, L.R. 1992. "The Said and the Unsaid." *Ohio State Working Papers in Linguistics* 40, 163–92.
- . 1989. *A Natural History of Negation*. Chicago: University of Chicago Press.
- . 1984. "Toward a New Taxonomy for Pragmatic Inference: Q-based and R-based Implicature." In D. Schiffrin (ed.), *Meaning, Form, and Use in Context: Linguistic Applications*. Washington, D.C.: Georgetown University Press.
- Levinson, S.C. 2000. *Presumptive Meanings: The Theory of Generalized Conversational Implicature*. Cambridge, Mass.: MIT Press.
- . 1995. "Three Levels of Meaning." In F.R. Palmer (ed.), *Grammar and Meaning*. Cambridge: Cambridge University Press.
- . 1987a. "Implicature Explicated?" *Behavioral and Brain Sciences* 10: 722–23.
- . 1987b. "Minimization and Conversational Inference." In J. Verschueren and M. Bertuccelli-Papi (eds.), *The Pragmatic Perspective*. Amsterdam: John Benjamins.
- . 1983. *Pragmatics*. Cambridge: Cambridge University Press.
- Pustejovsky, J. 1998. "The Semantics of Lexical Underspecification." *Folia Linguistica* 32: 323–47.
- Recanati, F. 1993. *Direct Reference: From Language to Thought*. Oxford: Blackwell.
- . 1991. "The Pragmatics of What Is Said." In S. Davis (ed.), *Pragmatics: A Reader*. New York: Oxford University Press.
- Ruhl, C. 1989. *On Monosemy: A Study in Linguistic Semantics*. Albany: State University of New York Press.
- Sperber, D., and D. Wilson. 1995. *Relevance: Communication and Cognition*, 2d ed. Oxford: Blackwell.
- . 1986. *Relevance: Communication and Cognition*. Oxford: Blackwell.
- Spooren, W. 1997. "The Processing of Underspecified Coherence Relations." *Discourse Processes* 24: 149–68.
- Stanley, J., and Z.G. Szabo. 2000. "On Quantifier Domain Restriction." *Mind & Language* 15: 219–61.
- van Deemter, K., and S. Peters, eds. 1996. *Semantic Ambiguity and Underspecification*. Stanford, Calif.: CSLI Publications.
- Vicente, B. 1998. "Against Blurring the Explicit/Implicit Distinction." *Revista Alicantina de Estudios Ingleses* 11: 241–58.

Distinguishing Semantics and Pragmatics

Kent Bach, *San Francisco State University*
and
Anne Bezuidenhout, *University of South Carolina*

SEMANTIC, PRAGMATIC: KENT BACH

THE DISTINCTION BETWEEN semantics and pragmatics has received a lot of bad press in recent years. It has been claimed to be faulty, confused, or even nonexistent. However, these claims are based on misconceptions of what the distinction is and of what it takes to show there to be something wrong with it. As I see it, the semantic-pragmatic distinction fundamentally concerns two types of information associated with an utterance of a sentence. Semantic information is encoded in the sentence; pragmatic information is generated by, or at least made relevant by, the act of uttering the sentence. This explains the oddity of such pragmatic contradictions as “I am not speaking” and “It is raining but I don’t believe it.” In “The Semantics-Pragmatics Distinction: What It Is and Why It Matters” (Bach 1999a), I develop this conception of the distinction and contrast it with alternatives. Here I will try to clarify that conception by showing how it avoids certain objections. Space will not permit going into much detail on the various linguistic data and theoretical considerations that have been thought to undermine the semantic-pragmatic distinction in one way or another.

Historically, this distinction has been formulated in various ways. These formulations have fallen into three main types, depending on which other distinction the semantic-pragmatic distinction was thought to correspond to:

- linguistic (conventional) meaning versus use
- truth-conditional versus non-truth-conditional meaning
- context independence versus context dependence

None of these distinctions does the job. The trouble with the first one is that there are expressions whose literal meanings are related to use. The second distinction is unhelpful because some expressions have meanings that do not contribute to truth-conditional contents. And the third distinction overlooks the fact that there are two kinds of context. This last point deserves elaboration.

It is a platitude that what a sentence means generally does not determine what a speaker means in uttering it. The gap between linguistic meaning and speaker meaning is said to be filled by “context”: what the speaker means somehow “depends on context,” or at least “context makes it clear” what the speaker means. But there are two quite different sorts of context, and they play quite different roles. What might be called “wide context” concerns any contextual information that is relevant to determining (in the sense of ascertaining) the speaker’s intention. “Narrow context” concerns information specifically relevant to determining (in the sense of providing) the semantic values of context-sensitive expressions (and morphemes of tense and aspect). Wide context does not literally determine anything. It is the body of mutually evident information that the speaker exploits to make his communicative intention evident and that his audience relies upon, taking him to intend them to do so, to identify that intention.

Another source of confusion is the phrase ‘utterance interpretation.’ Strictly speaking, sentences (and subsentential expressions), i.e., types, not tokens, have semantic properties. Utterances of sentences have pragmatic properties. Also, the term ‘interpretation’ is ambiguous. It can mean either the formal, compositional determination by the grammar of a language of the meaning of a sentence or the psychological process whereby a person understands a sentence or an utterance of a sentence. Using the phrase ‘utterance interpretation’ indiscriminately for both tends to confound the issues.

My conception of the semantic-pragmatic distinction involves certain assumptions about semantics and a certain view of communication. I take the semantics of a sentence to be a projection of its syntax. That is, semantic structure is interpreted syntactic structure. Contents of sentences are determined compositionally; they are a function of the contents of the sentence’s constituents and their syntactic relations. This leaves open the possibility that some sentences do not express complete propositions and that some sentences are typically used to convey something more specific than what is predictable from their compositionally determined contents. Also, insofar as sentences are tensed and contain indexicals, their semantic contents are relative to contexts (in the narrow sense). Accordingly, the following distinctions should be recognized:

- between a sentence and an utterance of a sentence
- between what a sentence means and what it is used to communicate

- between what a sentence expresses relative to a context and what a speaker expresses (communicates) by uttering the sentence in a context
- between the grammatical determination of what a sentence means and the speaker's inferential determination of what a speaker means (in uttering the sentence)

As for communication, when a speaker utters a sentence in order to convey something, the content of the sentence provides the basis for his audience's inference to what he is conveying and what attitudes he is expressing, e.g., belief in the case of assertion and desire in the case of requesting. In fact, as Bach and Harnish (1979, ch. 3) argue, because types of communicative speech acts may be individuated by the types of attitudes they express, their contents are simply the contents of the attitudes they express. That is one reason why the notion of the content of an utterance of a sentence has no independent theoretical significance. There is just the content of the sentence the speaker is uttering, which, being semantic, is independent of the speaker's communicative intention, and the content of the speaker's communicative intention. When one hears an utterance, one needs to understand the sentence the speaker is uttering in order to figure out the communicative intention with which he is uttering it, but understanding the sentence is independent of context except insofar as there are elements in the sentence whose semantic values are context relative. Recognizing the speaker's communicative intention is a matter of figuring out the content of that intention on the basis of contextual information in the broad sense.

This information does not literally determine that content. In no case does the semantic content of the uttered sentence determine what the speaker is communicating or, indeed, that he is communicating anything. That he is attempting to communicate something, and what that is, is a matter of his communicative intention, if he has one. If he is speaking literally and means precisely what his words mean, even that is a matter of his communicative intention. Communicative intentions are reflexive in the sense discovered by Grice: a communicative intention is one whose fulfillment consists in its recognition by the audience, partly on the basis that it is intended to be recognized. The role of Grice's maxims, or presumptions as they might better be regarded (Bach and Harnish 1979, 62–65), is to provide inference routes across any gap between what the sentence means and what the speaker aims to be communicating in uttering it.

This Gricean view of linguistic communication (it is developed in detail in Bach and Harnish 1979) lends itself to a certain conception of the semantic-pragmatic distinction. This distinction can be drawn with respect to various items, such as ambiguities, contradictions, implications, presuppositions, interpretations, knowledge, processes, rules, and principles, and, of course, 'semantics' and 'pragmatics' are also names for the study of these phenomena. For me

the distinction applies fundamentally to types of information. Semantic information is information encoded in what is uttered—stable linguistic features of the sentence—together with any extralinguistic information that contributes to the determination of the references of context-sensitive expressions. Pragmatic information is (extralinguistic) information that arises from an actual act of utterance, and is relevant to the hearer's determination of what the speaker is communicating.

This way of characterizing pragmatic information generalizes Grice's point that what a speaker implicates in saying what he says is carried not by what he says but by his saying it and sometimes by his saying it in a certain way (Grice 1989, 39). The act of producing the utterance exploits the information encoded but by its very performance creates new and otherwise invokes extralinguistic information. This extralinguistic information includes the fact that the speaker uttered *that* sentence and did so under certain mutually evident circumstances. This is context in the broad sense. Importantly, nonsemantic information is relevant to the hearer's inference to the speaker's intention only insofar as it can reasonably be taken as intended to be taken into account, and that requires the supposition that the speaker is producing the utterance with the intention that it be taken into account. There is no such constraint on contextual information of the semantic kind, which plays its role independently of the speaker's communicative intention. Contextual information in the narrow, semantic sense is limited to a short list of parameters associated with indexicals and tense, such as the identity of the speaker and the hearer and the time of an utterance. I may think I am Babe Ruth and be convinced that it is 1928, but if I say, "I hit 60 home runs last year," I am still using 'I' to refer to myself and 'last year' to refer to the year 2000.

Now let us consider some reasons that might be suggested for rejecting the semantic-pragmatic distinction. To the extent that the debate about it is not entirely terminological (e.g., many years ago 'pragmatics' was the name for indexical semantics), the main substantive matter of dispute is whether there is such a thing as "pragmatic intrusion," whereby pragmatic factors allegedly contribute to semantic interpretation. Here is an assortment of objections that are based on supposed pragmatic intrusion of one sort or another. Each of these objections is predicated on some misconception, as the responses indicate.

1. *Semantic phenomena are context independent, whereas pragmatic phenomena are context sensitive. But the meanings of certain expressions are context sensitive. Therefore, their meanings are not exclusively semantic.*

This objection assumes that anything pertaining to the use of an expression is automatically not semantic. However, the fact that the contents of certain expressions, notably indexicals and demonstratives, are context sensitive does not

show that their meanings vary with context. How their contents vary with context is determined by their fixed meanings, and that is a semantic matter. These variable contents are their semantic values.

2. *There are aspects of linguistic meaning that concern how a sentence is used, not its truth-conditional content. So linguistic meaning is not merely a semantic matter.*

This objection alludes to the fact that the meanings of certain expressions, what I call “utterance modifiers,” such as ‘to conclude,’ ‘frankly,’ and ‘to be precise’ (for a catalog of them see Bach 1999b, sec. 5), as well as grammatical mood, concern how a sentence is being used. However, all this shows is that semantics is not limited to what is relevant to truth-conditional content. There is no reason to assume that the linguistic meaning of a sentence cannot include information pertaining to how the sentence is used.

3. *Since language is rife with semantic underdetermination and vagueness, there is no such thing as literal meaning: sentence “semantics” is adulterated with pragmatics.*

These phenomena show only that sometimes the literal meaning of a sentence does not determine a complete proposition or a precise proposition. They do not show that there is no purely linguistic information on which language users rely. Take the case of semantically underdeterminate sentences, which do not express complete propositions, even modulo ambiguity and indexicality. Even though the following sentences do not express complete propositions,

- (1) Muggsy is too short/isn't tall enough
- (2) a. Kurt finished the picture
- b. Kurt finished the book draft
- c. Kurt finished the newspaper

they still have determinate semantic contents. However, these are not complete propositions. The semantics of (1) does not specify what Muggsy is too short or not tall enough for, and the semantics of the sentences in (2) do not specify whether Kurt finished painting, writing, reading or, for that matter, eating. However, what the speaker means must include some such thing. So the completion of what the speaker means involves the insertion of something that does not correspond to any constituent of the sentence. This does not show that there is something wrong with the semantic-pragmatic distinction but only that utterances of semantically incomplete sentences require pragmatic supplementation.

4. *There are many sentences whose typical use is not what, according to the compositional semantics of the sentence, the sentence means. Therefore, pragmatic information somehow blends into semantic information.*

This objection is illustrated by likely utterances of (3) and (4).

- (3) Jack and Jill went up the hill.
 (4) Jack and Jill are married.

(3) is likely to be used to assert that Jack and Jill went up the hill together and (4) that they are married to each other, even though this is not predictable from the meanings of the sentences. Nothing adverse to the semantic-pragmatic distinction follows from this, however. These examples show merely that some sentences are typically not used to mean what the sentences themselves mean. This is clear from the fact that the analogous uses are not typical for sentences like (3') and (4'),

- (3') Jack and Jill went up the hill separately/on different days.
 (4') Jack and his sister Jill are married.

5. *There are certain expressions that are generally not used strictly and literally, such as "empty," "everybody," and "circular". Therefore, their semantics does not determine how they are standardly used and pragmatics enters in.*

This is also true, but the existence of a distinction between semantics and pragmatics does not imply or even suggest that expressions must standardly be used literally. There may be a presumption of literality, but this presumption can easily be overridden, especially with words like the ones above.

6. *There are certain expressions that have a range of related meanings but are neither clearly ambiguous nor clearly unambiguous. What such an expression can be used to mean is always partly a pragmatic matter.*

This objection is based on examples like these:

- (5) a. Gus went from Natchez to New Orleans.
 b. The road went from Natchez to New Orleans.
 c. The show went from 7 to 11.
 d. Gus went from irritated to outraged.
 e. The house went from Gus to his wife.

The idea behind the objection is that as they occur in these sentences the words 'go,' 'from,' and 'to,' though semantically univocal, have distinct but related

meanings; that is, rather than being ambiguous their unitary linguistic meanings underdetermine what they are used to mean “in context,” hence that their pragmatics intrudes on their semantics.

This idea is a definite improvement over the view that the words ‘go,’ ‘from,’ and ‘to’ are used literally only in (5a), which involves movement from one place to another, and that their uses in the other sentences are in various ways “extended,” hence nonliteral uses. However, it does not follow that pragmatics intrudes on semantics. The existence of these various uses shows merely that the meanings of such polysemous terms are more abstract than the movement model would suggest (for further discussion see Bach 1994, sec. 7).

Nunberg (1979) offers a related objection, based on the multiplicity of uses of terms like ‘chair’ and ‘newspaper.’ ‘Chair’ can refer to particular chairs (chair tokens) or to chair types. But it is far from clear why this instance of the general type-token ambiguity poses a problem for the semantic-pragmatic distinction. The case of ‘newspaper’ is more interesting, because that term can refer either to particular copies of a newspaper, to specific issues or editions of a newspaper, e.g., the final edition of today’s *New York Times*, or to the publishing company. Nunberg claims that there is no basis for singling out one use as the conventional one and treating the others as derived from that. But surely the last use is a derived use, since the publishing company would not be referred to as the newspaper (e.g., the San Francisco Chronicle, which was recently bought by the Hearst Corporation). Indeed, it is arguable that this use of the term is elliptical for ‘newspaper publishing company.’ In any case, how to explain polysemy has no particular bearing on the semantic-pragmatic distinction, but is rather a problem in lexical semantics.

7. *There are certain complex expressions whose meanings are not predictable from the meanings of their constituents. Therefore, pragmatics impinges upon semantics.*

It is true that the meanings of phrases and compounds such as the following are not predictable, or at least not obviously predictable, from the meanings of their constituents:

- (6) a. sad girl
- b. sad face
- c. sad day
- d. sad music
- (7) a. child abuse
- b. drug abuse
- (8) a. election nullification
- b. jury nullification

- (9) a. slalom skiing
- b. snow skiing
- c. helicopter skiing
- (10) a. jellyfish
- b. goldfish
- c. catfish
- (11) a. clipboard
- b. diving board
- c. bread board
- d. game board
- e. headboard

As interesting as these examples are, they do not undermine the semantic-pragmatic distinction. All they show is that phrasal semantics is not straightforward. They suggest that compositionality is not as simple as it might seem, for there are different ways in which the meanings of words can combine. But this has nothing to do with pragmatics.

8. *A well-defined semantics/pragmatics distinction requires that semantics determine what is said. But intuitions about what is said indicate that it includes pragmatically determined elements.*

These “intuitions” are not necessarily about what is said (see Bach 1994 and “Seemingly Semantic Intuitions,” chapter 2 of this volume). They arise when one ignores the distinction between locutionary and illocutionary acts or assumes that the distinction between what is explicit in an utterance and what is implicated by it is exhaustive. Intuitions that are adverse to the semantic-pragmatic distinction are insensitive to the in-between category of what is implicit in what is said (more accurately, in the saying of it), and mistakenly include that in what is said. The same mistake is implicit in relevance theorists’ use of the term ‘explicature’ for aspects of utterance content that are not explicit and in the description of the entire content as “explicit.”

9. *The strict and literal semantic content of a sentence is often not calculated prior to the hearer’s determination of the content of a speaker’s utterance. Therefore, even if there is a theoretical role to the notion of semantic content, it has no psychological import, hence no empirical significance.*

Facts about “pragmatic processing” are not relevant to the semantic-pragmatic distinction. They are often cited in support of claims about what is and what is not said and even used to argue that there are pragmatic elements in what is said. However, nothing follows from such facts about what is or is not said, since that

is a matter of what a speaker does in uttering a sentence, not what his listeners do in understanding it. Moreover, the psychologically relevant category, if there is one, is information that is available to pragmatic processing, not what goes on in the processing itself.

None of the above considerations or phenomena poses a serious objection to the viability of the semantic-pragmatic distinction. What they do show is that it is important to avoid a simplistic approach to that distinction. Semantics and pragmatics are both complex, but this does not mean that they should be mixed up, much less that they overlap.

RADICAL PRAGMATICS: ANNE BEZUIDENHOUT

IN THE PREVIOUS SECTION, Kent Bach says: “To the extent that the debate about [the semantics-pragmatics distinction] is not entirely terminological (e.g., many years ago ‘pragmatics’ was the name for indexical semantics), the main substantive matter of dispute is whether there is such a thing as ‘pragmatic intrusion’, whereby pragmatic factors allegedly contribute to semantic interpretation” (p. 287). In these brief remarks I explore the idea that pragmatic considerations intrude into the domain of what is said, and so have an impact on the truth-conditional contents of sentences as uttered in particular conversational contexts. The view I have in mind belongs to the tradition of radical pragmatics. It goes against the grain of the view that most philosophers of language are inclined to accept. I will argue that more traditional views about how to draw the boundary between semantics and pragmatics by no means enjoy the advantage that many assume they do. I will also argue that the debate between traditionalists and intrusionists may be one that cannot be settled by *a priori* means, and that empirical tests of these rival views may ultimately have to decide the issue.

Radical Pragmatics and the Standard View

The term ‘radical pragmatics’ was used a fair bit in the late 1970s and early 1980s. Then it mostly meant what today we would call Gricean pragmatics. People went around wielding something called Modified Ockham’s Razor (MOR), a principle that says “Do not multiply meanings beyond necessity.” They tried to show that many phenomena that had been treated as cases of semantic ambiguity could be handled more economically by positing a more minimal semantics, and then using general conversational principles and maxims to show how further meanings could be derived.¹

There was predictably something of a backlash, with researchers arguing in favor of a radical *semantics*, and arguing that Griceans were too quick to wield MOR. Perhaps some of the disputed phenomena are after all best explained by an enriched semantics, in which many more meanings are governed by conven-

tions than allowed for by Griceans. Wierzbicka (1987), for instance, uses cross-linguistic data to argue that some phenomena that have been treated as pragmatic are in fact governed by language-specific conventions.

But that was the 1970s and 1980s. These days Gricean pragmatics no longer seems so radical, because it has become entrenched. Moreover, a certain view of the division of labor between semantics and Gricean pragmatics has become entrenched. Semantics is concerned with meanings that can be assigned to linguistic expressions independently of facts about a speaker's communicative intentions, whereas pragmatics is concerned with speaker meaning. This view (that semantics is independent of speakers' intentions) claims to be able to accommodate the fact that semantic values can be assigned to indexical expressions only relative to a context. The semantic content of an indexical may vary across contexts, but such variation is governed by a context-*independent* rule that fixes the content of the expression relative to a context. Context is here conceived of as a fixed set of elements that is needed for the semantic evaluation of indexical expressions, and hence of the sentences in which they occur. Context in this (narrow) sense may include the utterer of a sentence, the addressee, the time and place of utterance, and a few other such parameters. The important point is that knowledge of speakers' intentions is supposedly not required in order to apply the rules associated with indexical expressions.

Berg (forthcoming) offers the following characterization of the standard view of semantics (SV):

Every *disambiguated* sentence has a determinate semantic content, *relative* to an assignment of contents to its indexical expressions, and not necessarily identical to what may be conveyed (pragmatically) by its utterance.

SV limits the role of context to its role in disambiguation and reference assignment. Once these operations have been performed, relative to some context, we will arrive at what is (literally) said by the sentence in that context. Also, SV assumes that pragmatically conveyed content is distinct from semantic content even when pragmatically conveyed content coincides with semantic content (as will happen on those occasions on which a speaker means exactly what she says). Advocates of SV generally accept that Gricean pragmatics (with its appeal to conversational maxims and principles) gives an adequate account of any pragmatically conveyed content.

The Intrusion of Pragmatics into Semantics

Many competitors to Gricean pragmatics have emerged over the past decade or so. I am not going to survey all these alternatives here. Instead, I focus on just

one alternative, which Recanati (2000) calls truth-conditional pragmatics. Travis (1985), Recanati (1991), and Sperber and Wilson (1986), among others, have defended this view. According to this view:

when an utterance is made and a certain truth-conditional interpretation emerges for that utterance, it does so as a result of pragmatic processes that can be affected by any change in the wide context ('background'); it follows that *one can, by imaginatively altering the background, affect the truth-conditions of the utterance even though one leaves the narrow context [i.e., the information relevant to determining the semantic values of indexicals] untouched.*" (Recanati 2000, 5, his emphases)

The claim is that what is said (i.e., the proposition expressed) by a sentence is fixed only in some context of use, where context goes beyond "narrow" context and includes background information that a speaker intends her audience to access. This view is often supported by pointing to the phenomenon of semantic underdetermination. Some sentences are such that the information they semantically encode underdetermines the propositions they express on particular occasions of use. Semantic underdetermination must be distinguished from the phenomena of ambiguity and indexicality. Even after all ambiguities have been eliminated and all indexical elements in a sentence have been assigned a reference, the sentence may still fail to express a complete proposition, and hence may fail to be truth evaluable. To get a truth-evaluable content requires supplementation with further contextual information. The wider contextual information that is required for truth evaluability is limited only by the speaker's communicative intentions.

Relevance theorists such as Sperber and Wilson (1986) and Carston (1991, 1997), and Neo-Griceans such as Bach (1994a,b) and Recanati (1991, 1993, 1994), have done a lot to focus attention on a range of examples. Consider:

- (1) Bob and Susan are married {to one another}.
- (2) Betty has finished {washing the dishes}.
- (3) This table isn't strong enough {to support this pile of books}.
- (4) Everyone {in my department} came to my party.
- (5) You're not going to die {from this cut}.
- (6) I haven't had breakfast {today}.
- (7) This hamburger is {almost} raw.

Relevance theorists and Recanati identify what is said by the utterance of such sentences with the complete propositions they express in some (wide) context. Bach resists this identification. Instead he advocates a notion of what is said that corresponds more closely to that of sentence meaning, which has the consequence

that what is said can be something propositionally *incomplete*. So, although he has been one of the prime defenders of the notion of semantic underdetermination, he avoids the intrusion of pragmatics into what is said (although he does not avoid the intrusion of pragmatics into truth-conditional content).

Bach and relevance theorists agree that the propositions expressed by (1)–(3) are underdetermined by their semantically encoded meanings and that a pragmatic process of completion (Bach) or enrichment (relevance theorists) is needed to arrive at these propositions. (The difference is that Bach denies that these propositions correspond to what is said, and instead he calls them implicatures.) Recanati agrees that a process of completion or saturation is needed, but he argues that saturation is a matter of filling a slot that is semantically (though not syntactically) articulated.

There is also disagreement as to what is involved in examples such as (4)–(7). Bach would say there is a complete proposition that is expressed by such sentences. It is just that this proposition is not the one communicated, and hence a process of expansion is necessary to arrive at the communicated proposition. On occasion this communicated proposition will be logically weaker than the proposition expressed, as is the case in examples (5)–(7). Relevance theorists on the other hand would regard the information semantically encoded in (4)–(7) as underdetermining what is said. Enrichment processes are called for in all cases except (7), which calls for loosening. Recanati has yet a third view. He would classify (4) along with (1)–(3). Its interpretation involves filling a semantically articulated slot with a contextually supplied constituent. In the case of (5) and (6) the pragmatically supplied constituents are genuinely semantically unarticulated. Thus a pragmatic process of free enrichment is required for their interpretation.

Travis (1985, 1989) has argued at length for the context dependence of truth-conditional content. He focuses on a slightly different range of examples, and argues from a somewhat different starting point. He considers sentences such as:

- (8) This kettle is black.
- (9) Smith weighs 80 kilos.
- (10) Hugo is a sailor.

Here is a quotation in reference to example (8):

Suppose the kettle is normal aluminum, but soot covered; normal aluminum but painted; cast iron, but glowing from heat; cast iron but enamelled white on the inside; on the outside; cast iron with a lot of brown grease stains on the outside; etc. (Travis 1985, 197)

Travis's point is that there are many different circumstances under which we would be prepared to predicate blackness of the kettle, and in each of these what

we would be saying about the kettle would be different. This is not a matter of indexicality, as we can fix the reference to a particular kettle and a particular time, and yet what is said would vary if surrounding circumstances were varied. Travis points to a similar array of facts for his other examples and draws similar conclusions about the context variability of what is said by (9) and (10).

To make out the case for the variability of what is said by (8) we would have to spell out something about the circumstances in which the utterance of (8) takes place. Travis does not do this, but it is not hard to see what sort of thing would be required. Suppose that I am shopping for a kettle that will fit in with my kitchen décor, which is all black and white. My shopping companion spots a kettle on a shelf and pointing to it she utters (8). Arguably what she has said is true, even though the spout and the handle are chrome, the inside of the kettle is enameled white, and the bottom surface is copper. What counts as a black kettle in these circumstances is that its top outside surface (minus spout, handle, and possibly some trim) be painted black. On the other hand, suppose we are browsing through a thrift shop in search of a kettle and spotting a kettle on a shelf my friend utters (8). Reaching up I run my finger over the surface, revealing a red painted surface covered in fine black soot. In these circumstances what my friend said was *not* true. In yet other circumstances a kettle covered in black soot might count as a black kettle. Suppose it is my job to sort through kettles that have been donated to the thrift shop to see which ones need cleaning. I come across one covered in black soot and I utter (8). In these circumstances what I have said is true, even though underneath the soot the kettle is painted red.²

Call the view of saying that is advocated by Travis, Recanati, and relevance theorists a contextualist conception of saying. This conception contrasts with the more minimalist conception of saying defended by Griceans, by those who accept SV, and by Bach (1994a,b). Contextualists argue that because there are sentences that can be used to express different propositions in different circumstances and hence that can be true in some contexts and false in others, and because this difference in truth-value is not due to ambiguity or indexicality, the minimalist conception of what is said by such sentences must be mistaken. In order to establish the truth of the first premise of this argument contextualists frequently resort to giving examples, of the sort discussed above. In other words, they attempt to appeal to speakers' intuitions about the truth-conditions of sentences when these are uttered in particular circumstances. Berg (forthcoming) thinks that this is not a very satisfactory way to argue, as minimalists are unlikely to share contextualists' intuitions.

For instance, minimalists might suggest that what is strictly and literally expressed by (8) is something that could be paraphrased as:

- (11) The kettle is black in some way.

So the minimalist is likely to say that when my friend points to the red, soot-covered kettle and utters (8), what she strictly and literally says is *true*. On the other hand, what she says is misleading, because it conversationally implies something false about the kettle, namely that its top outside surface (minus spout, handle, and possibly some trim) is painted black. Contextualists are likely to respond that (8) and (11) do not say the same thing, as we can imagine circumstances in which the truth-conditions for (8) and (11) come apart. Suppose circumstances are as described above, viz., I am shopping for a kettle to go with my black-and-white kitchen décor. My friend points to a red kettle with a tiny black spec on it and utters (8). What she has said is surely false. However, if she had uttered (11) she would have said something true.

Clearly we are not making much progress in this debate at this point, but are simply trading intuitions back and forth. Is there any way to break this deadlock? Berg (forthcoming) suggests we can. He offers an argument that is a variation on an argument given by Kripke (1977). Berg asks us to imagine a language as much like English as possible, except that it is stipulated that (8), with references, times, etc., fixed, expresses the proposition *The kettle is black in some way*. Call this language Schmenglish. The next step is to ask whether, in a community that speaks Schmenglish, it would be possible to convey the full range of meanings that Travis claims can be expressed by means of (8). The answer is clearly 'Yes'. Even though (8) always expresses the same proposition in this language, all the other more specific interpretations that Travis claims are possible could be conversationally implicated. So the mere fact that we can imagine indefinitely many interpretations for (8) does nothing to impugn the claim that English itself has a minimalist semantics of the sort Schmenglish does.

But Berg's argument does not get us very much further in the debate between minimalists and contextualists about what is said. Let us concede that in Schmenglish it would be possible to communicate (indirectly) the same range of meanings that Travis argues can be directly expressed in English. From the point of view of the contextualist, all this establishes is that there are alternative ways of achieving the same communicative goal. It does not show that this is the way English speakers achieve their communicative goals. The unfortunate Schmenglish speakers are constrained by their language to achieve indirectly what we can achieve directly.³

Even if Berg's argument has not broken the deadlock between contextualists and minimalists, it does at least show how it could be broken. Unfortunately, it will not be broken by philosophical considerations. What might move us forward is an empirical psychological investigation of the pragmatic processes involved in language production and comprehension. Do minimal propositions like (11) play a role in the production and comprehension of utterances? If they do, this would be evidence in favor of the minimalist conception of saying.

Some empirical work on this topic has already been carried out. Bezuidenhout and Cutting (forthcoming) contains a summary of this work.

Breaking the Deadlock between Contextualists and Minimalists

In this section I consider three ways that might be suggested for breaking this deadlock by *a priori* means. First, some readers might have been thinking for a while now that there is an easy way to decide the issue between the contextualist and minimalist conceptions of saying. After all, Grice proposed various tests to see whether some communicated content was a conversational implicature or not, namely, the tests of cancellability, nondetachability, and calculability. Why can't these tests be applied to see whether or not the contents that Travis says are directly expressed in particular circumstances by sentences such as (8) are implicatures? However, as Travis (1985) shows, this is not possible to do without begging the question against the contextualist. All the tests presuppose a conception of what is said, and hence cannot be applied until we have settled the debate between contextualists and minimalists.

A second way of attempting to break the deadlock between minimalists and contextualists is to adopt the sort of "radical semantic" approach taken by Stanley and Szabo (2000). This approach would accept the intuitions of people like Travis as to the truth-conditional content of utterances such as (8), and would accept that this content varies from context to context. But it would deny that this variability arises because of semantic underdetermination. Rather, such variability would be assimilated to ordinary cases of indexicality. This proposal has been worked out in most detail for cases of quantifier domain restriction, such as is illustrated in (4) above. But Stanley (2000b) extends the proposal to cover other sorts of cases.

Note that as I am representing the contextualist position, what is said by the utterance of a sentence in an appropriate context is identified with the proposition it expresses, which in turn is identified with its truth-conditional content. Thus a third suggestion for breaking the deadlock is to argue that contextualists are conceptually confused about the notion of saying. These identifications involve a conflation of notions that ought to be kept distinct. Careful conceptual analysis will reveal the error. There are at least two ways of fleshing out this third suggestion.

One way is to argue for a conception of what is said that severs the connection between what is said by the utterance of a sentence in some context and the proposition it expresses.⁴ This way of fleshing out the third suggestion is adopted by Bach (1994a,b, 1999, 2001a). For Bach, what is said can be something less than propositional, and hence something that is not truth-evaluable. A second way of fleshing out the suggestion that contextualists are conceptually confused is to argue that we need to distinguish what is said by the utterance of a sentence in some context from its truth-conditional content. Cappelen and

Lepore (1997) adopt this position. They argue that our conception of what is said should line up with our practices of reporting speech. In reporting speech we are often sensitive to features of utterances other than their semantic content. This leads Cappelen and Lepore to defend a conception of what is said that incorporates many contextual and pragmatic elements. However, the task of characterizing what is said is severed from the task of giving the semantic (i.e., truth-conditional) content of utterances. With respect to this latter issue Cappelen and Lepore adopt the pure, pragmatically uncontaminated semantics accepted by advocates of SV (the standard view).

Although some of these suggestions are promising, the contextualist has resources to counter them. However, here I am unable to properly assess all these suggestions and will be able to focus just on Bach's arguments against the contextualist. Even so my responses will be sketchy.

Rejoinder to Bach

Bach (2001a) defends a syntactically constrained notion of what is said. What is said by a sentence is a projection of the syntax of the sentence as used in a given context, in the sense that the constituents of what is said correspond to elements of the sentence. This notion of what is said allows for indexicality and ambiguity, but excludes any content that does not correspond to some syntactic slot. However, Bach also believes that some sentences, such as (1)–(3) above, are semantically incomplete or underdetermined, in the sense that they do not express complete propositions. When used in context a pragmatic process of completion must operate in order to supply the missing propositional constituents that do not correspond to any syntactic elements. Thus it follows that what is said, which excludes such pragmatically determined aspects of content, can sometimes be propositionally incomplete, and hence something that is not truth-evaluable.

Contextualists assume that what is said is always something complete, even though the sentential vehicles used to say something may be semantically incomplete. In other words, contextualists have intuitions about the completeness of what is said that conflict with Bach's theoretical commitments.⁵ But Bach thinks we do not need to trade on intuitions to settle this issue. He claims to have a test to show that what is said can be incomplete. He calls this the Indirect Quotation (IQ) test. Take a sentence such as (2) above. A speaker of (2) can be indirectly quoted as follows:

(2_{IQ}) The speaker said that Betty has finished.

Bach takes this as evidence that what is said can be incomplete [i.e., can correspond to what Bach (1994a) calls a propositional radical]. He appears to assume that because the embedded sentence is semantically incomplete and (2_{IQ}) is

truth-evaluable, what is said can be incomplete. By an analogous argument we would be able to show that what people believe can be incomplete. We can report what the speaker of (2) believes as follows:

(2_{BR}) The speaker believes that Betty is finished.

However, it seems wrong to conclude that the speaker's belief contents can be incomplete. Rather, it seems more plausible both with respect to (2_{IQ}) and to (2_{BR}) to say that these sentences are themselves semantically underdetermined, because of the fact that their embedded sentences are. Hence what is expressed by such reports must be arrived at via a pragmatic process of completion or enrichment (or even in some cases by a pragmatic process of loosening).

Thus it is by no means clear that Bach's refusal to identify what is said with the proposition expressed in cases of semantic underdetermination leads to a conceptually more coherent position. Here again it looks as though *a priori* considerations are not going to break the deadlock between contextualists and minimalists.

Bach and contextualists also disagree on how widespread the phenomenon of semantic underdetermination is. There are sentences, such as (4)–(7) above, which contextualists claim are semantically underdetermined but which Bach would say express complete propositions. These propositions, which Recanati (1991) calls minimal propositions, are Bach's candidates for what is said by such sentences. What contextualists identify as what is said Bach prefers to call an implicature (i.e., something that is implicit is what is said). He believes that contextualists have failed to make fine-grained enough distinctions between various sorts of contents. For example, Bach claims that what is said by (6) ('I haven't had breakfast') is the minimal proposition that the speaker has never had breakfast. However, what is directly communicated is the proposition that the speaker has not had breakfast recently. This is an implicature, and not part of what is said, as contextualists would maintain.

Contextualists have argued against Bach's minimalist conception of what is said on the grounds that minimal propositions will not be accessed at any stage in the psychological processes of utterance production and comprehension, unless the context directly supports such an interpretation. Speaking from the hearer's point of view, if the context in which (6) is uttered is biased toward a pragmatically enriched interpretation, the comprehension process will not be mediated via grasp of the minimal interpretation that the speaker has never had breakfast. Similarly, from the speaker's point of view, when she decides how to encode her message, she will not set out to point her audience obliquely to her intended meaning via consideration of the thought that she has never had breakfast.

Bach suggests that psychological considerations are irrelevant to the discussion about what is said. He says that nothing about what is said follows from

facts about pragmatic processing, since as he notes in the previous section, what is said is “a matter of what a speaker does in uttering a sentence, not what his listeners do in understanding it” (p. 292). Elsewhere he complains: “It is a mystery to me why facts about what the hearer does in order to understand what the speaker says should be relevant to what the speaker says in the first place” (Bach 2001a, 25). These complaints appear to rest on the mistaken belief that the contextualist’s empirical account of pragmatic processing is solely an account about what a hearer must do to understand a speaker. But as I have tried to indicate above, the full empirical account offered by the contextualist would include a story both about comprehension and about production processes. If the minimal notion of what is said is bypassed both by speakers and hearers, this is surely relevant to an account of what is said.

Bach (1995) has an explanation as to why minimal propositions might sometimes be bypassed. He claims that in cases of sentences such as ‘I haven’t had breakfast,’ a process of standardization takes the listener directly from sentence meaning to the implicature, bypassing the minimal proposition. The inference from the minimal proposition (namely, that the speaker has never had breakfast) to the implicature has become “compressed by precedent”. Bach claims that his minimalist notion of what is said does not commit him to an account of the “temporal order or other details of the process of understanding” (Bach 2001a, 25). His account merely identifies the sort of information that is available to the performance system. Minimal propositions may be available even if they are not actually computed in the course of processing. They play a merely dispositional role, “waiting to be taken into account when there is special reason to do so.”

However, the contextualist can agree that minimal propositions are dispositionally available, and that a hearer will retrieve a minimal proposition when there is special reason to do so. For instance, if the speaker of (6) is someone who has lived his whole life apart from human civilization, say raised by wolves, then the hearer may retrieve the minimal proposition that the speaker has never had breakfast. Thus Bach would need to say more about his account in order to make it empirically distinguishable from the contextualist account. There must be some prediction it makes about how minimal propositions are processed that differs from the predictions made by contextualists, or else the distinctions Bach makes between propositional radicals, minimal propositions, and implicatures will be empirically empty. So, far from processing considerations being irrelevant, they may be crucial in giving empirical content to the various rival accounts. In this section I have argued that the debate between contextualists and minimalists is currently deadlocked, and that as long as it remains at the level of an appeal to intuitions about truth-conditions it will remain deadlocked. I speculated that progress might be made if the rival views could be subjected to experimental test. I then suggested three ways of moving the debate forward via

a priori means. I was not able to assess all these suggestions here. I have merely tried to deflect some of the considerations Bach (1994a,b, 1999, 2001a) has raised against contextualists.

**SEMANTICS AND WHAT IS SAID: REPLY TO BEZUIDENHOUT:
KENT BACH**

IN “SEMANTIC, PRAGMATIC” I endorse a certain conception of the semantic-pragmatic distinction and answer a number of possible objections. I do not dwell on the ongoing debate about what is said, but this is Anne Bezuidenhout’s focus. This debate concerns whether what is said is purely semantic, and clearly what one thinks about that depends in part on how one contrasts semantics with pragmatics.

Consider Jonathan Berg’s (forthcoming) formulation of the “standard view of semantics,” which she quotes:

- (SV) Every disambiguated sentence has a determinate semantic content, relative to an assignment of contents to its indexical expressions, and not necessarily identical to what may be conveyed (pragmatically) by its utterance.

Notice that SV does not require the semantic content of a sentence (relative to a context) to be a complete proposition, and that it does not indicate whether or not this content is to be identified with what is said by the speaker (in that context). I accept this identification (Bach 2001a); Bezuidenhout rejects it. Also, she suggests that “advocates of SV generally accept that Gricean pragmatics (with its appeal to conversational maxims and principles) gives an adequate account of any pragmatically conveyed content.” However, beyond her parenthetical allusion to maxims and principles [what Bach and Harnish (1979) call “conversational presumptions”], she gives no indication what she takes Gricean pragmatics to be. I suspect it is something like how the relevance theorists Deirdre Wilson and Robyn Carston characterized it at a recent pragmatics workshop in Oxford (Sept 29–Oct 1, 2000). Wilson and Carston imagine that according to Gricean pragmatics literal utterances involve only “encoding” and “decoding,” and that conversational maxims and inference do not enter in. They do not appreciate that even when a speaker means exactly what she says (when pragmatically conveyed content coincides with semantic content), the fact that the speaker means exactly this still depends on the speaker’s communicative intention and is still something for the hearer to rely on the maxims to infer.

There is a deeper bone of contention here. The underlying issue is whether the semantic content of a sentence, relative to a (narrow) context, is a complete proposition or whether pragmatic supplementation is generally required. But

this is a distinct issue, independent of SV, since SV does not require that the semantic contents of sentences be propositions. So when Bezuidenhout proceeds to focus on “truth-conditional pragmatics,” it is not clear what she takes the conflict to be between this and “Gricean pragmatics” or SV. She calls it “radical pragmatics,” but it is not clear what is radical about it beyond the fact that it supposes that sentences generally do not (relative to narrow contexts) express complete propositions. This supposition alone does not suggest that pragmatics intrudes into semantics.

Why does Bezuidenhout suppose that it does? Perhaps because she assumes that the semantics of sentences invariably assigns complete propositions to them. Perhaps it is because of how she characterizes semantic underdetermination: “Some sentences are such that the information they semantically encode underdetermines *the propositions they express* on particular occasions of use” (this volume, p. 294; my italics). But if they are semantically underdeterminate, they do not express (complete) propositions at all. Another explanation may be that she sometimes speaks of the contents of, and the propositions expressed by, utterances of sentences as opposed to sentences themselves. One can agree that in cases of semantic underdetermination “to get a truth-evaluable content requires supplementation with further contextual information . . . [which is] limited only by the speaker’s communicative intentions,” without supposing that the sentence itself has a truth-evaluable (i.e., propositional) content. A speaker’s communicative intention cannot affect the semantics of a sentence and is relevant not to the locutionary act (whose content is what is said) but only to the illocutionary act the speaker is performing. Moreover, if it is the content of an *utterance* that is in question, that content could be patently nonliteral, e.g., with a likely utterance of “You are a zombie,” in which case it is irrelevant to semantics. If we are concerned with semantics, there is no point in talking about contents of utterances unless we are prepared to enforce a distinction between literal and nonliteral content. Bezuidenhout’s discussion of utterances and their contents does not take this distinction into account.

She rightly notes that I “advocate a notion of what is said that corresponds closely to that of sentence meaning [and structure], which has the consequence that what is said can be something propositionally incomplete” (this volume, pp. 294–95). She seems to find this idea a bit idiosyncratic. But to me it is downright eccentric to suppose that any element of what is said in uttering a sentence can fail to correspond to some constituent of the sentence. After all, we’re talking about what is *said* in uttering the sentence, not about what might be conveyed in uttering it. I take the semantics of a sentence to be a projection of its syntax. That is, semantic structure is interpreted syntactic structure. The content of a sentence is determined compositionally as a function of the contents of its constituents and their syntactic relations. It just so happens that some syn-

tactically well-formed sentences do not express complete propositions. So if what is conveyed in their utterance must be a complete proposition, such sentences must be used to convey more than is predictable from their compositionally determined contents.

Consider Bezuidenhout's sentences (1)–(3):

- (1) Bob and Susan are married {to one another}.
- (2) Betty has finished {washing the dishes}.
- (3) This table isn't strong enough {to support this pile of books}.

Contrary to what she suggests, I do not take (1) to be semantically incomplete. It expresses the proposition that Bob and Susan are married, without specifying to whom. It may be implicit in the speaker's uttering (1) that they are married to each other, but that is not part of the sentence's content. As for (2) and (3), I am now inclined, contrary to what I said in Bach 1994, to suppose that argument slots complementing 'finish' and 'enough' are lexically mandated. If that is correct then, as Stanley (2000a) has argued, such sentences are covertly indexical. (I do not agree with Stanley's contention that *all* apparent cases of semantic underdetermination are really cases of covert indexicality.) Even so, I would still maintain that in uttering (2) the speaker did not *say* what Betty finished and that in uttering (3) the speaker did not *say* what the table was not strong enough for.

Regarding sentences (4)–(7),

- (4) Everyone {in my department} came to my party.
- (5) You're not going to die {from this cut}.
- (6) I haven't had breakfast {today}.
- (7) This hamburger is {almost} raw.

I do say that each of these sentences expresses a complete proposition (without the parenthetical qualification), one that is distinct from the one being communicated, so that a process of expansion is necessary to arrive at the communicated proposition. Relevance theorists would indeed regard what is semantically encoded in these sentences as underdetermining what is said in their utterance but, as far as I can tell, that is only because they have a loose conception of what is said, such that it can include what is not said. They do not put it that way, of course, yet for some reason they regard what is implicit in the making of an utterance to be part of its "explicit content."

Bezuidenhout reports that Travis has "argued at length for the context-dependence of truth-conditional content." But truth-conditional content of what? In the case of (8),

- (8) This kettle is black

“there are many different circumstances under which we would be prepared to predicate blackness of the kettle, and in each of these what we would be saying about the kettle would be different. . . . what is said would vary if surrounding circumstances were varied” (this volume, pp. 295–96). This is not quite right. The circumstances do not affect what is said. Rather, they affect what the speaker could reasonably be taken, and reasonably intend to be taken, to convey in uttering the sentence (see Berg 1993 and Bach 2001a). More importantly, the fact that people can mean different things in uttering (8) does not support a contextualist conception of saying. It just shows that there are different ways of being black [note the distinction in Harvey (2000) between “present color” and “official color”] and that there are ways of using (8) nonliterally, e.g., to mean that the kettle is predominantly black, that the main visible part of the kettle is black, or that the kettle is covered with black stuff. Moreover, even if the minimalist agreed with Travis and Bezuidenhout about the data, he would regard them as showing that (8) is semantically underdeterminate, hence that utterances of it require completion. The minimalist certainly does not have to concede that (8) should be paraphrased as “The kettle is black in some way” (see Bach 1994, 130–31, for a discussion of analogous cases). Clearly examples like (8) are very interesting, but the issue they raise is orthogonal to the dispute between minimalists and contextualists. It concerns the extent to which natural language sentences are semantically underdeterminate, not what theoretical stance to take toward the existence of such sentences. That (8) has various uses does not show that what is *said* in uttering it is affected by anything pragmatic, but this does suggest that its semantics is not as simple as it might seem.

I agree with Bezuidenhout that neither Grice’s cancellability test nor my IQ test, never mind intuitions (see “Seemingly Semantic Intuitions,” this volume), will break the deadlock between contextualists and minimalists. She suggests that matters might be resolved by an empirical psychological investigation of the pragmatic processes involved in language production and comprehension. This assumes, however, that minimalism requires that minimal propositions, such as those expressed by sentences like (4)–(7), play a role in the production and comprehension of utterances. Her contextualist case against a purely semantic conception of what is said rests not only on the empirical claim that minimal propositions (those correlated with sentence syntax) are not “accessed at any stage in the psychological processes of utterance production and comprehension, unless the context directly supports such an interpretation” (this volume, p. 300), but also on an implicit assumption. I agree with her about the importance of the production side, but in claiming that what is said is “a matter of what a speaker does in uttering a sentence, not what his listeners do in understanding it,” I was not harboring the “the mistaken belief that the contextualist’s empirical account of pragmatic processing is solely an account about what a hearer must do to understand a speaker.” I was merely responding to the many contextualists who

draw conclusions about what is said from (alleged) facts about comprehension processes. Indeed, contrary to what she suggests, I do not “suggest that psychological considerations are irrelevant to the discussion about what is said.” Rather, I reject her implicit assumption that psychological considerations about *processing* are relevant. Here I rely on the distinction between cognitive processes and information available to them (see Bach and Harnish 1979, 91–93, and Peacocke 1986). Insofar as communicative inference involves standardization of use (Bach 1995) and default reasoning (Bach 1984), both of which come in degrees, cognitive processes that implement such inference can be sensitive to certain sorts of information without actually computing it. This is why facts about processing itself are not decisive.

Moreover, Bezuidenhout’s appeal to processing considerations is in danger of proving too much. It would show that the conveyed content of a clearly non-literal utterance is the sentence’s semantic content. If I said to my wife, “Since you are the rudder of my life, you won’t steer me wrong,” she would probably take the word ‘rudder’ nonliterally well before she processed the sentence completely, in which case she would not compute the proposition expressed (literally) by the sentence. Should we thus conclude that the proposition I conveyed is a meaning of the sentence? With many nonliteral utterances hearers can figure out what a speaker is communicating without first identifying what the speaker is saying. Fortunately, adopting the semantic notion of what is said does not commit one to an account of the temporal order or other details of the cognitive processing involved in production or comprehension.

Contextualism, and the so-called radical pragmatics that goes with it, is based on some very interesting linguistic data, and these deserve detailed consideration (some of them are taken up in Berg (1993) and in Bach (2001a), as well as by contextualists). However, contextualism is not supported by any clear conception of semantics or of pragmatics. It needs to supplement the observation that a sentence like “The kettle is black” can be used in various ways with an account of how the meaning of such a sentence is built up from the meanings of its parts. Otherwise, contextualism would leave what is involved in understanding such sentences, and how they can be used to communicate, something of a mystery. In my view, though, once certain distinctions are taken into account, such as those between sentences and utterances, between linguistic contents and psychological contents, between locutionary and illocutionary acts, and between information available to cognitive processes and the processes themselves, the rationale for contextualism is lost, although the linguistic data it highlights remain to be explained.

NOTES (BEZUIDENHOUT)

1. Another view defended under the rubric of radical pragmatics is the view espoused by Nunberg (1979). In this paper he defines radical pragmatics as the doctrine according to which “the semantics/pragmatics distinction cannot be validated even in principle: there is

no way to determine which regularities in use are conventional and which are not.” The “regularities in use” that Nunberg focuses on are cases of polysemy, such as:

The newspaper {publication} is on the table.
 The newspaper {publisher} fired John.
 The chair {token} is broken.
 The chair {type} was common in 19th century parlors.

Nunberg asks whether each of these uses is governed by a separate convention. He cites cross-linguistic evidence, evidence from deferred ostension, and syntactic evidence to support the claim that lexical conventions should not be multiplied. He concludes that words like ‘newspaper’ have only one conventional use with the other use generated pragmatically. Now the question arises as to which of these uses is the conventional one and which is the derived one. Nunberg argues that there is no principled way of deciding this issue.

2. This discussion of Travis’ views might have reawakened memories of Searle’s (1980, 1983, 1985) notion of the background against which we interpret sentences like:

(*) The cat is on the mat.

Searle’s point is similar to Travis’, at least as I interpret him. (Berg 1993 interprets Searle’s views differently.) The claim is that truth-conditions can be assigned to a sentence only relative to a background of assumptions. Moreover, this background cannot ever be completely and explicitly spelled out. In particular, the normal truth-conditions that we would assign to (*) are only relative to certain assumptions about the normal relations of cats to mats. We can always imagine circumstances that are special in certain ways, and in which the conditions for the truth of (*) would be different from the usual ones.

3. One might challenge the distinction I am assuming here between directly saying something and indirectly implicating that thing. Bach (1995) insists on a threefold distinction, between what is said, what is directly communicated and what is indirectly communicated. Something other than what is strictly and literally said by a sentence can sometimes be directly communicated by that sentence. A process of standardization can make this nonliteral interpretation more accessible. What is strictly and literally said will be bypassed in the interpretive process, because precedent has compressed the inference from what is said to the nonliteral interpretation. Something like Bach’s view might be advocated for the cases discussed by Travis, such as (8) above. The claim would be that although what (8) strictly and literally says is something like (11), this minimal proposition will be bypassed in the inferential process that results in the recovery of some appropriate enriched interpretation. This is a matter of standardization—the inference from (11) to any given enriched interpretation has become compressed by precedent. The enriched interpretation is now directly communicated, and this may create the illusion that it is directly expressed. The difficulty with this suggestion is that some of these enriched interpretations are very specific, and I might never have encountered a situation before in which that was the correct enriched interpretation. Nevertheless, my intuition is that I would understand this very specific enriched interpretation to have been directly expressed. Precedent cannot explain this intuition, since there is no precedent in this and indefinitely many other specific cases.
4. Bach disapproves of saying that utterances have propositional content. For him it is either sentences or the communicative intentions of speakers that have such content. But if an utterance is something produced with a certain communicative intention, then it seems harmless enough to talk of the utterance itself having content. It has the same content as the communicative intention with which it is produced.
5. One might worry that in claiming that what is said must be fully propositional the contextualist has conflated saying with stating. However, there is no reason why the contextualist should not be able to honor Austin’s distinction between locutionary and illocutionary acts. Saying is the performance of a locutionary act, which Austin defined as the uttering of certain words with a certain sense and reference. Austin was clearly assuming that disambiguation and reference assignment are sufficient for expressing a

complete proposition. In light of the recognition of the phenomenon of semantic underdetermination, the contextualist will need to revise this slightly. A locutionary act is the uttering of certain words with a certain sense and reference and pragmatically enriched/loosened in a certain way.

REFERENCES

- Bach, K. 2001a. "You Don't Say?" *Synthese* 128: 15–44.
- . 2001b. "Seemingly Semantic Intuitions." In J. Keim Campbell, M. O'Rourke, and D. Shier (eds.), *Meaning and Truth: Investigations in Philosophical Semantics*. New York: Seven Bridges Press.
- . 1999a. "The Semantics-Pragmatics Distinction: What It Is and Why It Matters." In K. Turner (ed.), *The Semantics-Pragmatics Interface from Different Points of View*. Oxford: Elsevier.
- . 1999b. "The Myth of Conventional Implicature." *Linguistics and Philosophy* 22: 327–66.
- . 1995. "Standardization vs. Conventionalization." *Linguistics and Philosophy* 18: 677–86.
- . 1994a. "Conversational Implicature." *Mind & Language* 9: 124–62.
- . 1994b. "Semantic Slack: What Is Said and More." In S.L. Tsohatzidis (ed.), *Foundations of Speech Act Theory: Philosophical and Linguistic Perspectives*. London: Routledge.
- . 1984. "Default Reasoning: Jumping to Conclusions and Knowing When to Think Twice." *Pacific Philosophical Quarterly* 65, 37–58.
- Bach, K., and R. Harnish. 1979. *Linguistic Communication and Speech Acts*. Cambridge, Mass.: MIT Press.
- Berg, J. forthcoming. "Is Semantics Still Possible?" *Journal of Pragmatics*.
- . 1993. "Literal Meaning and Context." *Iyyun, The Jerusalem Philosophical Quarterly* 42: 397–411.
- Bezuidenhout, A. 2000. "Attitude Ascriptions, Context and Interpretive Resemblance." In K.M. Jaszczolt (ed.), *The Pragmatics of Propositional Attitude Reports*. Oxford: Elsevier.
- Bezuidenhout, A.L., and J.C. Cutting (forthcoming). "Literal Meaning, Minimal Propositions and Pragmatic Processing." *Journal of Pragmatics*.
- Cappelen, H., and E. Lepore 1997. "On an Alleged Connection Between Indirect Speech and the Theory of Meaning." *Mind & Language* 12: 278–96.
- Carston, R. 1997. "Enrichment and Loosening: Complementary Processes in Deriving the Proposition Expressed?" *Linguistische Berichte*: 103–27.
- . 1991. "Implicature, Explicature, and Truth-theoretic Semantics." In S. Davis (ed.), *Pragmatics: A Reader*. New York: Oxford University Press.
- Grice, P. 1989. *Studies in the Way of Words*. Cambridge, Mass.: Harvard University Press.
- Harvey, J. 2000. "Colour-Dispositionalism and Its Recent Critics." *Philosophy and Phenomenological Research* 61: 137–55.
- Kripke, S. 1977. "Speaker's Reference and Semantic Reference." *Midwest Studies in Philosophy* 2: 255–76.
- Nunberg, G. 1979. "The Non-Uniqueness of Semantic Solutions: Polysemy." *Linguistics and Philosophy* 3: 143–84.
- Peacocke, C. 1986. "Explanation in Computation Psychology: Language, Perception and Level 1.5." *Mind & Language* 1: 101–23.
- Recanati, F. 2000. "Unarticulated Constituents." Ms. available at www.phon.ucl.ac.uk/home/robyn/workshop/papers/recanati.htm
- . 1994. "Contextualism and Anti-Contextualism in the Philosophy of Language." In S.L. Tsohatzidis (ed.), *Foundations of Speech Act Theory: Philosophical and Linguistic Perspectives*. London: Routledge.
- . 1993. *Direct Reference: From Language to Thought*. Oxford: Blackwell.
- . 1991. "The Pragmatics of What Is Said." In S. Davis (ed.), *Pragmatics: A Reader*. New York: Oxford University Press.
- Searle, J.R. 1985. *Expression and Meaning: Studies in the Theory of Speech Acts*. Cambridge: Cambridge University Press.
- . 1983. *Intentionality*. Cambridge: Cambridge University Press.

- . 1980. "The Background of Meaning." In J.R. Searle, F. Kiefer and M. Bierwisch (eds.), *Speech Act Theory and Pragmatics*. Dordrecht: Kluwer Academic Publishers.
- Sperber, D., and D. Wilson. 1986. *Relevance: Communication and Cognition*. Oxford: Blackwell.
- Stanley, Jason. 2000a. "Context and Logical Form." *Linguistics and Philosophy* 23: 391–434.
- Stanley, J. 2000b. "Nominal Restriction." Ms.
- Stanley, J., and Z.G. Szabo. 2000. "On Quantifier Domain Restriction." *Mind & Language* 15: 219–61.
- Travis, C. 1989. *The Uses of Sense: Wittgenstein's Philosophy of Language*. Oxford: Clarendon Press.
- . 1985. "On What Is Strictly Speaking True." *Canadian Journal of Philosophy* 15: 187–229.
- Wierzbicka, A. 1987. "Boys Will Be Boys: 'Radical Semantics' vs. 'Radical Pragmatics'." *Language* 63: 95–114.

