

Christopher Viger
Robert J. Stainton
Editors

Studies in Linguistics and Philosophy

85

Compositionality, Context and Semantic Values

Essays in Honour of Ernie Lepore



Springer

COMPOSITIONALITY, CONTEXT AND SEMANTIC VALUES

Studies in Linguistics and Philosophy

Volume 85

Managing Editors

KAI VON FINTEL, *M.I.T., Cambridge*
LISA MATTHEWSON, *University of British Columbia*

Editorial Board

JOHAN VAN BENTHEM, *University of Amsterdam*
GREGORY N. CARLSON, *University of Rochester*
DAVID DOWTY, *Ohio State University, Columbus*
GERALD GAZDAR, *University of Sussex, Brighton*
IRENE HEIM, *M.I.T., Cambridge*
EWAN KLEIN, *University of Edinburgh*
BILL LADUSAW, *University of California at Santa Cruz*
TERRENCE PARSONS, *University of California, Irvine*

For further volumes:
www.springer.com/series/6556

COMPOSITIONALITY, CONTEXT AND SEMANTIC VALUES

Essays in Honour of Ernie Lepore

edited by

ROBERT J. STANTON

University of Western Ontario, London, ON, Canada

and

CHRISTOPHER VIGER

University of Western Ontario, London, ON, Canada

 Springer

Editors

Robert J. Stainton
Department of Philosophy
University of Western Ontario
Talbot College, London
Ontario, N6A 3K7, Canada
rstainto@uwo.ca

Christopher Viger
Department of Philosophy
University of Western Ontario
Talbot College, London
Ontario, N6A 3K7, Canada
cviger@uwo.ca

ISBN: 978-1-4020-8309-9

e-ISBN: 978-1-4020-8310-5

DOI 10.1007/978-1-4020-8310-5

Library of Congress Control Number: 2008927518

© Springer Science+Business Media B.V. 2009

No part of this work may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, microfilming, recording or otherwise, without written permission from the Publisher, with the exception of any material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work.

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

Contents

Part I Compositionality

- “If”, “Unless”, and Quantification** 3
Sarah-Jane Leslie
- Bridging the Paratactic Gap** 31
Daniel Blair

Part II Context and “What Is Said”

- On the Epistemic Utility of *What is Said*** 61
Sanford C. Goldberg
- In Defense of Context Shifting Arguments** 79
Lenny Clapp
- Contextualism, Skepticism and Objectivity** 105
David Hunter
- On Failing to Capture Some (or Even All) of What is Communicated** 129
Kent Johnson

Part III Semantic Values

- The Disunity of Truth** 147
Josh Dever
- Descriptions, Negation, and Focus** 193
Michael Glanzberg
- Evidentials: Some Preliminary Distinctions** 221
James Higginbotham

The Direct Expression of Metaphorical Content 237
Marga Reimer

The Empirical Case for Bare Demonstratives in Vision 255
Zenon Pylyshyn

Index 275

Contributors

Daniel Blair, Ph.D.

810 Whittier PL NW, Washington DC 20012, USA, danb124@gmail.com

Lenny Clapp, Ph.D.

Investigador, Instituto Investigaciones Filosóficas, Circuito Mario de la Cueva S/N, Zona Cultural, Ciudad Universitaria, La UNAM, CP 04510, Col. Coyoacán, Del. Coyoacán, El D.F., Mexico, lclapp@minerva.filosoficas.unam.mx

Josh Dever, Ph.D.

Department of Philosophy, University of Texas at Austin, 1 University Station C3500, Austin, TX 78712, USA, dever@mail.utexas.edu

Michael Glanzberg, Ph.D.

Department of Philosophy, University of California, Davis, One Shields Avenue, Davis, CA 95616, USA, mglanzberg@ucdavis.edu

Sanford Goldberg, Ph.D.

Department of Philosophy, Crowe 3-179, 1880 Campus Drive, Northwestern University, Evanston, IL 60208-2214, USA, s-goldberg@northwestern.edu

James Higginbotham, Ph.D.

School of Philosophy, University of Southern California, 3709 Trousdale Parkway, Los Angeles, CA 90089-0451, USA, higggy@usc.edu

David Hunter, Ph.D.

Philosophy Department, Ryerson University, 350 Victoria Street, Toronto, ON, M5B 2K3, Canada, dhunter@philosophy.ryerson.ca

Kent Johnson, Ph.D.

Department of Logic and Philosophy of Science, University of California, Irvine, 3151 Social Science Plaza A, Irvine, CA 92697-5100, USA, johnsonk@uci.edu

Sarah-Jane Leslie, Ph.D.

Department of Philosophy, 1879 Hall, Princeton, NJ 08544, USA,
sjleslie@princeton.edu

Brian P. McLaughlin, Ph.D.

Department of Philosophy, Rutgers University, 26 Nichol Avenue, New Brunswick,
NJ 08901, USA, brianmc@rci.rutgers.edu

Zenon Pylyshyn, Ph.D.

Rutgers Center for Cognitive Science, Rutgers, the State University of New Jersey,
New Brunswick, NJ 08854, USA, zenon@rucss.rutgers.edu

Marga Reimer, Ph.D.

Department of Philosophy, University of Arizona, Tucson, Arizona 85721-0027,
USA, reimer@u.arizona.edu

Robert J. Stainton, Ph.D.

Department of Philosophy, University of Western Ontario, Talbot College, London,
Ontario, N6A 3K7, Canada, rstainto@uwo.ca

Christopher Viger, Ph.D.

Department of Philosophy, University of Western Ontario, Talbot College, London,
Ontario, N6A 3K7, Canada, cviger@uwo.ca

Ernie

I met Ernie in 1965 on the wrestling mats of our high school in North Bergen, New Jersey, a township on top of the plateau overlooking Hoboken and across the Hudson River from Manhattan. Hoboken then was still the Hoboken of Elia Kazan's "On the Waterfront" (1954).¹ Even though the Hudson was less than a mile across at that point, it was a wide spiritual divide. We were Jersey boys, not New Yorkers.

Ernie was as ambitious as I was about wrestling, and, so, after the season was over, we used to take a bus to Journal Square in Jersey City, and then walk about eight city blocks to a gym to lift weights. In those days, high schools didn't have weight rooms; and gyms were scarce, men only, quite filthy, and entirely devoid of cardio equipment and Nautilus machines. They were all sweat, grunts, groans, and clanking iron. By 1968, on Tuesdays and Thursdays, after a grueling wrestling practice at the high school, we would take a bus to New York City (it took about a half hour to get into "the City" by bus, less if the Lincoln Tunnel was not crowded), and then a short subway ride up to the New York Athletic Club on 59th street, across from Central Park, to spend a couple of hours working out with former university wrestling stars—guys in their mid-twenties from places like Oklahoma, Nebraska, and Iowa—who were training to make the Olympic team. Even with all of this wrestling time, we were frustrated by the fact that there was nowhere to work out on Sundays. We investigated and found out that the Jersey City YMCA had a wrestling mat and was open on Sundays. We then spent our Sunday afternoons working out there, so as not to miss a day of wrestling. Wrestling was our savior: a healthy way to get out anger.²

But it wasn't all wrestling. We did something else too: We talked. We spent many hours together introspecting out loud, and just trying to make sense of things. Ernie has mentioned in print one early topic of discussion: "We spent years trying to solve various logical *conundra* like how on earth the Virgin Mary could have been a virgin

¹ The film has special meaning for me. My family lived near the building in which a man is thrown from the roof in an early scene, and my big brother Mike was an altar boy in the church in which the dock workers had their meeting. We moved from Hoboken to Union City and then, when I was 13, to more "upscale" North Bergen—the home town of the Cinderella Man, James J. Braddock.

² We are now faculty advisors to the Rutgers wrestling team.

and still have given birth to Jesus.”³ Being Irish and Italian, respectively, we were, of course, Catholic boys, even though not quite good Catholic boys.

Ernie came up the hard way. He is part of our meritocracy, not our aristocracy. His father left when he was an infant. On welfare until he was 18 years old, he was raised by his mother, Irma, an Italian immigrant who never learned to read and write English. She loved him dearly. But at home at night no one nagged him to do his homework. He made sure he did his homework. In high school, during wrestling season, he did it after an exhausting 2-hour practice. And he did it even though he'd gone the whole day without much to eat, not because he needed to make weight for wrestling, but because he didn't have any money. There were no free school lunches in those days. Making weight was easy.

Not that we never had money. There was plenty of work in the area, which was why it was so heavily populated. One way we earned money off-season was by working in a frozen food warehouse that was directly next to the Holland Tunnel in Jersey City. My cousin Riche McEnroe was the night foreman, and he let us work there part-time, from 8:00p.m. until midnight; we either loaded trucks or pulled carts full of boxes of frozen food out of the warehouse to be loaded onto trucks. It was the warehouse that Helene Stapinski's father worked in.⁴ He must have been working there when we worked there; in fact, I seem to remember him, but it may be a false memory. The full-time workers were members of the Teamster's Union, and so we all got a twenty-minute break every hour and forty minutes. That made the warehouse a wonderful place to study. You could work on a textbook for twenty minutes, taking notes, and then spend the next hour and forty minutes mulling over what you had read while doing mindless labor, occasionally even pulling a crumpled piece of paper out of your pocket to glance at the notes you wrote. I can't speak for Ernie, but of all the textbooks that I worked through in high school, I remember most the ones that I studied in the warehouse. We weren't the only ones who treated the breaks as study hall. I remember one guy, Greg, who worked full time in the warehouse and also went to college full time; he slept about 4 hours a day. I also remember one early April night when we were both in the backroom of the warehouse, and he took off his boot and dropped the corner of a heavy box on his big toe to break it, so that he could go on disability and have time to write his final papers and study for his exams. He aced all of his courses that semester.

In high school, Ernie was determined to be a lawyer someday. What made college financially possible for him was a wrestling scholarship, student loans (which he researched and applied for all on his own), and a scholarship from the Italian-American Club. The men in the Club kept an eye on the smart Italian kids in the neighborhood and loved it whenever one of them wanted to be a doctor or a lawyer.

³ LePore, *Meaning and Argument* (Oxford: Blackwell, 2000), p. xiii.

⁴ Helene Stapinski, *Five-Finger Discount: A Crooked Family History* (New York: Random House, 2002). This is a wonderful book if you have any interest in Jersey City. It is loaded with information about the Hague political machine. The only mistake in the book is that she confuses the Immaculate Conception with the Virgin Birth. (She describes North Bergen as “a scary place.”)

They wanted the kids to be able to leave the city, to have houses instead of apartments, and to be able to wear suits instead of the blue uniforms most of the factories in Hoboken and Jersey City required their workers to wear. There's a better life, they'd say. Although Ernie never became a lawyer, doesn't wear suits, and now once again lives in an apartment in North Bergen, I'm sure he hasn't disappointed them. He's a university professor, they'd brag.

Not all the kids from the neighborhood have done well. We know from first-hand observation the sorts of horror stories you'd expect poor inner-city kids to know. Growing up on mean streets can make you sentimental. A favorite film of ours is the 1938 "Angels with Dirty Faces," which starred James Cagney, Pat O'Brien, Humphrey Bogart, and the Bowery Boys. It starts with two kids running away from the police; one makes it over the fence, the police catch the other kid's ankle before he makes it over. The kid who makes it over the fence, the Pat O'Brien character, grows up to become a priest, Father Jerry. The kid whose ankle gets snared, the James Cagney character, Rocky Sullivan, grows up to be public enemy number one. In that role, he is the hero of the Bowery Boys, a bunch of teen-aged hoods from Rocky's old neighborhood. Rocky is eventually captured and sentenced to death in the electric chair. Father Jerry visits him in his cell at Sing Sing, on the Hudson, and asks a favor of his childhood friend. He asks him to go to the chair looking like a coward, so that the Bowery Boys will no longer view him as a hero. Rocky explains why he just can't do that. But when the time comes for him to walk to the chair, he is sobbing and has to be dragged kicking and screaming. The newspapers report the cowardly way "tough guy" Rocky went out, and the Bowery Boys are crushed. Father Jerry visits them in their basement hangout. He talks to them gently, and the last line of the movie, which he utters as he leads them from the basement, is, "Let's go say a prayer for a kid who couldn't run as fast as I could." Ernie and I weren't quite good Catholic boys, and now we are no longer boys or Catholic. But we still sometimes say a prayer for the kids who couldn't run as fast as we could.

During his first year at the University of Massachusetts, someone told Ernie that philosophy is the best major for law school; and so he became a philosophy major. He soon loved philosophy for its own sake. I visited him one weekend after he had become a philosophy major, and we spent almost the entire time walking around talking about a metaphysics course he was taking with Bruce Aune. I returned from that weekend thinking about his new fascination, philosophy, something I knew only a little bit about from my mother, who, as an English major at Seton Hall University, had taken an epistemology course when she was pregnant with me. I felt Ernie's involvement with philosophy was making us drift apart. So I signed up for a philosophy course for the following semester to find out what it was all about. The course was taught by Chris Hill, now of Brown University.

Irma never got to see Ernie as a philosopher, or even as a philosophy major. She died at the age of 47. Ernie was at school in Amherst, Massachusetts when he heard that she had to be rushed to a hospital. He took a bus to the Port Authority in New York City, a bus from there to North Bergen, and then walked a number of city blocks. He made it to the hospital to see her just before she passed away. He was

heartbroken, all on his own, and 18 years old. He mourned deeply. Then, he picked himself up by his own bootstraps and embraced life.

Ernie is my remaining friend from high school. But ‘my remaining friend from high school’ would be an improper description in Ernie’s mouth. Once Ernie’s friend, you’re his friend for life. You don’t need to make the effort to stay in touch. He’ll take the initiative to stay in touch with you. By the early 1970s, Ernie had a network of correspondence that I matched only after I got email in the 1990s. But by then Ernie had gotten email too, and his correspondence reached astronomical proportions. Friends of Ernie might not reach the cardinality of Friends of Bill, but I wouldn’t bet on that. I’ve had a number of philosophers tell me with a smile how Ernie emails them regularly. They are part of a large crowd. Ernie also stays in regular email contact with Steve Herman, his beloved undergraduate mentor, who has been out of philosophy for over twenty five years. He keeps up with Michael Keating, a high school friend who also went on to major in philosophy at college, but then moved into the world of business. And he keeps tabs on the two Tabs: Ernie and Tony Tabbacino, fellow high school wrestlers who have devoted their lives to teaching high school students. I could go on and on. When I wonder how some mutual high school friend is doing, I ask Ernie.

Ernie genuinely enjoys the successes of his friends: it makes him feel larger. And he tries to make his friends sharers and partners in his own successes. He loves to help people, and does so from respect, feeling and affection.

Ernie is completely devoted to his students. His sphere of caring, however, extends well beyond his students. Since 2004 he has been one of the twelve members of the New Jersey State Board of Education. He spends over 40 Wednesdays a year in Trenton fighting for resources for inner-city schools. And he also finds time to mentor young professorial philosophers. He remembers how Barry Loewer and others helped him when he was a struggling young assistant professor, and he passes it on. He has a keen, unflinching eye for genuine philosophical talent, and is always on the lookout for it. He has called me up many times to talk excitedly about some talented young philosopher he just met at a conference or whose article he just read, saying: “I’m going to invite him/her to the workshop.” For over a decade, he has been running an annual (and occasionally semiannual) philosophy of language workshop at the Rutgers Cognitive Science Center, where he is Associate Director. The workshop brings together promising young philosophers of language and leading senior figures in the field such as Gilbert Harman and James Higginbotham. Most of the contributors to this volume have participated in one or more of Ernie’s workshops, affectionately called “Erniefests.”

What goes around comes around. Some of the good Ernie has spread around has now come back to him in this superb, richly deserved volume on his philosophical work. Irma would be proud.

Part I

Compositionality

“If”, “Unless”, and Quantification

Sarah-Jane Leslie

Abstract Higginbotham (1986) argues that conditionals embedded under quantifiers (as in ‘no student will succeed if they goof off’) constitute a counterexample to the thesis that natural language is semantically compositional. More recently, Higginbotham (2003) and von Stechow and Iatridou (2002) have suggested that compositionality can be upheld, but only if we assume the validity of the principle of Conditional Excluded Middle. I argue that these authors’ proposals deliver unsatisfactory results for conditionals that, at least intuitively, do not appear to obey Conditional Excluded Middle. Further, there is no natural way to extend their accounts to conditionals containing ‘unless’. I propose instead an account that takes both ‘if’ and ‘unless’ statements to restrict the quantifiers in whose scope they occur, while also contributing a covert modal element to the semantics. In providing this account, I also offer a semantics for unquantified statements containing ‘unless’.

Keywords Conditionals · quantification · compositionality · modality · ‘unless’

1 Introduction: Quantified Conditionals and Compositionality

A language is semantically compositional if the meanings of its complex expressions are wholly determined by the meanings of their parts, and the manner in which those parts are combined. The belief that natural languages are semantically compositional has played a central role in contemporary semantics.

The belief is not stipulative, but is an empirical claim. It thus is conceivable that we might discover a counterexample to the thesis that natural languages are semantically compositional. We might, for example, discover that there are complex natural language constructions whose meanings do not depend solely on their parts, and the way in which those parts combine. A few such putative counterexamples have been discussed over the last thirty years, and one highly influential example was discussed by James Higginbotham in 1986. Higginbotham argued

S.-J. Leslie

Department of Philosophy, 1879 Hall, Princeton, NJ 08544, USA

that, when a conditional containing either “if” or “unless” is embedded under a quantifier, as in “no student will succeed if they goof off”, the meaning of the conditional varies depending on the nature of the quantifier in whose scope it occurs. Much discussion has come in the wake of Higginbotham’s 1986 article, such as Pelletier (1994a, b), Janssen (1997), von Fintel (1998), von Fintel and Iatridou (2002), and Higginbotham (2003). That quantified conditionals pose a challenge to the idea that natural languages are semantically compositional has acquired an almost folkloric status, and is frequently discussed in surveys and encyclopedia entries on compositionality. Pelletier discusses the possibility in his 1994 survey article on compositionality (1994b); in his *Handbook of Logic and Language* paper on compositionality, Janssen discusses the possibility that quantified conditionals may constitute a counterexample to the thesis that natural languages are compositional, and in a *Stanford Encyclopedia of Philosophy* entry on compositionality, Zoltan Szabo discusses quantified conditionals as a possible counterexample to this thesis.

I do not believe that quantified conditionals behave in a non-compositional manner. I will begin by considering conditionals that contain “if”, and consider a very simple account of their compositional structure. This simple account, which treats embedded conditionals as predicates of their quantified subjects, delivers satisfactory truth conditions for the most part, but runs into difficulties with conditionals that do not obey the principle of Conditional Excluded Middle. For that reason, I reject this simple account, and instead argue that an account that takes “if”-clauses to restrict quantifiers delivers the desired results, so long as we recognize that there is a covert modal element in the semantics of quantified “if”-statements.

I then consider quantified “unless”-statements, and propose a parallel account. We should understand quantified “unless”-statements as restricting the quantifying determiners in whose scope they occur, while also contributing a covert modal element to their semantics. In order to provide such an account, however, we need to understand the semantics of the unquantified versions of these statements, and so I develop a semantics for unquantified “unless”-statements. The account of quantified “if” and “unless”-statements I propose here provides a uniform meaning of “if” and “unless”; their semantics do not vary depending on the nature of the quantifier in whose scope they occur. We need not ascribe any sort of chameleon-like semantics to “if” and “unless”, which would have their meaning depend on the nature of the quantifier under which they are embedded.

Some of the discussion of Higginbotham’s claim has centered on the question of whether a chameleon-like semantics for conditionals would constitute a genuine counterexample to compositionality, or whether the principle of compositionality is sufficiently vague as to absorb the possibility (Pelletier, 1994a, b; Janssen, 1997). The principle of compositionality is sufficiently vague so as to encompass a variety of precisifications. Some of the more liberal formulations of the principle are arguably compatible with an item’s possessing a chameleon-like semantics, though the stricter formulations are not. I will not take up the question of whether compositionality is compatible with a chameleon-like semantics for an item, but will rather

argue that the proposed chameleon-like semantics does not even accurately capture the truth conditions of the relevant English sentences, and will offer a uniform semantics in its place.

2 The Puzzle of Quantified Conditionals

Higginbotham (1986) claims that “if” makes a different semantic contribution in (1) and (2) below, as does “unless” in (3) and (4):

- (1) Every student will succeed if they work hard.
- (2) No student will succeed if they goof off.
- (3) Every student will succeed unless they goof off.
- (4) No student will succeed unless they work hard.

He claims that, while the “if” and “unless” in (1) and (3) have the semantic values they would have were they not embedded under quantifiers, the “if” and “unless” in (2) and (4) have different semantic values altogether. It is important to notice here that Higginbotham (1986) is assuming that indicative conditionals have the semantics of material conditionals:

Elementary inferences involving these [subordinating conjunctions] proceed very well when they are understood as truth functional connectives, the material conditional [for ‘if’] and the non-exclusive ‘or’ [for ‘unless’] . . . The puzzle that I wish to discuss is independent of the issues of most prominent concern in that literature [on the semantics of conditionals], and it will be just as well to state it initially with the understanding that these classical terms of logical theory are truth functional. The puzzle is this: the words ‘if’ and ‘unless’ seem to have different interpretations, depending on the quantificational context in which they are embedded.

Higginbotham claims that (1) can be understood to contain a material conditional, and (3) an inclusive logical disjunction, and so no puzzle arises for those sentences. But if (2) were to contain a material conditional, then (2) would be true if and only if every student goofed off and didn’t succeed. Similarly, if (4) were to contain a disjunction, it would be true if and only every student both failed to work hard and failed to succeed.¹ Those truth conditions are not appropriate to the English sentence, however: (2) does not seem to entail that every student goofs off, and

¹ *No student will succeed if they goof off* is equivalent to: *for every student, it’s false that he will succeed if they goof off*, which in turn is equivalent to: *for every student, he will goof off and he won’t succeed*. Similarly, *No student will succeed unless he works hard* is equivalent to: *for every student, it’s false that he will succeed unless he works hard*, which, on the assumption that “unless” means *or*, is equivalent to: *for every student, he will not succeed and he will not work hard*. Here and for the rest of the paper I will make occasional reference to the truth functional equivalence of ‘no x (A)’ and ‘every x (not A)’ when both quantifiers have wide scope over the sentence, as do Higginbotham (2003) and von Stechow and Iatridou (2002). This is not intended as a claim about the semantics of ‘no’, nor as a claim that the two constructions are everywhere intersubstitutable, but merely as the observation that they are truth functionally equivalent when they have wide scope over the sentence in which they occur.

(4) does not entail that every student will fail to work hard. In fact both (2) and (4) are intuitively compatible with every student's recognizing them to be true and working hard as a result. Higginbotham (1986) notes this, and suggests that the truth conditions of (2) and (4) are rather given by (2') and (4'):

- (2') No student goofs off and succeeds.
- (4') No student succeeds and doesn't work hard.

These truth conditions contrast with the truth conditions of (1) and (3), where "if" and "unless" contribute a material conditional and an inclusive "or" respectively. Higginbotham concludes then that "if" and "unless" make different contributions depending on the nature of quantifier they are embedded under. This, he claims, is a counterexample to compositionality.

We should wonder whether Higginbotham's (1986) analysis adequately captures the truth conditions of (1)–(4). He proposes that (1) and (2) can be analyzed as (1') and (2'):

- (1) Every student will succeed if they work hard.
- (1') Every student will either succeed or not work hard.
- (2) No student will succeed if they goof off.
- (2') No student goofs off and succeeds.

(For clarity, I have formulated the material conditional in (1') as a disjunction.) These putative paraphrases do not adequately capture the truth conditions of the English sentences (1) and (2).²

To see the intuitive non-equivalence of (1) and (1'), consider poor Bill, who, no matter how hard he works, will never succeed at calculus. Bill knows this, and does not in fact try hard in his calculus class since he knows it is futile. Bill will then satisfy the material conditional in (1'), since he does not satisfy its antecedent – the equivalent disjunction "will either succeed or not work hard" is satisfied by Bill in virtue of his failing to work hard. Thus (1') may be true of a class containing Bill, since Bill presents no obstacle to its truth. But is (1) true if Bill is among the relevant students? The answer is quite clearly *no*. Bill is a student in that class, and so it is simply not true that every student will succeed if they work hard. Bill is a clear counterexample to this; no matter how hard he works, he won't succeed in this class.

Counterexamples to the paraphrasing of (2) by (2') also exist. Imagine a student in a New Jersey high school – let's call her Meadow – whose father has managed to scare the life out of her teacher. This teacher has no intention of giving Meadow anything less than an A in his class, no matter what she does. So it is simply not true that no student in the class will succeed if they goof off, for Meadow will succeed no matter what she does. It so happens, though, that Meadow is quite interested in

² Higginbotham (2003) and von Stechow and Iatridou (2002) discuss counterexamples of this nature, though they use them to object to 'restrictive analyses', which I will consider below. I am indebted to them for the structure of the counterexamples presented in this section of the paper.

the subject matter, and does not in fact goof off. Meadow is no obstacle to the truth of “no student goofs off and succeeds”, then, since she does not goof off, and so does not satisfy the conjunction “goofs off and succeeds”. Thus it can be true of a domain containing her that no student in it goofs off and succeeds. While (2) cannot be true of a class that includes Meadow, (2') can be, so we must reject (2') as an analysis of (2).

These same counterexamples tell against Higginbotham's (1986) analysis of (3) and (4) as (3') and (4'):

- (3) Every student will succeed unless they goof off.
- (3') Every student either succeeds or goofs off.
- (4) No student will succeed unless they work hard.
- (4') No student succeeds and doesn't work hard.

Intuitively, (3) cannot be true of any class that contains Bill, who will fail no matter what he does. But if Bill is again aware of his predicament, and so resolves not to waste his time trying in vain, then (3') may be true of a class containing Bill. Bill satisfies the disjunction “succeeds or goofs off”, and so poses no obstacle to the truth of (3'). Thus we might have a class that contains Bill, of which (3) is false but (3') is true.

Similarly, if Meadow is amongst the relevant students, (4) cannot be true, since she will succeed no matter what. It is intuitively false that no student will succeed unless they work hard, if Meadow is one of the students. If Meadow is once again interested in the subject matter and so elects to work hard, however, (4') may be true of a class containing her. If Meadow works hard, then she will not satisfy the conjunction “succeeds and doesn't work hard”, and so (4') might still be true of Meadow's class. Thus neither (3) and (3'), nor (4) and (4') are equivalent. Thus Higginbotham's (1986) non-compositional account does not even adequately capture the truth conditions of quantified conditionals, and so is untenable.

3 The Semantics of Conditionals Containing “If”

Let us set aside conditionals that contain “unless” for now, and focus on ones that contain “if”. “Unless”-statements are considerably more complex than “if”-statements, so it will be helpful to first formulate an account of “if”-statements. I will take up “unless”-statements in Part 4 of this paper.

3.1 A Simple Solution

There is a tempting solution to Higginbotham's puzzle of quantified conditionals, which would seem to let us deliver a fully compositional account in a most straightforward manner. To see this “Simple Solution”, let us put aside worries specific to conditionals for a moment, and consider the truth conditions of quantified sentences in general. Standard accounts of quantified sentences of the form “Q Ns VP” assign

to them truth conditions that depend on how many of the Ns possess the property denoted by the VP – in particular on whether the number or portion of Ns required by the quantifying determiner Q possess the property denoted by the VP. A relevant question to ask, then, is whether the truth of quantified conditionals depends on how many of the relevant items possess the conditional property, or – to put it in terms that do not make reference to conditional properties—how many of the relevant items satisfy the open conditional in question? Von Stechow and Iatridou (2002) argue that we can indeed provide a fully compositional account of quantified conditionals in this manner, and I argued as much myself in Leslie (2003a, b). While the Simple Solution offers an elegant, appealing and uniform treatment for the majority of cases, I will argue in the next section that, if we pursue the Simple Solution, we will be forced once again to adopt a chameleon-like semantics for “if” in a limited number of cases. I will take this to be good reason to look for an alternative account.

Let us consider in more detail how the Simple Solution would proceed. We saw above that the unfortunate Bill raised difficulties for Higginbotham’s non-compositional account of conditionals, since his presence is enough to render false “every student will succeed if they work hard”, *even if Bill does not in fact work hard*. On the Simple Solution, we would predict that Bill would falsify “every student will succeed if they work hard” iff Bill fails to satisfy “x will succeed if x works hard”. Intuitively, Bill does not satisfy this conditional: it is false that Bill will succeed if he works hard. Thus we would predict that Bill’s presence would be incompatible with the truth of “every student will succeed if they work hard”.

Similarly, we would predict that Meadow would indeed be a counterinstance to the claim “no student will succeed if they goof off”. For the quantified statement to be true of a domain containing Meadow, Meadow would have to fail to satisfy “x will succeed if x goofs off”. However, on any natural interpretation of the conditional, it is true that Meadow will succeed if she goofs off. It is clear, then, why “no student will succeed if they goof off” cannot be true of a class that includes Meadow.

This treatment is completely compositional with respect to the contribution of the conditional to the truth conditions of the entire sentence. It is also completely independent of any particular semantic treatment of conditionals themselves. We have offered no explanation of when an object satisfies the embedded conditional; this account of how conditionals compose appears to be independent of whatever the ultimate account of semantics for conditionals turns out to be. Just as it is not necessary to provide an account of *when* an item satisfies the predicate “is F” in order to highlight the compositional structure of “Q Ns are F”, if the Simple Solution was to succeed, it would not be necessary to provide an account of when an item satisfies an open conditional in order to see that a compositional analysis of quantified conditionals is possible. The Simple Solution, then, is an appealing option, and thus far it appears to handle our data correctly.

There is, however, a class of “if”-statements that are not well handled by the Simple Solution, namely those “if”-statements that do not obey the Law of Conditional Excluded Middle. I will also argue that “unless”-statements are simply not amenable to anything like the Simple Solution, but first let us consider those quantified “if”-statements that resist the Simple Solution.

3.2 *Conditional Excluded Middle*

Higginbotham (2003) and von Fintel and Iatridou (2002) each claim that we can give compositional interpretations to the troublesome quantified conditionals, but only if we assume that “if”-statements obey the law of Conditional Excluded Middle (CEM). Higginbotham describes this principle as follows:

Writing the Stalnaker conditional as ‘ \Rightarrow ’, we have the validity of (CEM), or Conditional Excluded Middle: (CEM) $(\varphi \Rightarrow \psi) \vee (\varphi \Rightarrow \neg\psi)$

(2003, p. 186)

Higginbotham is reluctant to endorse CEM, but takes it to be the only means of giving a compositional account of quantified conditionals. He writes, “Compositionality can be restored under certain assumptions [namely CEM] about the meaning, or the presuppositions, of conditionals. However, I am not aware at present of any way of grounding these presuppositions that is not stipulative” (p. 182). Von Fintel and Iatridou (2002) are more enthusiastic in their endorsement of CEM, since it is part of a theory of conditionals to which von Fintel is antecedently committed. Neither von Fintel and Iatridou, nor Higginbotham provide much explanation of why they believe CEM is a necessary assumption when analyzing *quantified* conditionals in particular, however. That von Fintel and Iatridou would assume CEM is perfectly understandable, since one of the authors has defended such an analysis of conditionals elsewhere. It is less than clear from his 2003 paper, though, why Higginbotham feels obliged to accept CEM.

The Simple Solution is just a way of dealing with *quantified* conditionals, and so should be neutral on the truth of CEM. If CEM is a true principle governing unquantified conditionals, then it should also govern quantified ones, but if certain unquantified conditionals do not obey CEM, we have no explanation of why these conditionals should suddenly obey it when they appear under a quantifier.

The Simple Solution made no assumptions whatsoever about the semantics of unquantified conditionals – we gave the truth conditions of quantified conditionals solely in terms of how many items satisfied or failed to satisfy the embedded conditional. If we encounter a conditional that does not obey CEM, then, we should be able nonetheless to analyze quantified versions of that conditional compositionally. Suppose, for example, (5) is a conditional that does not obey CEM:

(5) a is Q, if it is P.

Then by assumption “it’s false that a is Q, if it is P” is not equivalent to “if a is P, then it’s false that a is Q”, though “it’s false that: a is Q, if it is P” is nonetheless interpretable and acceptable. Then (6) should also be interpretable and acceptable:

(6) No x is Q if x is P.

(6) should be true just in case none of the relevant items satisfy the open conditional “x is Q, if x is P”. That an item can fail to satisfy the open conditional without satisfying “if x is P then it’s false that x is Q” should not affect our analysis. There is

nothing in the account presented here that even *suggests* that CEM is an assumption required to provide a semantics for quantified conditionals.

It is a controversial matter whether CEM is a principle governing all conditionals, or whether there are some that do not obey it. A good candidate for a conditional that does not obey CEM is (7):

(7) This fair coin will come up heads if flipped.

Suppose that we have a fair coin before us, and we are contemplating what will happen if we decide to flip it. On the assumption that the coin in question really is fair, (7) is intuitively false. Since (7) is a false conditional, if it obeyed CEM, then (8) would be true:

(8) This fair coin will not come up heads if flipped

However, (8) seems to be false also; it seems that we have a conditional that does not obey CEM.³

Let us now consider how the Simple Solution handles quantified conditionals whose embedded conditionals do not satisfy CEM. The above discussion suggests that the arbitrary fair coin fails to satisfy “x will come up heads if x is flipped”. The Simple Solution would then predict that (9) would be true of any given collection of fair coins, since each fair coin will fail to satisfy the embedded conditional:

(9) No fair coin will come up heads if flipped.

But (9) strikes us as false under these circumstances!⁴ (9) expresses a much stronger claim: (9) would be true only if each coin was sure *not* to come up heads if flipped. That is, (9) is true iff each coin satisfies the open conditional “x will not come up heads if x is flipped”.

A friend of the Simple Solution might respond by invoking CEM here. The conditional “x will not come up heads if x is flipped” is related to (9)’s embedded conditional via CEM: if CEM holds, then an item can fail to satisfy “x will come up heads if flipped” *if and only if* it satisfies “x will not come up heads if flipped”.

³ One might deny that (7) and (8) really are false, and claim instead, for example, that they are simply indeterminate, or lack a truth value. Certainly the defender of CEM as a general principle should argue for some such claim. I will not discuss such a possible defense here, but rather the discussion will proceed on the highly intuitive assumption that this is a genuine counterexample to CEM. It is worth noting, though, that it is far easier to convince oneself that (7) and (8) are indeterminate, than it is to convince oneself that their quantified counterparts (9) and (10) are:

(9) No fair coin will come up heads if flipped.

(10) Every fair coin will come heads if flipped.

(9) and (10) strike most people as quite clearly false. Thus even if one is inclined to reject (7) and (8) as counterexamples to CEM on the grounds that they are indeterminate rather than false, one still needs an explanation of why (9) and (10) seem quite clearly false and not at all indeterminate. Any natural extension of the Simple Solution to cases of indeterminacy would predict that the quantified statements should be indeterminate if their embedded conditionals are indeterminate.

⁴ I am indebted to Jim Higginbotham and David Chalmers for pointing this out to me.

Intuitively, it appears that (9) requires something like this for its truth: the coins in question must all fail to come up heads if flipped for (9) to be true. Thus it seems that the Simple Solution will be adequate only if we assume that an item fails to satisfy an open conditional “if $P(x)$ then $Q(x)$ ” if and only if it satisfies the conditional “if $P(x)$ then not $Q(x)$ ” – i.e. if we do assume that all conditionals obey CEM.

Higginbotham (2003) notes that this assumption is strange and stipulative; we have no explanation of why we would need to assume CEM for our analysis. On the account sketched here, we would in fact predict, on the face of it, that we would *not* need to assume CEM. If our difficulties were resolved by assuming that quantifiers demand that conditionals in their scope obey CEM, though, perhaps this would justify our adopting the stipulation. The situation, however, is not quite so straightforward.

To provide an adequate analysis of (9), we were forced to assume that the coins in question failed to satisfy “ x will come up heads if x is flipped” if and only if they satisfied “ x will not come up heads if flipped”. Fair coins do not intuitively satisfy “ x will come up heads if flipped”, but “no fair coin will come up heads if flipped” is clearly false. We explained this by assuming that, in order to *fail* to satisfy “ x will come up heads if flipped”, an item must *satisfy* “ x will *not* come up heads if flipped”. Fair coins clearly do not satisfy this latter conditional, so we concluded that, despite appearances, fair coins must satisfy “ x will come up heads if x is flipped” after all. We were then able to explain the falsity of (9), which is true if and only if none of the coins satisfy “ x will come up heads if flipped”. But this explanation of why (9) is false unfortunately predicts that (10) will be true:

(10) Every fair coin will come up heads if flipped.

In our explanation of (9)’s falsity, we stipulated that to fail to satisfy “ x will come up heads if flipped” *just is* to satisfy “ x will not come up heads if flipped”, and used that equivalence to arrive at the conclusion that each of the relevant coins must, in fact, satisfy “ x will come up heads if flipped”. These conditions, though, are exactly ones in which (10) ought to be true; thus our analysis predicts the truth of (10), despite its obvious falsity. We have purchased our explanation of (9)’s falsity only at the price of predicting (10)’s truth. Out of the frying pan and into the fire.

It is clear that (10) is false as long as (at least some of) the coins fail to satisfy the open embedded conditional, *even though they also fail to satisfy the CEM-equivalent conditional*. (9), however, is only true if the coins satisfy this CEM-equivalent conditional; it is not enough that they simply fail to satisfy the open embedded conditional. The proposed defense of the Simple Solution has led to the awkward position of requiring that our quantified conditionals both obey and fail to obey CEM. It appears that CEM is a necessary stipulation when we are providing a semantic analysis of conditionals under quantifiers such as “no”, but not if the quantifier is one such as “every”, CEM applies only if it applies to the unquantified version of the conditional. Thus if conditionals such as (7) and (8) do not obey CEM, we are forced to alter their semantics so as to conform to CEM when they occur under quantifiers like “no”, but not when they occur under quantifiers like

“every”. In this way we find ourselves back at square one; one semantic analysis applies to conditionals under “every”, and another to conditionals under “no”.

This suggests, I think, that we have not properly understood the logical form of conditionals embedded under quantifiers. The Simple Solution so far fares considerably better than Higginbotham’s original account – it provides adequate truth conditions in the vast majority of cases, and the violations are localized to those marginal and controversial conditionals that fail to obey CEM. Nonetheless, we have no explanation of why CEM is a necessary assumption for providing the semantics of conditionals embedded under “no”. We have even less of an explanation of why this assumption does not apply to conditionals embedded under “every”. The Simple Solution, though initially most appealing, is not ultimately adequate. Another approach is called for.

3.3 *A Modalized Restrictive Account*

A popular account of quantified conditionals emerges from the tradition that began with David Lewis (1975), which takes “if”-statements to restrict quantifiers and quantificational adverbs.⁵ Lewis argued that “if”-statements that occur in the scope of quantificational adverbs restrict the domain of quantification of that adverb. For example, we would analyze “always, if m and n are positive integers, the power m^n can be computed by successive multiplication” as involving quantification over pairs of positive integers. The sentence is analyzed to mean that, for all pairs of positive integers m and n , the power m^n can be computed by successive multiplication. Thus the “if”-clause “if m and n are positive integers” provides the domain of quantification for the adverb “always”.

Most contemporary theorists in this tradition assume that, if no explicit adverb of quantification is present in a conditional statement, then a covert universal quantifier over possible situations⁶ occurs in the sentence’s logical form. On this view, conditionals serve to restrict the domain of possible situations over which the quantifier ranges – be it an explicit quantificational adverb or a covert universal quantifier. It is almost always assumed that, if an explicit adverb of quantification occurs in the sentence, then the conditional will restrict that adverb, and no covert universal will occur in the sentence’s analysis.

On such an account, an “if”-statement of the form “If R , then M ” (i.e., in which no explicit quantificational adverb occurs) would be analyzed as:

⁵ A quantificational adverb is an adverb such as “always”, “sometimes”, “often”, “never”, and so on. Lewis (1975) argued that these adverbs quantify over cases or situations. Thus, for example, the sentence “John always wins” is to be analyzed to mean that all relevant situations involving John are ones in which he wins.

⁶ I.e. parts of possible worlds; see Kratzer (1989). In our discussion, nothing will hang on the use of situations rather than worlds. (An account that uses situations rather than worlds is useful in dealing with so-called ‘donkey’ sentences, such as “if a farmer owns a donkey, he beats it” (Heim, 1990). We will not be concerned with such sentences here.)

All $[C \cap R] [M]$

where C denotes the set of contextually relevant situations, and R and M are the interpretations of the antecedent and consequent respectively. Thus “if R , then M ” is true iff all of the contextually relevant situations in which “ R ” is true are ones in which “ M ” is true. If an explicit adverb of quantification occurs in the sentence, then that adverb will take the place of the covert universal quantifier. For example, “Never, if R , then M ” would be analyzed as:

No $[C \cap R] [M]$

Thus “never, if R , then M ” is interpreted to mean that no relevant situations in which “ R ” is true are situations in which “ M ” is true.

Lewis confined his original discussion to adverbs of quantification, but it is a natural further step to treat “if”-statements as restricting quantificational NPs such as “no students”, if the “if”-statement occurs in the scope of such an NP (see, e.g. Kratzer 1991; von Stechow 1998). On this view, we would construe (1) and (2) as (1*) and (2*) below:

- (1) Every student will succeed if they work hard.
- (1*) Every student who works hard will succeed.
- (2) No student will succeed if they goof off.
- (2*) No student who goofs off will succeed.

Or more formally:

- (1*LF) Every x [x is a student and x works hard] [x will succeed]
- (2*LF) No x [x is a student and x goofs off] [x will succeed]

Kratzer’s treatment of “if”-statements as restricting quantificational operators has been very influential. As it stands, though, it does not accurately capture the truth conditions of (1) and (2), since it is susceptible to the same counterexamples as Higginbotham’s (1986) account. Let us consider Bill once again – doomed to failure regardless of how hard he works – whose presence suffices to falsify (1). Should Bill decide not to work hard, though, then he poses no obstacle to the truth of (1*): he is not among the students who work hard, and so is irrelevant to (1*)’s truth or falsity. Thus (1) will be false while (1*) may yet be true. Similarly, the inclusion of the fortunate Meadow – who will succeed no matter what – among the relevant students is enough to render (2) false. Should Meadow decide not to goof off, though, then (2*) may well still be true, since only those students that actually goof off are relevant to the truth of (2*). This analysis, then, does not fare any better than Higginbotham’s original (1986) account.

It should be clear, though, exactly what the root of the difficulty is for this version of the restrictive account – the analysis is ignoring possible circumstances that are relevant for the truth of the quantified conditional because they are merely possible, and not actual. This difficulty does *not* arise for the restrictive analysis when the quantificational element is an adverb of quantification or a covert universal, because we are taking those quantifiers to range over *possible* situations. The truth conditions

of conditionals such as “Bill will succeed if he works hard” do not simply depend on the happenings of the actual world, because the covert universal is taken to range over possible situations. If we restricted the domain of the universal to actual situations, then we would predict inappropriate truth conditions if, as it happens, Bill never *actually* works hard. A modal element is needed to deliver the correct truth conditions for conditionals.

This suggests that our objection to treating *quantified* conditionals as restricted quantifiers, then, would be defeated were we able to include such a modal element in their truth conditions. Meadow falsifies “no student will succeed if they goof off” even if she does not actually goof off, because *were she to goof off, she would succeed nonetheless*. This modal fact is enough to guarantee that Meadow falsifies the quantified conditional, regardless of how events in the actual world unfold. We need to take these possible events into account when giving the truth conditions of quantified conditionals, just as we must when we are giving the truth conditions of conditionals that contain quantificational adverbs. Indeed, it would be rather surprising were the two types of constructions *not* to require such parallel treatment.

There are a variety of ways, it would seem, in which this idea might be implemented. We might take the quantifier to range over possible individuals, for example. Here, I will pursue a particular means of implementing the idea, which fits rather well with some recent work by Bart Geurts (m.s.), though there are other ways that one might implement the idea.

Geurts (m.s.) argues that, even when a conditional statement contains an explicit quantificational NP or quantificational adverb, the conditional may still serve to restrict a covert universal, in the same way that it does when no explicit quantifier or quantificational adverb is present. Geurts asks us to consider the following sentence:

(11) If Beryl is in Paris, she often visits the Louvre.

Geurts points out that (11) can be read as saying that on many of the occasions in which Beryl is in Paris, she visits the Louvre, or as saying that whenever Beryl is in Paris, she pays many visits to the Louvre. The first reading is obtained by taking the “if” clause to restrict the overt quantificational adverb “often”, while the second is obtained by taking the “if” clause to restrict a covert universal, of the sort that is standardly taken to occur in the absence of a quantificational adverb.

Geurts’ account differs from some more conventional views in that he claims that a conditional may have a covert universal associated with it, *even when the sentence contains an explicit quantifier or quantificational adverb*. Thus Geurts does not take such explicit items to block the emergence of a covert universal. Geurts, though, only discusses this covert operator in contexts where it is the operator that the conditional is restricting.

However, there is no reason that I know of that would prevent this covert universal quantifier over possible worlds from occurring in the logical form of a conditional statement, *even though the conditional is itself restricting an explicit quantifying determiner*. I propose that a conditional may contribute a covert universal quantifier to the semantics, even though the conditional itself serves to restrict an explicit quantifier. Further, I suggest that, when a conditional restricts an

explicit quantifier, this covert universal takes wide scope over the entire statement. Thus, when a conditional restricts a quantifier, every relevant possible world must be such that the quantified statement holds in it, for the entire statement to be true.

This account delivers the correct results for the quantified conditionals we have been considering. The unfortunate Bill – doomed to failure regardless of how hard he may work – posed a problem for a straightforward account of the conditional as restricting the quantifier. If we understood the conditional as restricting the quantified NP, with no modal element present, we would predict that the quantified conditional would be true, so long as Bill did not in fact work hard. The quantified conditional is not, however, intuitively true under those circumstances. The fact that, had Bill worked hard, he still would not have succeeded is enough to falsify the quantified conditional. I propose that we amend the above analysis, so as to include wide-scope quantification over contextually relevant possible worlds:

$\forall w \text{ Cw, } w_0$: Every x [x is a (relevant) student in w & x works hard in w] [x will succeed in w]

“ $\text{Cw, } w_0$ ” picks out a contextually determined restriction on the possible worlds over which we are quantifying. These truth conditions correctly predict that “every student will succeed if they work hard” will be false if Bill is among the relevant students. Since there are relevant worlds in which Bill works hard but does not succeed, the statement is false. Similar remarks apply to the quantified conditional “No student will succeed if they goof off”, which is falsified by Meadow’s presence, regardless of how hard she actually works. Since there are relevant possible worlds in which Meadow goofs off and still gets an A, the quantified conditional cannot be true.⁷

⁷ There is a fair amount of contextual variability associated with the restricting nominal “student” here. I have been eliding the details of this restriction, other than including a parenthetical ‘relevant’ in my representation of the logical form of these statements. There is far more that needs to be said here. In particular, it seems that some contextual restrictions allow the extension of the restricted nominal to change across the possible situations, while others do not. For example, if I say “every student will succeed if they work hard” with my introductory logic class in mind, there is a reading of the sentence on which it applies to any students who might possibly take my class. The utterance would then be a commentary on how I run my course. On this reading, the statement is false if the likes of Bill is even a possible member of my class. There is another reading of the sentence, though, on which it only applies to the students that have actually enrolled in my class, and thus understood is a commentary on the intellectual abilities of these students. On this reading, it does not matter whether Bill might have enrolled – that he has not in fact enrolled is enough to discount him from the evaluation of the statement. We should, I think, understand this variability as part of the general phenomenon of contextual variability in nominals – the property picked out by “is a student” might be such that its extension does not vary across the relevant possible situations, or it might be less rigid. (We could also locate difference between the readings in the set of relevant possible worlds we are considering. The proposal presented here is neutral between the two implementations, however, I am inclined to locate the restriction in the restricted nominal.) It should be noted, though, that it is less clear how these two readings would be generated, if we understood the statement to be quantifying over actual individual students, and attributing conditional properties to them, as we would under the Simple Solution. Unless we take the quantificational NP to range over possible individuals, it may be hard to avoid the consequence that the only available readings of the statement should be ones that pertain to the students that are, in fact, members of my class.

This Modalized Restrictive Account is thus able to deliver the correct truth conditions for (1) and (2). The Simple Solution, of course, was also able to handle these sentences correctly. However, our Modalized Restrictive Account, unlike the Simple Solution, delivers the intuitively correct results when faced with conditionals that do not obey Conditional Excluded Middle, without employing ad hoc assumptions.

Higginbotham (2003) claimed that a compositional account of quantified conditionals is forthcoming only if we assume that conditionals under quantifiers obey CEM. He was rightly uncomfortable with this result, feeling it to be little more than stipulation. As we saw above, the troubles run deeper than unexplained stipulation; the stipulation only applies to conditionals embedded under quantifiers such as “no”. If a conditional occurs under “every”, it obeys CEM only if its unquantified counterpart obeys CEM. Thus, in our above example, “every coin will come up heads if flipped” is a false claim, even though there is no coin in the domain that satisfies “x will not come up heads if flipped”. If CEM held here, the universally quantified claim would only be predicted to be false if there were such a coin. Thus CEM has to be imposed differentially on conditionals, depending on the nature of the quantifier they are embedded under. It was just this sort of chameleon-like semantics, though, that we set out to avoid.

Our Modalized Restrictive Account yields the right predictions without recourse to such uncomfortable assumptions and chameleon-like analyses. The Modalized Restrictive Account would render “no fair coin will come up heads if flipped” as:

$\forall w Cw, w_0$: No x [x is a fair coin in w & x is flipped in w] [x will come up heads in w]

On this analysis, the statement is true iff in all relevant possible circumstances, none of the coins that are flipped will come up heads. These are the truth conditions we have been seeking, and we are able to arrive at them without making questionable assumptions about the plausibility of CEM in such a case.

Similarly, we have at hand a straightforward, parallel analysis for “every coin will come up heads if flipped”:

$\forall w Cw, w_0$: Every x [x is a fair coin in w & x is flipped in w] [x will come up heads in w]

It is clear that this analysis correctly predicts that “every coin will come up heads if flipped” will be false. The Modalized Restrictive Account is able to capture the strong truth conditions of *both* the quantified conditionals. The Simple Solution issued in overly weak conditions for the conditional under “no”, unless CEM was assumed to apply. However, once CEM was assumed, the truth conditions for the conditional under “every” were predicted to be overly weak. Only a differential application of CEM captured the strong truth conditions of both statements. Our restrictive analysis allows us to avoid any such differential assumptions. This consideration constitutes good reason to prefer a restrictive analysis of conditionals to the Simple Solution.

Furthermore, we will see in the next section that no version of the Simple Solution is applicable to quantified “unless”-statements, while a Modalized Restrictive

Account delivers the desired results. Treating quantified “if” and “unless”-statements in a uniform manner constitutes further reason to prefer the restrictive account to the Simple Solution in the case of “if”-statements.

4 The Semantics of Conditionals Containing “Unless”

The Simple Solution to the puzzle of quantified conditionals treated embedded “if”-statements as predicating a conditional property of the quantified NP subject. The truth value of a sentence of the form “Q Ns P, if R” would then depend on how many of the relevant Ns possess the conditional property. The logical form of such a sentence, we have claimed, might be given as follows:

Q [N] [P if R]

A parallel account for “unless” would render the logical forms of (3) and (4) as follows:

(3) Every student will succeed unless they goof off.

(3 LF) Every x [x is a (relevant) student] [x will succeed unless x goofs off]

(4) No student will succeed unless they work hard.

(4 LF) No x [x is a (relevant) student] [x will succeed unless x works hard]

(3 LF) handles Bill’s case adequately: Bill does not satisfy “x will succeed unless x goofs off”, since it’s false that Bill will succeed unless they goof off. Thus if (3 LF) is the logical form of (3), we would predict that (3) would not be true if Bill is among the relevant students. But what of (4 LF)? We wish to predict that a sentence whose logical form is given by (4 LF) will not be true if Meadow is among the relevant students. (4 LF) is true if none of the relevant students satisfy the open “unless”-statement, or alternatively if all of the relevant students fail to satisfy it. Meadow will present an obstacle to the truth of (4 LF) iff she satisfies “x will succeed unless x works hard” . . . and here we encounter a difficulty.

“Meadow will succeed unless she works hard” is intuitively false. This is not a true sentence in the scenario we have described. Meadow will succeed no matter what she does, so *it’s false that Meadow will succeed unless she works hard*. Thus, if the logical form of (4) was given by (4 LF), Meadow would pose no obstacle to the truth of (4). She fails to satisfy the open “unless”-statement, and so it is quite possible that *no student* in a class containing her would satisfy it.

The Simple Solution, then, does not even begin to accommodate “unless”-statements. It appears that the truth conditions of quantified “unless”-statements do *not* depend on how many members of the domain satisfy the open “unless”-statement. Statements of the form “No Ns P unless they R” cannot be understood to mean that *No Ns satisfy “P unless they R”*. In the case of “unless”-statements, we do not need to invoke conditionals that fail to obey CEM to raise difficulties for the Simple Solution. It cannot handle these rather basic examples of quantified “unless”-statements.

One might, of course, just deny that “Meadow will succeed unless she works hard” really is false. One might argue that it is merely pragmatically unacceptable for one reason or another. I find such a solution deeply unsatisfying. To my ear, and the ears of my informants, it is simply false that Meadow will succeed unless she works hard. Any account that validates that intuition should be preferred to one that dismisses it. In what follows I will propose such an account, and to the extent that it is successful, it provides us with a far more satisfactory account than one that chalks up the appearance of falsity here to mere pragmatic factors.⁸

Let us then accept at face value the intuition that it’s false that Meadow will succeed unless she works hard. One way that we might frame our puzzle is as follows: for unquantified “unless”-statements, there appears to be a “uniqueness” requirement. This “uniqueness” requirement has it that, for the “unless”-statement to be true, it would have to be the case that working hard is the only relevant way in which Meadow will fail to succeed. Since this is false in the case described, the “unless”-statement is predicted to be false. This uniqueness requirement, however,

⁸ Treating ‘unless’ as meaning ‘if . . . not’ is the most obvious way to fill out the claim that Meadow really does satisfy the relevant ‘unless’-statement: It’s true that Meadow will succeed if she doesn’t work hard. Higginbotham (2003) proposes that we handle ‘unless’ in this manner, and claims that a compositional treatment of quantified ‘unless’-statements is possible so long as ‘unless’ is assimilated to ‘if . . . not’. (Higginbotham provides few details, so it is not clear whether he proposes this to deal with situations such as Meadow’s, or for some other reason.) Besides a general desire not to simply dismiss as pragmatic any phenomenon that threatens semantic simplicity, there are other considerations that weigh against treating ‘unless’ as ‘if . . . not’. Geis (1973) produces a battery of reasons not to equate ‘unless’ with ‘if . . . not’, and I refer my reader to his excellent article for more detailed discussion than I can provide here.

Geis notes that ‘unless’ and ‘if . . . not’ behave differently with respect to the possibility of coordinate structures. There is no obstacle to conjoining clauses containing ‘if . . . not’, but we cannot do the same with clauses containing ‘unless’. Compare, for example:

John will succeed if he doesn’t goof off and if he doesn’t sleep through the final.

*John will succeed unless they goof off and unless he sleeps through the final.

‘Unless’ and ‘if . . . not’ also interact differently with negative polarity items. Naturally, negative polarity items can occur in the scope of ‘if . . . not’. They cannot, however, occur in the scope of ‘unless’:

John won’t succeed if he doesn’t ever attend class.

*John won’t succeed unless he ever attends class.

As a final point against the identification of ‘if . . . not’ and ‘unless’, we should note that clauses containing ‘if . . . not’ can be modified by ‘only’, ‘even’, ‘except’, while clauses containing ‘unless’ cannot:

John will succeed only if he doesn’t goof off.

*John will succeed only unless they goof off.

John will succeed even if he doesn’t work hard.

*John will succeed even unless he works hard.

John will succeed except if he doesn’t work hard.

*John will succeed except unless he works hard.

I will take these considerations and others in Geis (1973) to tell strongly against the identification of ‘unless’ with ‘if . . . not’ that Higginbotham (2003) suggests, and so this particular means of deriving the falsity of “Meadow will succeed unless she works hard” is untenable. Perhaps other means might be proposed, but I do not know of any other such proposals.

seems to disappear when “unless”-statements are embedded under some quantifiers such as “no students”. “A will succeed unless A works hard” is false if A can succeed while working hard. However, the mere fact that working hard and succeeding are compatible for each student does not suffice to make true “No student will succeed unless they work hard”. The quantified statement is much stronger. It is not made true by the mere *compatibility* of working hard and succeeding. For Meadow, hard work and success are certainly compatible, but this does not mean that “No student will succeed unless they work hard” can be true of a class containing her. It cannot. “No student will succeed unless they work hard” specifically rules out the possibility of students like Meadow, who may succeed without hard work.

I will argue that a Modalized Restrictive Account of quantified “unless”-statements will deliver the results we are seeking. Once we have a satisfactory account of “unless”-statements that occur in the presence of adverbs of quantification, it will be a simple matter to extend this account to handle “unless”-statements that are embedded under quantifiers.

4.1 Von Fintel’s Account of “Unless”

To make progress here, we will need to understand the semantics of unquantified “unless”-statements in more detail. There has been relatively little contemporary discussion of “unless”, but fortunately von Fintel (1992, 1994) offers an excellent discussion that will be extremely helpful to us here. Von Fintel’s account extends and formalizes Geis (1973), and includes a uniqueness condition that explains why “Meadow will get an A unless she works hard” is false.

Von Fintel’s account of “unless”-statements follows in the Lewis-Kratzer tradition of treating conditionals as restrictions on quantificational adverbs, and he assumes, along with most theorists, that a covert universal quantifier occurs in the absence of an explicit quantificational adverb.

Let us begin by considering cases in which no adverb of quantification is present in the sentence, and so the quantifier in question is a covert universal. Von Fintel’s account of “unless”-statements has two parts. The first part treats “unless”-statements as having as part of their meaning something akin to “if . . . not”. Thus “M unless R” has its interpretation given in part by:

All [C – R] [M]

It is thus part of the truth conditions of “M unless R” that all relevant situations in which “R” is false are ones in which “M” is true. It should be clear that this is extensionally equivalent to the analysis we would give for “If not R, then M”, and so the common intuition (see, e.g., Higginbotham 2003) that “unless” is akin to “if . . . not” is captured by this part of von Fintel’s treatment.

“Unless” does not simply mean “if . . . not” (Geis, 1973, see also fn 8). Von Fintel recognizes this, and so includes a so-called uniqueness condition, which he formulates as follows:

$$\forall S (\text{All } [C - S] [M] \rightarrow R \subseteq S)$$

Thus, for any set of situations S , if all relevant situations that are not S situations are also M situations, then S includes R as a subset. It is this condition that explains the falsity of, e.g., “Meadow will get an A unless she goofs off”. It is certainly true that all relevant situations in which Meadow does not goof off are situations in which she will get an A, thus the first condition of the analysis is satisfied. But the uniqueness condition will not be satisfied. Consider a proper subset of the (possible) situations in which she goofs off – say, situations in which she both goofs off and chews gum in class. Clearly, the set of situations in which she goofs off is not a subset of the situations in which she both goofs off and chews gum in class, at least on the very natural assumption that there are some relevant, possible situations in which she goofs off but does not chew gum. However, since Meadow will get an A in all relevant situations, she will a fortiori get an A in situations in which she *either* doesn’t goof off, *or* doesn’t chew gum in class. But this disjunctive set of situations just consists of the situations denoted by $[C - S]$, where S is the set of situations in which she both goofs off and chews gum. Thus, we have found a set S of situations such that all the relevant non- S situations are situations in which Meadow gets an A, but the set of situations in which Meadow goofs off is not a subset of this set S . Thus the uniqueness condition is not satisfied. The uniqueness condition will only be satisfied if all the situations in which “ R ” holds are situations in which “ M ” does not hold. If there are any R -situations that are also M -situations, then if we subtract these situations from R , we will obtain a set S that falsifies the uniqueness condition.

Von Fintel’s account of statements of the form “ M unless R ” thus contains two conjuncts:

$$\text{All } [C - R] [M] \ \& \ \forall S (\text{All } [C - S] [M] \rightarrow R \subseteq S)$$

It should be obvious by now that we will not be able to use this analysis to give an account of quantified “unless”-statements in any straightforward manner. If we try to treat “No students are M unless they are R ” as

$$\text{No } x [x \text{ is a student}] [x \text{ is } M \text{ unless } x \text{ is } R]$$

and use von Fintel’s analysis of the “unless”-statement, we will obtain:

$$\text{No } x [x \text{ is a student}] [\text{All } [C - \{s: x \text{ is } R \text{ in } s\}] [\{s: x \text{ is } M \text{ in } s\}] \ \& \ \forall S (\text{All } [C - S] [\{s: x \text{ is } M \text{ in } s\}] \rightarrow \{s: x \text{ is } R \text{ in } s\} \subseteq S)]$$

(where C is the set of relevant situations.) But as long as all the students fail to satisfy at least one of the conjuncts of the analysis, the statement will be true. As before, this predicts that Meadow will pose no obstacle to the truth of “no student will get an A unless they work hard”, since she will not satisfy the uniqueness condition of the “unless”-statement. Once again, the uniqueness condition – essential for an account of “unless”-statements that do not occur under quantifiers – creates difficulties once we try to embed the statement under “no”.

4.2 A Modified Account of “Unless”

Before we return to our main project of giving an account of “unless”-statements under quantifiers, let us consider how von Fintel’s account fares when there is an explicit adverb of quantification present in the “unless”-statement. Von Fintel (1992) formulates his account in general terms as follows:

$$Q [C - R] [M] \ \& \ \forall S (Q [C - S] [M] \rightarrow R \subseteq S)$$

where “Q” is the relevant quantifier – either a covert universal as before, or an adverb of quantification that explicitly occurs in the sentence. Let us see how his account handles a statement such as (12):

(12) John never succeeds unless he works hard.

(Or to make the scope of the adverb more apparent, we may substitute the more awkward “Never, unless he works hard, does John succeed”.) Clearly, (12) cannot be true if there are any possible, contextually relevant situations in which John succeeds without working hard. Von Fintel (1994) claims that (12) also requires for its truth that any time John works hard, he succeeds, but this seems to me too strict a requirement for the truth of (12). (12) may be true, yet there be some relevant situations in which even hard work does not suffice for John’s success. My intuitions, and those of my informants, have it that it should not be part of the truth conditions of (12) that every situation in which John works hard is one in which he succeeds. If John is someone who finds his coursework extremely difficult, and so never succeeds without hard work, (12) will be true, *even if John sometimes finds the work so difficult, that he fails despite working hard*.

Von Fintel’s account, however, predicts that the truth conditions of (12) would include such a strict requirement. His above analysis, applied to (12), would be as follows:

$$\text{No } [C - \{s: \text{John works hard in } s\}] [\{s: \text{John succeeds in } s\}] \ \& \ \forall S (\text{No } [C - S] [\{s: \text{John succeeds in } s\}] \rightarrow \{s: \text{John works hard in } s\} \subseteq S)$$

The first conjunct above is perfectly correct – it states that no relevant situation in which John does not work hard is a situation in which John succeeds. The second conjunct – the uniqueness condition – imposes an overly demanding condition, however. The second conjunct is not satisfied as long as there is some set of situations S such that none of the relevant non-S situations are situations in which John succeeds, and yet S does not contain the situations in which John works hard. Suppose, for example, that amongst the contextually relevant situations are ones in which the subject matter is just too difficult for John to master. No matter how hard he works, he won’t succeed in those situations. Intuitively, (12) can be true despite the possibility of such situations, but the uniqueness clause in von Fintel’s account is violated under these circumstances.

To see that this is so, let us take S to be the set of situations in which the subject matter is *not* too difficult for John. Let us further suppose that there are some relevant possible situations in which John works hard, even though the subject matter,

regrettably, is just too difficult for him. (This supposition is just the one described in the preceding paragraph.) Then the set of situations in which John works hard will not be a subset of S , and so for this S it is false that:

$$\{s: \text{John works hard in } s\} \subseteq S$$

However, since S is the set of situations in which the subject matter is not too difficult for John, $[C - S]$ is the set of relevant situations in which the subject matter *is* too difficult for John. In the scenario we are describing, none of these situations are situations in which John succeeds. Thus it is true that:

$$\text{No } [C - S] [\{s: \text{John succeeds in } s\}]$$

Thus von Fintel's uniqueness clause is violated, and so (12) is predicted to be false, so long as there are some situations in which John works hard but still doesn't succeed. Intuitively, however, it may be true that John never succeeds unless he works hard, even though sometimes his hard work isn't enough to secure his success. Sometimes the subject matter is simply beyond him. Thus von Fintel's account does not correctly handle "unless"-statements that contain the quantificational adverb "never".

If von Fintel's account included only its first part – the requirement that no situations in which John does not work hard be ones in which he succeeds – we would have the intuitively correct truth conditions for "John never succeeds unless he works hard". As we saw above, however, the uniqueness clause is needed to provide adequate truth conditions for "unless"-statements that contain universals, be they covert or overt. How can we accommodate this data in a compositional manner?

I believe that the problem lies in the formulation of von Fintel's uniqueness clause. Further evidence that it is not properly formulated emerges when we consider "unless"-statements that contain adverbs of quantification such as "usually" or "rarely", as in (13) and (14). Von Fintel (1994) claims that statements such as (13) and (14) are ill-formed and semantically deviant. I must admit that I simply do not share this intuition, nor do my informants. Since (13) and (14) are perfectly fine to my ear, I will aim to provide an account of "unless" that captures their truth conditions adequately.

(13) John usually succeeds unless they goof off.

(14) John rarely succeeds unless he works hard.

(Or to make it absolutely clear that the quantificational adverbs have scope over the whole statements:

(13') Usually, unless John goofs off, he succeeds.

(14') Rarely, unless John works hard, does he succeed.

I cannot find anything objectionable about these sentences.)

Von Fintel's account cannot be successfully applied to (13) and (14); this is natural since von Fintel does not intend that it should apply to them. The account applied to (13) would yield the following:

Most $[C - \{s: \text{John goofs off in } s\}] [\{s: \text{John succeeds in } s\}] \& \forall S (\text{Most } [C - S] [\{s: \text{John succeeds in } s\}] \rightarrow \{s: \text{John goofs off in } s\} \subseteq S)$ ⁹

As before, I cannot find fault with the first conjunct of the account – it is certainly necessary for the truth of (13) that most situations in which John does not goof off are ones in which he succeeds. It is the uniqueness condition that gives us cause for concern.

Suppose, for example, that for the most part, when John doesn’t goof off, he succeeds, and again for the most part, when John does goof off, he doesn’t succeed. Let us say, though, that once in a while John bribes his teacher, in which case he usually succeeds, no matter what he does. It seems that (13) is true under these circumstances, as long as John very rarely bribes his teacher, but the uniqueness clause is violated. We may take our S to be the set of situations in which John does not bribe his teacher. John almost never bribes his teacher if he is planning to work hard – what would be the point? – so the situations in which he does bribe his teacher are generally ones in which they goof off. Thus the set of situations in which John goofs off are *not* a subset of S , i.e. of the situations in which he refrains from bribing his teacher. However, as we have described the example, most of the situations in which John *does* decide to bribe his teacher are ones in which John succeeds, so it is true that:

Most $[C - S] [\{s: \text{John succeeds in } s\}]$

Once again, the uniqueness clause is not satisfied, and so we would predict that “John usually succeeds unless they goof off” would be false as described. It is enough to falsify von Fintel’s analysis that John very occasionally bribes his teacher and, having done so, usually succeeds as a result. Intuitively, however, “Usually, John succeeds unless they goof off” is not so strong a claim as to be incompatible with such circumstances.

In the above example, it was important that we stipulated that John only occasionally bribes his teacher to succeed. If this was common practice for him, then (13) would not be true. It does not seem correct to say, for example, that Meadow usually succeeds unless she goofs off. Thus in the case of “most” or “usually”, some second conjunct is needed, for it is not enough for the truth of the claim that most of the situations in which Meadow does not goof off be ones in which she succeeds. The first conjunct of von Fintel’s analysis alone would not suffice here, though it seemed that it would suffice when the adverb of quantification was “never”.

We can mount a similar argument against the appropriateness of the uniqueness clause in the case of (14), which is an “unless”-statement that contains the adverb “rarely”. Suppose, for example, that it only occasionally happens that John succeeds without working hard – for the most part, he only succeeds when he works hard. Then (14) is intuitively true. If we suppose further that, sometimes, the subject

⁹ I am assuming here that ‘usually’ can be understood as ‘most’, and so am setting aside any additional normative or otherwise modal import ‘usually’ may possess; nothing will hang on this simplifying assumption.

matter is simply too hard for John, and despite his best efforts, he does not succeed, then (14) remains intuitively true. Once again, however, von Fintel’s semantics predicts that the statement will be false. I shall not go through the details again, but my reader may convince herself that this is so by taking the set S to be the set of situations in which John works hard, and the subject is not too difficult for him.

Let us then summarize our desiderata for an account of “unless”-statements of the form “ $Q M$ unless R ”. The first conjunct of von Fintel’s analysis was absolutely correct in all cases:

$$Q [C - R] [M]$$

It would be surprising if this part of the analysis was not correct, given the strong intuitions that “unless” is semantically similar to “if . . . not”. The uniqueness clause so far has proved tricky, however. We would like it to be equivalent to von Fintel’s uniqueness clause when the quantifier is a universal, but we would like it to effectively evaporate when the quantificational adverb is “never”. We would like to have some version of a uniqueness clause when the adverb is “usually”, but we would like it to amount to the requirement that most of the situations in which M holds are ones in which R doesn’t hold, so as to allow the truth of “John usually succeeds unless they goof off” if John very occasionally bribes the teacher, but not if he does so as a matter of course.

I propose that we analyze statements of the form “ $Q M$ unless R ” as follows:

$$Q [C - R] [M] \ \& \ Q [M \cap C] [C - R]$$

“ $Q M$ unless R ” is true, then, if and only if Q of the relevant non- R situations are M situations, and Q of the relevant M situations are non- R situations.

If there is no explicit adverb of quantification in the sentence, then I assume that a covert universal is present. Thus “John will succeed unless they goof off” is analyzed as:

$$\text{All } [C - \{s: \text{John goofs off in } s\}] [\{s: \text{John succeeds in } s\}] \ \& \ \text{All } [\{s: \text{John succeeds in } s\} \cap C] [C - \{s: \text{John goofs off in } s\}]$$

The sentence is true just in case all relevant situations in which John doesn’t goof off are ones in which he succeeds, and all relevant situations in which John succeeds are ones in which he doesn’t goof off. These truth conditions are equivalent to the ones that von Fintel provides for “unless”-statements that contain universal quantifiers.¹⁰

¹⁰ The formulation of von Fintel’s uniqueness clause needs to be amended in order for these to be strictly equivalent, but it is a minor adjustment, and is independently motivated. As it stands, von Fintel has the following as his uniqueness clause:

$$\forall S (Q [C - S] [M] \rightarrow R \subseteq S)$$

However, the clause, as it stands, is violated if there are ‘irrelevant’ R -situations (i.e. situations that are in R , but not in C). That is, statements such as “John will succeed unless he doesn’t work hard” would be predicted to be false if there are possible situations outside of the contextually relevant ones in which John doesn’t work hard – situations in which, e.g., John dies in a freak accident. To see that the uniqueness clause is violated, take S to be the contextually relevant situations in

Our first desideratum, then, is satisfied: If the quantifier is a universal, our account is equivalent to von Fintel’s.

What, though, of the quantifier “no”, or “never”? We wished that the uniqueness clause would “evaporate” in such cases, and this is indeed what we obtain. The determiner “no” is a *symmetric* determiner (Barwise & Cooper, 1981); “no As are Bs” is true iff “no Bs are As” is also true. We may “swap” the material in the restrictor with that in the scope, without changing the truth value of the claim, if the quantifier is “no” or its equivalent. But the uniqueness clause we are considering amounts to just this exchange! We might just as easily have written “ $M \cap C$ ” in place of “ M ” in the first conjunct, and so formulated our analysis as follows:

$$Q [C - R] [M \cap C] \& Q [M \cap C] [C - R]$$

If Q is symmetric, then the two conjuncts are equivalent. We are thereby able to capture the intuition that there is no real uniqueness clause when the quantifier is “no” – it is enough for the truth of the “unless”-statement that no relevant non- R situations be M situations. The uniqueness clause does not in fact evaporate in a non-compositional manner, but simply becomes redundant if the quantifier in question is “no”.

There is, I think, a suggestion of sorts to the effect that there are *some* situations in which M and R both hold, and the “uniqueness” clause provides an explanation of this suggestion, so it is not completely vacuous. For example, “John never gets an A unless he works hard” suggests that there are some possible situations in which John gets an A by working hard. This would seem to be related to the implication or presupposition carried by “if”-statements that contain “never”, such as “John never gets an A if they goof off”. We would analyze the “if”-statement as

$$\text{No } [C \cap \{s: \text{John goofs off in } s\}] [\{s: \text{John gets an } A \text{ in } s\}]$$

Strictly speaking, this analysis predicts that “John never gets an A if they goof off” is true if there are simply no relevant possible situations in which John goofs off. Intuitively, though, the English conditional suggests that it is a live possibility that John will goof off, and, of course, that in such possible situations, John will fail to get an A .

I do not think that the situation is much different in the case of “John never gets an A unless he works hard”, which we would represent as:

$$\text{No } [C - \{s: \text{John works hard in } s\}] [\{s: \text{John gets an } A \text{ in } s\}] \& \text{No } [C \cap \{s: \text{John gets an } A \text{ in } s\}] [C - \{s: \text{John works hard in } s\}]$$

which John doesn’t work hard (i.e. $S = C \cap R$). This difficulty is easily remedied by rendering the uniqueness clause as:

$$\forall S (Q [C - S] [M] \rightarrow C \cap R \subseteq S)$$

The adjustment is minor, and surely reflects von Fintel’s original intentions. Once we have made this adjustment, the two clauses are provably equivalent when the quantifier in question is a universal.

The second conjunct here is a perfect parallel to the analysis of the “if”-statement above, and it carries with it a similar suggestion (implication or presupposition, depending on the details of one’s account) that there are some live possibilities in which John gets an A. Those possibilities cannot be ones in which John *doesn’t* work hard, so we derive the suggestion that it’s possible for John to work hard and get an A. We do not need to make any assumptions specific to “unless”-statements here – however we account for the parallel suggestion with “if”-statements should carry over here. (Von Fintel’s account of “if”-statements is an example of an account that treats this suggestion as a presupposition.)

The account set out so far also provides an appealing analysis of “unless”-statements that contain “usually”, as in “usually M, unless R”. We wanted our account to require that most of the situations in which “M” holds be situations in which “R” holds. For example, we wished to explain why “John usually succeeds unless they goof off” was compatible with John’s occasionally slipping the teacher a bribe, but not with his doing so as a matter of course. We are now able to do so; “John usually succeeds unless they goof off” will be analyzed as:

$$\text{Most } [C - \{s: \text{John goofs off in } s\}] [\{s: \text{John succeeds in } s\}] \ \& \ \text{Most } [\{s: \text{John succeeds in } s\} \cap C] [C - \{s: \text{John goofs off in } s\}]$$

The “unless”-statement will be false if most of the relevant situations in which John succeeds are just ones in which he bribes the teacher, then kicks back, since the second conjunct will not be true under those conditions. As long as we are only considering the occasional bribe, however, the “unless”-statement will be true.

We thus arrive at a plausible and appealing account of “unless” by adopting this formulation of the uniqueness clause. Our account now yields the right results even when the sentence in question contains a quantificational adverb that is not a universal.

4.3 *Aside: Uniqueness Clauses and Coordinate Structures*

One might worry that, in allowing the uniqueness clause to evaporate in “unless”-statements containing “never”, we lose an explanation of why “unless” clauses cannot be conjoined. Geis (1973) points out that coordinate structures with “unless” are not permissible, for example:

- (15) *John will get an A unless they goof off and unless he sleeps through the final.

Von Fintel (1991, 1994) proposes that his uniqueness clause explains the impermissibility of this statement – since the “unless” clause expresses the unique minimal restriction that makes the conditional true, there cannot be another such clause. If both restrictions made the conditional true, then the uniqueness clause would not be satisfied in either case. Von Fintel’s account of “unless”-statements containing “never” also features a uniqueness clause, and so he claims that it also predicts the unacceptability of conditionals such as

- (16) *John never gets an A unless he works hard and unless he bribes the teacher.

I have argued that his uniqueness clause issues in overly strong truth conditions for “unless”-statements containing “never”, and so is not a desirable component of an account of such statements. The uniqueness clause also did violence to “unless”-statements containing “usually” and “rarely”, which, *pace* von Fintel, are quite acceptable. But do we achieve our intuitively correct truth conditions at the cost of losing our explanation of the impermissibility of coordinate structures like the ones above?

On closer inspection, it is far from clear that von Fintel’s uniqueness clause does in fact explain the unacceptability of the coordinate structures. His semantics predict *only* that the two sets of situations denoted by the two “unless”-clauses are coextensive. If they are not coextensive, the statement is false, not defective. And if they are coextensive, then his semantics predicts that the statement will be true! Consider, however, statements such as:

- (17) *I will respect the list of endangered species unless it contains renates and unless it contains cordates.

This statement is just as unacceptable as the two above, but it is far from clear why this should be so on von Fintel’s account. Since the two “unless”-clauses denote states of affairs that are coextensive in the possible worlds that are likely to be relevant, there is no obstacle to the uniqueness clause being satisfied for both “unless” clauses. Relatedly, it is not clear why statements such as:

- (18) *John will get an A unless he goofs off and unless he sleeps through the final.

are *impermissible*, as opposed to simply entailing that John will goof off if and only if he sleeps through the final.

We should also be hesitant to offer a straightforwardly semantic explanation for the impermissibility of these constructions, since conjoined “unless” clauses are far more acceptable when they occur at the left periphery of the sentence. Consider, for example:

- (19) Unless they goof off, and unless he sleeps through the final, John will succeed.

Rearranging the sentence in this way significantly increases its acceptability, but von Fintel’s account, or any obvious variation on it, would predict that such rearrangement would not impact the sentence’s acceptability. Unless we assume that moving the clauses to the left periphery alters the truth conditions of the statement, a truth conditional explanation of the permissibility of coordination will not be forthcoming.¹¹

¹¹ I am indebted to John Hawthorne for bringing this phenomenon to my attention.

We should also not lose sight of the lingering phenomenon that von Fintel (1992, 1994) points to – namely that “unless” clauses can be disjoined, as in:

(20) I won’t go to the party unless Bill comes, or unless there is free beer.

These data together suggest that the behavior of “unless” clauses in coordinate structures needs considerably more investigation before it will be understood. Positing a uniqueness clause does not provide us with the explanation we seek. While more work is certainly in order, the phenomenon of coordination does not provide us with a reason to prefer von Fintel’s account to mine.

4.4 “Unless”-Statements Embedded under Quantifiers

We have formulated a promising account of “unless”-statements that occur with quantificational adverbs. The “unless”-statements serve to restrict the range of possibilities that fall under the domain of the adverbial quantifiers. It is not difficult, then, to extend this account so that “unless”-statements embedded under quantifying determiners serve to restrict those quantifiers. As in the case of “if”-statements, a modal element must be introduced into the semantics, and I will continue to do so by means of a wide-scope covert universal quantifier over possible worlds.

I propose that we analyze quantified “unless”-statements of the form “Q Ns M, unless they R” by letting the antecedent R restrict the quantifier in the same way that it restricts an adverb of quantification in the account provided above. The logical form of such a statement would then be:

$$\forall w Cw, w_0: Qx [Nx - Rx] [Mx] \& Qx [Nx \& Mx] [Nx - Rx]$$

Or more perspicuously:

$$\forall w Cw, w_0: Qx [Nx \& \text{not } Rx] [Mx] \& Qx [Nx \& Mx] [Nx \& \text{not } Rx]$$

We will thus treat, e.g., “no student will succeed unless they work hard” as:

$$\forall w Cw, w_0: \text{No } x [x \text{ is a (relevant) student in } w \& x \text{ does not work hard in } w] \\ [x \text{ succeeds in } w] \& \text{No } x [x \text{ is a (relevant) student in } w \& x \text{ succeeds in } w] \\ [x \text{ is a (relevant) student in } w \& x \text{ does not work hard in } w]$$

We correctly predict that Meadow’s presence is enough to falsify the claim, no matter how hard she in fact decides to work. Since there is a relevant possible world in which Meadow succeeds without working hard, the statement is false. Let us recall the intuition we had earlier: that “unless”-statements embedded under “no” seem to be in some sense equivalent to “if . . .not”-statements. We had no sense that there was a uniqueness clause making a contribution to the truth conditions of the statement. On this analysis we can understand why this is so. Just as “unless”-statements that contain the quantificational adverb “never” seemed to be equivalent to “if . . . not” statements, the same is true for ones embedded under “no”, since in both cases the quantifiers are symmetric, and so the uniqueness clause does not add any additional demands to the truth conditions.

We are also able to predict at last that both the over-protected Meadow and the unfortunate Bill suffice to falsify “every student will get an A unless they goof off”.

$\forall w Cw, w_0$: Every x [x is a (relevant) student in w & x does not goof off in w]
 $[x$ succeeds in $w]$ & Every x [x is a (relevant) student in w & x succeeds in $w]$
 $[x$ is a (relevant) student in w & x does not goof off in $w]$

No matter how hard Bill and Meadow actually work, the statement is false if either is among the relevant students. The statement is false in Bill’s case because there is a relevant possible world in which he does not goof off and yet does not succeed, and so the first conjunct of the analysis is false in that world. It is false in Meadow’s case because of the relevant possibility of her succeeding without working hard, and so falsifying the second conjunct in that world.

“Unless”-statements containing “most” and “few” can be given a parallel analysis. “Most students will succeed unless they goof off” will be analyzed as:

$\forall w Cw, w_0$: Most x [x is a (relevant) student in w & x does not goof off in $w]$
 $[x$ succeeds in $w]$ & Most x [x is a (relevant) student in w & x succeeds in $w]$
 $[x$ is a student in w & x does not goof off in $w]$

Similarly, “few students will succeed unless they work hard” is to be analyzed as:

$\forall w Cw, w_0$: Few x [x is a (relevant) student in w & x does not work hard in $w]$
 $[x$ succeeds in $w]$ & Few x [x is a (relevant) student in w & x succeeds in $w]$
 $[x$ is a student in w & x does not work hard in $w]$

We have thus managed to provide an account of quantified “unless”-statements that adequately captures their truth conditions, without attributing a chameleon-like semantics to “unless”. Quantified statements containing “unless” could not be understood as attributing conditional properties to a particular number or proportion of items in a domain restricted by the relevant nominal, as the Simple Solution would have it. They can, however, be analyzed by way of taking the “unless”-statement to restrict the quantifier, albeit in a somewhat complex manner. We have seen, though, that an adequate account of unquantified “unless”-statements extends naturally to accommodate quantified “unless”-statements.

5 Conclusion

A uniform semantic analysis of “if” and “unless” embedded under quantifiers is possible. These constructions thus do not pose a threat to the thesis that natural language is semantically compositional. The semantics of these statements is not, however, a straightforward matter. The Simple Solution to Higginbotham’s puzzle – according to which their truth and falsity depend on the number or proportion of the relevant items that satisfy the open conditional – ran into difficulty when we considered “if”-statements that do not obey Conditional Excluded Middle, and was wholly unable to deal with “unless”-statements. Ultimately, we found that both types of conditional ought to be treated as restricting their quantified NP subjects, while

also contributing a covert modal element to the semantics. Given the complexity of the analysis required to give adequate truth conditions for these constructions, it is hardly surprising that theorists have doubted that a successful, uniform account of them would be possible. I hope to have shown that, despite these doubts, we can indeed provide such an analysis.

Acknowledgment Many thanks to John Hawthorne, Richard Larson and Paul Pietroski, Christopher Viger, and an anonymous reviewer for extremely helpful discussion and comments on an earlier draft of this paper. A much earlier version of the paper was presented in 2003 to the Australian National University Philosophical Society, and at the Carleton University Conference on Recent Research in Semantics, under the title of “Compositional Conditionals”. I am indebted to both audiences for their questions and comments.

References

- Barwise, J., & Cooper, R. (1981). Generalized quantifiers and natural language. *Linguistics and Philosophy*, 4, 159–219.
- Geis, M. (1973). *If and Unless*. In B. Kachru, R. Lees, Y. Malkiel, A. Petrangeli, & S. Saporta (Eds.), *Issues in linguistics: Papers in Honor of Henry and Renee Kahane* (pp. 231–253). Urbana: University of Illinois Press.
- Geurts, B. (m.s.). *On an ambiguity in quantified conditionals*. University of Nijmegen.
- Heim, I. (1990). E-type pronouns and donkey anaphora. *Linguistics and Philosophy*, 13, 137–177.
- Higginbotham, J. (1986). Linguistic theory and Davidson’s program in semantics. In E. Lepore (Ed.), *Truth and interpretation: Perspectives on the philosophy of Donald Davidson* (pp. 29–48). Oxford: Basil Blackwell.
- Higginbotham, J. (2003). Conditionals and compositionality. In J. Hawthorne (Ed.), *Philosophical perspectives*, 17, pp. 181–195.
- Janssen, Theo M. V. (1997). Compositionality. In J. van Benthem & A. ter Meulen (Eds.), *Handbook of logic and language* (pp. 417–475). Amsterdam: Elsevier.
- Kratzer, A. (1989). An investigation of the lumps of thought. *Linguistics and Philosophy*, 12, 607–653.
- Kratzer, A. (1991). Conditionals. In A. von Stechow & D. Wunderlich (Eds.), *Semantics: An international handbook of contemporary research* (pp. 651–657). New York: de Gruyter.
- Leslie, S.-J. (2003a). *Compositional conditionals*. Paper presented to the Australian National University Philosophical Society.
- Leslie, S.-J. (2003b). *Compositional conditionals*. Paper for the Conference on Recent Research in Semantics, Carleton University.
- Lewis, D. (1975). Adverbs of quantification. In E. Keenan (Ed.), *Formal semantics of natural language*. Cambridge: Cambridge University Press.
- Pelletier, F. J. (1994a). On an argument against semantic compositionality. In D. Prawitz & D. Westerstaal (Eds.), *Logic and philosophy of science in uppsala* (pp. 599–610). Dordrecht: Kluwer.
- Pelletier, F. J. (1994b). The principle of semantic compositionality. *Topoi*, 13, 11–24.
- Von Fintel, K. (1992). Exceptive conditionals: The meaning of ‘Unless’. *Proceedings of the North East Linguistics Society*, 22, 135–151.
- Von Fintel, K. (1994). *Restrictions on Quantifier Domains*. Ph.D. Thesis, University of Massachusetts, Amherst.
- Von Fintel, K. (1998). Quantifiers and if-clauses. *The Philosophical Quarterly*, 48(191), 209–214.
- Von Fintel, K. & Iatridou, S. (2002). *If and when if-clauses can restrict quantifiers*. Paper for the Workshop in Philosophy and Linguistics, University of Michigan.
- Szabó, Z. G., Compositionality. In Edward N. Zalta (Ed.), *The stanford encyclopedia of philosophy (spring 2005 edition)*, <http://plato.stanford.edu/archives/spr2005/entries/compositionality>.

Bridging the Paratactic Gap

Daniel Blair

Abstract One common objection to Donald Davidson's view of indirect discourse sentences is that the logical grammar that he argues for is inadequate to handle a large number of semantic and grammatical phenomena that otherwise naturally occur in such contexts, most famously the binding of pronouns by quantificational antecedents. This paper modifies some aspects of Davidson's view in such a way as to permit such phenomena to occur while preserving the underlying idea of 'parataxis.' I focus on the case of binding. I sketch some extensions of the proposed modification and discuss some of the more far reaching consequences of modified theory.

Keywords Binding · Donald Davidson · grammar · indirect discourse · intensional contexts · logical form · parataxis · pronouns · quantifiers · sentence

1 Introduction

This paper defends and elaborates Donald Davidson's theory of indirect discourse and propositional attitude sentences, the so-called 'paratactic theory' of indirect discourse sentences originally presented in his classic article 'On Saying That,' focusing on one sort of objection. The objection, in its most general form, is this: however suggestive Davidson's theory may be as a theory about the semantics of indirect discourse and propositional attitude sentences, it is a non-starter as a theory concerning the properties of propositional attitude and indirect discourse sentences as they appear in natural language. My argument will be developed via a consideration of one classic expression of this concern originating in James Higginbotham's seminal 1986 essay 'Linguistic Theory and Davidson's Program in Semantics.' In many ways, this objection is both the most technically and philosophically challenging objection the theory faces since it is a matter of showing how anything like quantification and binding is so much as *possible* within the set of assumptions

D. Blair
810 Whittier PL NW, Washington DC 20012, USA

Davidson minimally needs to get his theory off the ground,¹ if they are to be captured as recognizably *semantic* phenomena.² I shall argue that what is widely judged to be the most toxic aspect of Davidson's analysis actually can be preserved in a form which sustains Davidson's core idea that the content of a propositional attitude or indirect discourse sentence is projected from the content of more than the logical form of a *single* sentence.

The paper is structured as follows: first, I briefly review the analysis presented in Davidson's article and the objection that is the focus of this paper, then I will show how one might elaborate a notion of 'parataxis' in such a way that the core idea of Davidson's treatment is sustained, albeit with a more refined view of the relation between logical form and grammatical structure than anything found in Davidson's essay.

There is a more general point that I hope emerges. Davidson's view has been taken to be the poster child of philosophical theories of language that are almost comically unhinged from empirical reality. While there is some truth to that assessment, the situation is more complicated than has been acknowledged and it is complicated in interesting ways. While Davidson's original view is almost certainly false, thinking about how one might construct a successor notion of 'parataxis' can lead us to consider some aspects of linguistic structure and form that receive little if any consideration in the philosophical literature. While attention to empirical detail is essential in the philosophy of language, it is just as important to be attentive to the rather wide range of ways that exist for understanding the relation between grammar and meaning. The grammar of natural language does not, by itself, resolve the issue between Davidson and his opponents.³

2 Davidson's Analysis

I want to start off by clarifying what the paratactic theory actually is and outlining the analysis Davidson actually gives. Davidson's analysis is given in his characteristic way, an idea of ingenious simplicity, wittily and succinctly presented after a survey of failed attempts to address a longstanding problem. I shall state the problem, as Davidson does, as one involving non-substitutability of proper names within the embedded complement clauses of verbs of propositional attitude and indirect speech. I shall foreground considerations of grammatical context with a view to later discussion.

¹ This is not to say that, without addressing the complaint at issue here, Davidson would not be able to account for quantification and pronominal binding at all. There are many contexts in which intensionality is not at issue.

² And not, for example, pragmatic or discourse related features of a sentence. See Botterell and Stainton (2005) or Lepore and Loewer (1989).

³ See Ludwig and Ray (1998) for a listing of some of the other objections to the theory. I shall have (virtually) nothing to say about those objections here.

Statement 1, together with 2, licenses the conclusion in 3, given a suitable formulation of Leibniz' law:

1. John walked to the store.
2. John = Johann
3. Johann walked to the store.

3 is true if 1 and 2 are. Suppose that we add to 1–3 a premise such as 4:

4. [S [NP John] [VP walked to the store]] is true if and only if John walked to the store.

Premise 2 can also be given an explicitly linguistic parallel:

5. Ref [NP John] = Ref [NP Johann]⁴

Read this as saying that the reference⁵ of the linguistic expression [NP John] is the same as the referent of the linguistic expression [NP Johann]. Given the relations between 1, 2 and 3, it would appear that replacing [NP John] with [NP Johann] does not affect the truth (or falsity) of the sentences containing the latter. 6 seems equivalent to 4 in every respect for which the question of truth is important:

6. [S [NP Johann] [VP walked to the store]] is true if and only if John walked to the store.

6 differs from 4 only in the substitution of one NP for another which has the same semantic value.

Consider a simplified structure of the material surrounding the subject NP of 1, 4 and 6:

7. [S [NP Subj] [VP Pred]]

Given what was seen above, it seems that the truth-value of a sentence having the grammatical structure displayed in 7 is indifferent to substitutions in the position occupied by '[NP Subj]' when those substitutions do not change the semantic value of the expressions occupying that position.

When a broader range of constructions are considered, it appears that there are grammatical contexts for which a guarantee of referential sameness in one position does not guarantee sameness of truth-value overall. Consider

8. Galileo said that the earth moves.

⁴ More fully, axioms for proper names might be: Ref[x, [NP John], σ] ('x is the referent of [NP John] relative to a sequence of objects σ iff $x = \text{John}$ '). Talk of sequences is important when considering quantification and I shall discuss that further below.

⁵ Or 'semantic value' in the terminology of Larson and Segal (1995).

Supposing for a moment that all noun phrases⁶ have referents, consider the following premise:

9. Ref [_{NP} The earth] = Ref [_{NP} The only planet broadcasting baseball]

Using this identity to substitute NPs with the same reference in the context:

10. [_S [_{NP} Galileo] [_{VP} said [_S that [_{NP} the earth] [_{VP} moves]]]]

does not preserve truth; 11 seems clearly false:⁷

11. [_S [_{NP} Galileo] [_{VP} said [_S that [_{NP} the only planet broadcasting baseball] [_{VP} moves]]]]

So much is numbingly familiar. Substitution need not preserve truth when the NP in the position occupied by **SUBJ**₂ in 12 is exchanged for an NP that may differ from it only in non-semantic ways:

12. [_S [_{NP} **SUBJ**₁] [_{VP} Pred₁] [_S that [_{NP} **SUBJ**₂] [_{VP} Pred₂]]]]

Although it still appears that substitution is fine with respect to the position occupied by **SUBJ**₁, just as it is in 7.

One of the characteristics of the problem is that different sorts of structures are involved in the two cases. 12 involves one sentence embedded inside of another and it is within this embedded sentence that **SUBJ**₂ appears. It is a natural enough thought that it is somehow the nature of the *structure* in 12 – the formation and embedding of one clause within another clause – that is responsible for the behavior we are concerned with here. When a ‘that-clause’ is used as an argument, or is apparently the argument of a verb like ‘say,’ problems arise.⁸

There are many, many responses to those problems, but Davidson’s response is unique in denying that a sentence like 8 above actually is a single context upon which substitution can be performed:

The proposal . . . is this: sentences in indirect discourse, as it happens, wear their logical form on their sleeves (except for one small point). They consist of an expression referring to a speaker, the two place predicate ‘said,’ and a demonstrative referring to an utterance. Period. What follows gives the content of the subject’s saying, but has no logical or semantic connection with the original attribution of a saying: This last point is no doubt the novel one, and upon it everything depends: from a semantic point of view the content sentence in indirect discourse is not contained in the sentence whose truth counts, i.e., the sentence that ends with ‘that.’⁹

⁶ Or perhaps Determiner Phrases (DPs), which embed Noun Phrases in many contexts. I shall stick to the more familiar idea that it is NPs that have referents here and below; see Longobardi (2001), for discussion.

⁷ I am ignoring here those theorists who have doubted whether a semantic theory should give much credence to the intuition that substitution really fails here.

⁸ Although not all such predicates show the same properties, e.g., ‘It is true that P.’ Not all uses of the ‘that clauses’ induce opacity.

⁹ Davidson (1968), 106.

If there is no sentence, or, better, no linguistic expression assigned a logical form corresponding to the *whole* of ‘Galileo said that the earth moves,’ then there is no worry about substitution being licit in *it*. The question simply doesn’t arise. Logical forms and the formulae upon which logical operations are defined are to be assigned to something, no doubt. But what Davidson denies is that what those formulae and logical forms are assigned to are what would most usually be termed ‘the sentences’ of a natural language.¹⁰ His contention is that for the purposes of constructing a theory of meaning for a natural language, there is no single item that contains both the content clause and the reporting clause.

The other part of Davidson’s view is that the ‘that’ appearing after the verb ‘say’ – what is usually called a complementizer, a C^0 (**THAT_C**) – is the nominal demonstrative ‘that’ (**THAT_{DEM}**). While not much has been made of the semantic properties, if any, of complementizers, the semantic properties of demonstratives are, of course, one of the most contested areas in the philosophy of language.¹¹ Davidson contends that the homophony between the two items is more than surface deep. They are the same item. It follows from this that the complementizer has a semantic value like that of other instances of **THAT_{DEM}**. The referent of this demonstrative, according to Davidson, is the utterance of the clause which follows it, ‘the earth moves’ in the example above.¹²

Combining these two theses, *any* change in the clause that follows the demonstrative, e.g., substitution of co-referring names, pronunciation, etc., will change its referent. And if the referent of the demonstrative changes, then the truth-value of the reporting clause may as well. Indeed, there need be *no* change in the linguistic properties of the sentence type tokened at all, semantic and non-semantic properties inclusive, even after the semantic values of all context sensitive expressions has been determined: simply in virtue of being two distinct tokenings of the same ‘sentence,’ the referent of the demonstrative in the reporting clause changes.¹³

Note that there are essentially *two* routes by which substitution is blocked within Davidson’s theory: first, there is no linguistic context upon which logical operations can be defined that includes *both* the reporting and content clauses and so there is no sentence within which substitution fails; secondly, there is the demonstrative part of the theory, which makes the semantic contribution of one item of the reporting clause sensitive to factors other than the semantic properties of the items within the content clause. Each factor is independent of the other. One could have a paratactic theory, with the benefits claimed, without the contention that the complementizer is

¹⁰ Or, more precisely, *root* clauses.

¹¹ I am assuming that the complementizer is ‘that.’ There are other views. One might suppose, following Pesetsky and Torrego (2001), that complementizers are always unpronounced and that the ‘that’ following verbs like ‘say’ is a kind of tense element.

¹² Reference to utterances can be problematic for a variety of reasons. As mentioned in the introduction and further below, I am focused on one objection and am putting to one side a good many other objections.

¹³ A point stressed by Tyler Burge (1986).

a demonstrative or even that there is a demonstrative in the reporting clause at all. Or one could forego separating the clauses and keep the demonstrative part of the theory in any of several different ways. There is room for keeping one of the parts of Davidson's analysis while jettisoning the other, especially if one of those parts becomes untenable for one reason or another.¹⁴

2.1 *Some Caveats*

The utterance of the reporting clause and the content clause are part of the same *speech act*, the same assertion. While it might be useful to think of the utterance of the content clause as utterly external to the speech act that consists in the assertion of the reporting clause, nothing requires this. One can be misled into thinking that it must be separate by thinking about how demonstrative reference works in general: the demonstratum is separate from the demonstration. 'That car is going too fast' does not require, for its truth or even its legibility, that the car somehow be a part of the action or of the sentence itself. But even if it is not generally true that the referents of demonstratives are parts of the same assertion which includes them, they sometimes are: 'I know you're going to hate me for saying **this**, but those pants make you look fat.' The demonstrative in bold here refers to a subpart of the whole assertion. The reporting clause(s) and the content clause(s) are two parts of the same action, the set up and the follow through, so to speak.

In the interests of being complete, it is also worth making a short digression to say something about the notion of 'samesaying.' The reason why is that it appears several times in the development of Davidson's view, playing a substantial role in elucidating the analysis, even if its role in the formal development of view is unclear. Its role in the paper does reveal aspects of Davidson's view that routinely get lost in other expositions. According to Davidson, one does not make one's self a samesayer *simply* by demonstrating an utterance, as if the analysis applied indifferently to both attitude ascriptions and pointing at a message on a blackboard. One makes one's self a samesayer *by uttering the content clause*. No such requirement is in force for simple demonstrations of content, such as pointings at signs. While I can say what John said by demonstrating, e.g., a transcript of his utterance, I will not thereby have made myself a samesayer with John. But I can do so by reporting the content of John's saying by saying, for example: John said that it was raining. Davidson's otherwise cryptic contention that attitude sentences and indirect reports are 'performative' makes sense in this respect. And the utterance that makes the speaker a samesayer with his target is the one that directly follows the demonstrative. There isn't room for one to utter 'Galileo said that the earth moves' while demonstrating a token of 'Snow

¹⁴ For discussion of a version of sententialism that preserves reference to linguistic form but abandons 'parataxis,' see Higginbotham (2006).

is white' with the demonstrative, not at least on the view that Davidson actually gives.¹⁵

This suggests, at least, that there are some differences between complementizers and demonstratives even if one accepts that complementizers are context sensitive referring singular terms. It also suggests that the relationship between the successive clauses of an indirect discourse report is fairly intimate one, even accepting the paratactic aspect of the theory.

One way to understand the latter point is take both the content clause and the reporting clause to be assigned logical forms within the same theory of meaning (with a caveat to be noted shortly). To do so, it would appear that they must be constructed from linguistically homogenous elements, i.e., that they must be formed from the same grammatical and lexical resources.

The familiar words coming in the train of the performative of indirect discourse do, on my account, have structure, but it is familiar structure and poses no problem for the theory of truth not there before indirect discourse was the theme.¹⁶

The elements within the content clause must themselves be capable of appearing un-embedded as well as embedded. While *having* a referent is a semantically significant fact, *being* a referent isn't. Another reason why both clauses should be capable of appearing both in embedded and un-embedded contexts is that the content clause itself might contain reference to another utterance, as in multiply embedded reports like 13:

13. John believes that Mary said that Bill thinks everyone is out to get him.

Here are the expressions assigned logical forms within 13:

- 13a. John believes **THAT₁**
- b. Mary said **THAT₂**
- c. Bill thinks **THAT₃**
- d. Everyone is out to get him

Logical forms will be assigned to each of 13a–d. The paratactic theory ought to apply to these cases just as they do cases of single embedded clauses.¹⁷ If that is right, then whether or not something is a content clause or a reporting clause is a *relational* matter, a matter of where it sits with respect to its syntactic neighbors.¹⁸

¹⁵ See Botterell and Stainton, *op cit*.

¹⁶ Davidson (1968), p. 108.

¹⁷ The speaker has to say something to provide a value for the demonstrative. The importance of the notion of samesaying is something I have left for another occasion but it is important here to recall that the complex notion of 'making myself a samesayer' sums up a way of responding to an objection to previous theories of propositional attitude sentences that occupies fully half of Davidson's essay.

¹⁸ And so the designation 'content clause' and 'reporting clause' are merely heuristic: they are not part of the theory itself.

Things are a bit more intricate when multiply embedded clauses are involved. For while the individual parts are assigned logical forms, it does not follow that the semantic values of the complementizers or demonstratives are somehow supplied by the theory of logical form. A case in which there is only one embedded clause can make it appear as though the referents are tokens of items that have their own logical forms. But thinking about multiply embedded clauses shows that matters are more complex than this. The referent of **THAT**₁ in 13a is not the utterance in 13b. Rather it is something that itself does not have a single logical form, i.e., it is an utterance of ‘Mary said that Bill thinks that everyone is out to get him.’ From the point of view of the truth conditions of 13, what we have is the following for the semantic values of the demonstratives:

- 13a'. Ref[**THAT**₁, ‘Mary said that Bill thinks that everyone is out to get him’]
- b'. Ref[**THAT**₂, ‘Bill thinks that everyone is out to get him’]
- c'. Ref[**THAT**₃, ‘Everyone is out to get him’]

Except for c', none of the referents here are linguistic structures which are given logical forms nor are they tokens of items which are given logical forms. It is true, however, that all of the individual expressions of 13 are items that have logical forms. The demonstratives refer to linguistic expressions that have been uttered in a context.

2.2 To Sum Up

I have now gone over Davidson's view in some detail to show how, exactly, the theory is supposed to work and what is left unspecified. This will help show why exactly the theory is vulnerable to some objections and not to others. The formulation of Davidson's theory has three components: the notion of parataxis, the idea that complementizers are demonstratives and the notion of samesaying, or, more specifically, the notion of making one's self a samesayer. I have not said very much at all about the last notion, and will not be elaborating on it further below. I have responded to a few criticisms of Davidson's theory along the way with the idea of playing up those features of Davidson's theory that specifically have to do with the notion of logical form and grammatical context, and what sorts of expressions are to be given unitary, single logical forms. Davidson's theory is rather underspecified, as is perhaps already clear. But it should also be clear that this leaves a lot of room for developing the theory in different ways, not all of them obvious.

3 Binding Problems

To have ameliorated some of the problems that indirect discourse and propositional sentences present for semantics by showing them to be essentially epiphenomena of the way logical form is individuated in relation to grammatical structure on the basis of a surprisingly simple, if not completely intuitive set of assumptions, would have been a neat trick. If only it had worked.

Higginbotham (1986) pointed out some phenomena, apparently unrelated to intensionality, that present problems for Davidson's view of how sentential complementation works. Consider the following sentence and the graphical representation of its paratactic analysis:

14. Every boy said that he is happy.

15. Every boy said **THAT**.

→ He is happy

Concentrate on the following interpretation of 14: Every boy said that he himself is happy. In more familiar terms, suppose that the pronoun in the 'that-clause' of 14 is *bound* by subject quantifier phrase 'every boy.' The problem is that that reading doesn't seem to be available in 15 and, despite its informality, 15 is the parsing apparently forthcoming from the premises of the paratactic theory.¹⁹ Once separated from its binder by the gap, the pronoun seems to have only a single type of interpretation, the singular, deitic reading. Underlying syntactic structure appears to be a necessary condition for the availability of bound readings.²⁰

¹⁹ I say 'apparently' because the missing reading seems much more readily available, intuitively at any rate, if one adopts Davidson's more prolix analysis. That would yield:

i. $[\forall x: \text{boy}(x)][\exists y] [x \text{ said } y \text{ and my next utterance makes us samesayers.}] \text{ He is happy}$

That is: for every boy, there is something he said and my next utterance makes us, he and I, samesayers: He is happy. As long as each boy said something which samesays the demonstrated utterance of mine, then the reading in question is available. i. is then 'definitionally abbreviated' back to 15 above. If 15 is understood in this latter way, then we can set aside the claim the 'that' that appears in 15 just is the demonstrative and take it to be a defined term of the theory.

²⁰ More precisely, a pronoun can be bound by a quantificational antecedent if that pronoun is c-commanded by that quantificational antecedent, where c-command is a relation that is defined over a single syntactic structure and the categories which contain the pronoun as well as the antecedent. There are well known exceptions to this principle and I am ignoring virtually all the intricacies of binding phenomena. For recent work, see Safir and Büring.

A few further points are worth mentioning. First, I am assuming here that the pronoun in the demonstrated clause in a case like

i. 'Jan thought **THAT**, he was late'

does not present the same problems that a pronoun bound by a quantificational antecedent would. Intuitively, the pronoun can have the same semantic value as the subject NP in the first clause. The extent to which coreference relations are syntactically encoded is an open question; see again Safir and Büring.

Secondly, quantifiers can sometimes at least appear to bind pronouns intersententially, as an anonymous reviewer stresses. Indeed, it may well be the case that binding is freely available across sentential borders and even without grammatical subordination and that it is the cases where binding is *not* available that is the concern of *grammar* and of how *grammar* interacts with binding. The theory that I am developing here can be used to say something about those negative conditions, as conditioned by the grammar of a language. I discuss this further below.

The lack of a syntactic bridge between clauses disrupts not just the binding of nominals but any sort of linguistic relation that is defined over structures that incorporate structures larger than single clauses.²¹ So more than just cross-clausal binding is problematic. Once the bridge has been cut, these semantic relations are jeopardized as well. Any semantic relation between expressions X and Y in this grammatical context:

$$\dots X \dots [_{CP} \dots Y \dots$$

is compromised by the very premises Davidson deploys to account for other semantic phenomena. Severing the clause into pieces inoculates propositional attitude sentences from undercutting extensionality but at a rather steep price, one that cannot be met if the theory is to capture the data at issue.

This isn't the only sort of worry one might have about parataxis. The core of Davidson's analysis is that there is a kind of 'semantic mitosis' involved in the analysis of indirect discourse sentences: an indirect discourse report consists of two (or more) semantically independent units, the borders between them being delineated by a complementizer. With this idea on board, there simply is no way to define, in terms of the structure of natural languages sentences, a grammatical context upon which substitution can be denied. The problem is simply unstateable in a theory of the logical forms of natural language sentences incorporating this assumption. One might worry here that this is an altogether too *easy* – or perhaps too *glib* – way to make the problems disappear though. To be told, at the end of the day, that intensionality is something of an illusion brought about by a faulty way of parsing sentences is unsatisfying at best. The problems have simply been defined away. There now simply is no context like 12 (repeated below) to which one needs to assign a logical form:

12. [_S [_{NP} **SUBJ**₁] [_{VP} Pred₁ [_S that [_{NP} **SUBJ**₂] [_{VP} Pred₂]]]]]

How worrisome is this? One should bear in mind what a theory incorporating 'parataxis' could plausibly entail. Applied across the board to all 'that clauses,' one might think substitution failure ought to be found everywhere. And it isn't. But parataxis of the sort that parses multi-clausal sentences into sets of semantically 'independent' units doesn't require that it should be either. It merely *allows* that intensionality can arise in virtue of the *structure* of sentences. Whether it *does* arise in a particular case, for example, the truth predicate, also involves consideration

Lastly, I take it that Higginbotham's intuitions about his example are correct. Combining syntactic parataxis with demonstrative reference blocks the availability of bound readings.

²¹ For example, any relation of scope between two quantifiers in different clause seems to be problematic, not to mention the use of WH-phrases that have been extracted from deeply embedded clauses, negative polarity licensing and so on. I discuss this briefly below. Thus, even if one doubted that quantificational binding really were a matter for a theory of sentential grammar to say much about, relations of syntactic subordination surely are something which a theory of the interaction of grammar and semantic structure must address.

of the meaning of the truth predicate and how the logic of truth presupposed by a theory of truth is applied to the structures of natural language sentences.

Parallel remarks apply in the case of other predicates such as 'believe' and 'know.' The logical anomalies that initially motivated the paratactic theory end up being a confluence of different factors, a good proportion of which have to do with our theories of belief, of speech act content, and so on, in addition to theories of natural language structure. As dramatic as parataxis might seem on its own, even more dramatic is the way in which the burden of explaining certain anomalies within philosophical semantics is shifted away from the concerns of linguistics per se.

So while there appear to be quite a few examples of theorems like

16. [_S Snow is white] is true if and only if snow is white.

to be derived within a Davidsonian-style theory of truth for arbitrarily many different types of sentences, it appears to be a consequence of the paratactic view that nothing parallel holds with respect to 'Galileo said that the earth moves.' If so, then there will be no theorem of the form:

17. [_S Galileo said [_S that the earth moves]] is true if and only if Galileo said that the earth moves.

Nothing like 17 can be derived within the same theory of truth for which there is a theorem for 16. And if Davidson is right, this is not a *defect* of the theory. One can repeat the same lesson for the great variety of other propositional attitude sentences, including multiply embedded sentences and the other types of 'that clauses' that give rise to intensionality.²² Grammatical structure and logical form are apparently radically divorced.²³

Alongside Higginbotham's problematic cases, there are also substantial hurdles for the idea that complementizers and demonstratives are identical, an idea whose radical-ness has not passed unnoticed.²⁴ I will not be concerned with the problems associated with it too much here, e.g., the homophony between complementizers and demonstratives is limited and is not as cross-linguistically robust as one might like and one can typically delete complementizers without affecting the felicity of an utterance in a way that is impossible with nominal demonstratives. But, in contrast to the issues regarding 'parataxis,' there are some more or less obvious ways of defending this aspect of the view without compromising the spirit of the analysis.

²² If one adopts the paratactic across the board, the types of linguistic expressions that are given theorems within a semantic theory are roughly parallel to the 'kernel sentences' in Chomsky's 1955/1975. In the later case, kernel sentences are joined into larger units via 'Generalized Transformations.' What we need now is a way of joining the kernel logical forms together in a way that allows for binding.

²³ This consequence, all by itself, has been enough for some to reject the theory, cf. Ręcanski (2000). Below I shall argue that there are ways of understanding separating clauses that is not as theoretically naïve as some have thought.

²⁴ Segal and Speas, 1986 detail the many problems, some obvious, some more subtle.

In the end, the claim about the identity of ‘**THAT_{DEM}**’ and ‘**THAT_C**’ is too strong. Davidson states what he really *requires*:

What follows the verb ‘said’ has only the structure of a singular term²⁵ . . .

It must be a singular term such as a proper name, or a variable, and not a semantically complex expression such as a definite description, and, like the demonstrative ‘that,’ must have the potential for different semantic values for each of its occurrences. It is only this weaker claim that Davidson requires, even if claims a good deal more.

Here is how Davidson himself explains the nature and role of the element that follows that predicate:

Imagine an altered case. Galileo utters his words ‘Eppur si muove,’ I utter my words, ‘The earth moves.’ There is no problem yet in recognizing that we are samesayers; an utterance of mine matches an utterance of his in purport. I am not now using my words to refer to a sentence; I speak for myself, and my words refer in the usual way to the earth and to its movement. If Galileo’s ‘Eppur si muove’ made us samesayers, then some utterance or other of Galileo’s made us samesayers. The form ‘ $(\exists x)(\text{Galileo's utterance } x \text{ and my utterance } y \text{ makes us samesayers})$ ’ is thus a way of attributing to any saying I please to Galileo provided I find a way of replacing ‘y’ by a word or phrase that refers to an appropriate utterance of mine. And surely there is a way I can do this: I need only produce the required utterance and replace ‘y’ by a reference to it.²⁶

Taking ‘say’ to be a two-place predicate with an argument x for sayers and an argument y for things said, Davidson’s thought is that one should not substitute ‘[that S]’ for the variable y . ‘[That S]’ is not a term or even a constituent of an indirect speech report for Davidson.

But this still leaves us with the paratactic parsing and the consequence that the elements within the content clause do not make a contribution to the compositionally determined content of the reporting clause. Of course, the fact that there are certain sorts of discrepancies between semantics and syntax is well known and the facts reviewed here may just be part of this general phenomena. But whatever the virtues of this response, it still leaves us with the datum above, i.e., that bound readings vanish in the relevant contexts if one follows through on the paratactic parsing of these sentences. Further, radical divorces between semantics and syntax have a way of turning out to be unhappy ones. We are left with no way to explain how meaning depends on grammatical structure in a *regular* way. We simply cannot collapse *intra*-sentential phenomena into *inter*-sentential phenomena and let demonstrative reference be the *only* link between clauses. These problems cannot be finessed – they are not *mere* details: the ability to bind into a ‘that clause’ subordinated by a verb of propositional attitude is not a species of logical structure that can be sacrificed.

To sum up this section: Indirect speech reports exhibiting bound readings between two (or more) expressions across a clausal boundary are disabled along with substitution by the premises of the paratactic theory. There are certainly other reasons

²⁵ OST, p. 108.

²⁶ OST, p. 105.

why one might be worried about the paratactic parsing, although there are more or less plausible responses to those worries. Even one of the more outlandish features of Davidson's analysis can be given a less toxic version without any obvious harm to the overall package. However, the parataxis part is the most important part of the view and it is the most empirically vulnerable as well. The temptation to write off this part of the analysis is thus understandable: since relations of subordination affect not just pronominal binding but also, e.g., anaphoric relations between tenses, negative polarity licensing and much else, it is hard to say what significant part of natural language the theory would apply to.

But there may still be ways of proceeding. Just as one can conceptualize the role of complementizers in various ways that are compatible with the overall thrust of the theory, so one can re-conceptualize how the semantic structure is related to grammatical structure as well.

4 Sentences, Logical Forms and 'Parataxis'

Consider the way that Davidson describes how the paratactic theory solves the problems of opaque contexts.

The appearance of failure of the laws of extensional substitution is explained as due to our mistaking what are really two sentences for one: we make substitutions in one sentence, but it is the other (the utterance of) which changes in truth. Since an utterance of 'Galileo said that' and any utterance following it are semantically independent.²⁷

Indirect discourse sentences come in chunks, as it were: substitution does not affect the truth value or truth conditions of the chunk in which it is performed, but it does affect the chunk on which substitution is not performed. One should keep in mind that even though the content and reporting clauses are, in Davidson's phrase above, 'semantically independent,' this is no hindrance to their being ordered in a fairly rigid way with respect to one another. So much is clear from Davidson's own semi-serious stab at formalizing the suggestion:²⁸

The earth moves. $(\exists x)$ (Galileo's utterance x and my last utterance make us samesayers.)

If we view 'my next utterance' as a kind of rule of interpretation for 'THAT_C' akin to the rules of interpretation associated with indexicals like 'I' or 'here,' then we can get a lot of what is needed from making this informal gloss a bit more specific.

The only thing we can get from the semi-formal gloss though is that there is a certain linear (or perhaps just temporal) order and linear order by itself will not be of much help with the problems at hand. Binding typically involves relations of subordination between a binder and its potential bindees. And there are endlessly many relations of linear order between phrases that do not involve subordination: the successive sentences of a discourse, for example. What is needed is a type of

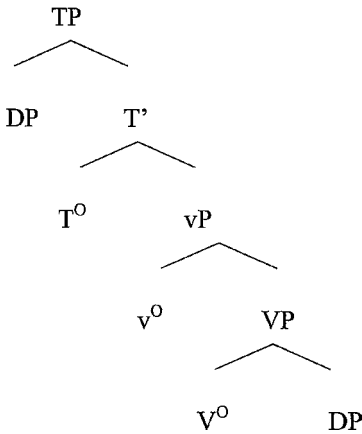
²⁷ OST, p. 107.

²⁸ OST, p. 105.

ordering that captures what is special about subordination. The notion of ‘Parataxis’ is not terribly helpful as it could signify any of a variety of things, e.g., sometimes just a juxtaposition of sentences, or arrangement of linguistic fragments whose combination is not mediated grammatically but rather rhetorically or pragmatically. More importantly, ‘parataxis’ often just means coordination as opposed to subordination.

To see what sort of order might be sufficient, we need to return to the earlier discussion about the grammatical contexts in which intensionality arises. I want to be a bit more precise now about what is meant by ‘that clauses.’ Clauses are usually categorized as Tense Phrases or ‘TPs’ so that the clausal expression ‘snow is white’ is categorized as a tense phrase or a ‘TP.’²⁹ A schematic tree diagram is given below:

18:

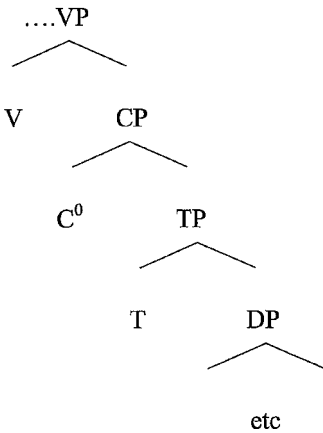


This is a very rough sketch of the structure of a transitive clause, but it should be sufficient for our purposes. The position marked by the first, uppermost DP is what is most commonly called the ‘subject’ position, the lower DP taking up the ‘object’ position. The phrase ‘vP’ may be less familiar, but its presence is closely linked to auxiliaries and other ‘light’ verbs, e.g., the ‘make’ of periphrastic causatives in sentences like ‘That song always makes me cry.’

When a clause is embedded within another sentence, usually as the argument of a noun, verb or adjective, that clause is still a TP, but the whole expression comprising the argument of the verb is now a ‘complementizer phrase,’ or a ‘CP.’ This means that the relevant embedding structure has the properties indicated in 14:

²⁹ This helps with understanding the variety of different kinds of clauses since clauses seem to vary in the kinds of tenses that they have in addition to their mood and argument structure.

19:



The CP layer itself, intervening between TP and V, has internal structure of a complex sort.³⁰ Some of this structure is directly relevant and I shall discuss it further below. Getting back to the main topic, it is the positions underneath the CP node I am concerned with, i.e., C⁰ and TP and everything contained within the latter. It is those positions for which the logical and semantic problems arise, when the CP containing it is itself contained within a VP as a complement to a verb like ‘say’ or ‘believe.’ The grammatical contexts for which substitution is problematic are positions within TP when TP is in construction with the complementizer ‘that.’ Intuitively, while positions within ‘John is here’ permit substitution, those same positions do not permit substitution in ‘That John is here.’

We can rephrase our problem: binding seems to require that binder and bindee are part of the same structure. It is a seemingly necessary condition for the kind of binding in Higginbotham’s example that the bindee must be contained within a structure that is itself contained within another structure, one that also contains its binder. If one supposes that this also means that the logical form given to the whole sentence must also have this same structure, the objection stands since preserving the unity of the syntactic structure within logical form is precisely to deny parataxis. But if both assumptions aren’t required, or at least there is a way of describing syntactic structures and the relationships between clauses when one is subordinated to another that does not require that both clauses make contributions to the compositional semantics of the same logical forms, then one preserves the core idea there is no single logical form underlying canonical indirect discourse reports.

There are more sophisticated notions of order that might help here. There is the old intuition that a sentence like 20a is derived from a structure like 20b:

³⁰ Luigi Rizzi, (1997).

- 20a. Who did John kiss
 b. John kissed who

via the displacement and reordering of elements. A more articulated structure shows this (ignoring tense):

21. [_{CP} Who_i [did [_{TP} John [_{VP} kiss *t_i*]

In this structure, CP is derived from TP via the extraction of 'who' and its displacement at the left most part of the sentence. The extraction must obey certain rules of grammar, of course. But the structure in 21 depends upon the node labeled CP standing in a certain relation to the part of structure labeled TP. Suppose that we think of the structure in 21 as being built up in steps, starting, for example, with the verb 'kiss', and adding 'who' to form a larger more complex phrase, labeled VP. Subsequent additions of material to this structure yield larger and larger phrases, one phrase succeeding the next via the rules of grammar. Then we can think of CP as being the *successor* phrase to TP via the movement of 'who.' I am simplifying a much more complex theory. In general though, ordering relations between elements in linguistic structures are crucial for the development of a syntactic theory for natural language, an idea that goes back 50 years. At the root of the projection of grammatical structure into larger and more intricate structures is an iterated concatenation of linguistic expressions according to syntactic rules.

The example in 20 involves ordering *within* a clause. But the same model can be applied to inter-clausal relations as well. Suppose that concatenating clauses together is a way of ordering them in one or another grammatically well formed way. Usually, such concatenation is a matter of, e.g., supplying a verb with its argument or, the semantic correlate, supplying an argument for a function. Let us say that *within* a clause or a TP, the semantic correlate of concatenation works more or less like function application: thus, agreement and predicate argument structure work as they would seem otherwise to do. But *between* clauses, the semantic significance of concatenation lies in establishing some aspects of the interpretation of the subordinated clauses and in supplying a value for a variable.³¹

Concatenation is not a matter then of supplying an argument for a function so much as it a matter linking the interpretation of one expression to that of another. Before we go any further, one might worry about the idea that concatenation within a clause seems to be something different than it is between clauses. Surely, so basic an operation as syntactic concatenation should be associated with the same basic semantic interpretation, e.g., functional application. All things being equal, don't we want to have same sort of mechanism underlying the combination of all sorts of expressions?

³¹ In fact, on the view I develop below, variables in both clauses are supplied with values: the argument of 'say' in the reporting clause and, by default, the specification for tense in the content clause.

Perhaps. But nearly all applications of the function/argument model require other sorts of departures, even if those departures are usually elsewhere in the theory, a prime example being the various forms of type shifting used to capture the semantic properties of quantifiers in subject and object positions within a sentence.³² I see no real difference in theoretical parsimony between the two approaches on this score. If it is unity that is wanted here, one could encode concatenation as a kind of ordering even within a clause as well, although the details would differ depending upon the properties of the expressions concatenated.

Those details to one side, my task here is to make plausible that there are ways of capturing the grammatical facts about binding into clausal complements of indirect discourse that do not require containment relations between logical forms. The point is not to devise a new theory of grammatical structure or even a new theory of logical form. It is to show that the assumption that the semantic structure of an indirect discourse sentence involves embedding one clause inside of a predicate so that the embedded clause is proper subpart of a larger sentence is strictly unnecessary. In one sense, it is to downplay the role that talk of ‘sentences’ play within semantic theory. But it is not to completely abandon the notion of a sentence. It remains the case that ordering relations between clauses derived from syntactic rules have different semantic properties than do, e.g., the ordering of clauses within a multi-sentence discourse. What is needed is a way of assuring that relations of precedence between linguistic expressions – more specifically, between clauses – are sufficient to capture the semantic properties of subordination.

The question is: can one give descriptions of the *semantic* properties of one clause and its *syntactic* predecessor, so to speak, without assuming that one clause is *contained* within the other? More to the point for my purposes here: can one describe the grammatical facts as involving sequences of syntactic structures without thereby supposing that one expression is *contained* within the other? It matters that they are *connected* in a certain way, of course. One might still talk of containment here, sure enough, but at least in the grammatical case, such talk is doing no work if all the relevant syntactic structure is established in a local, point-to-point way.

Those local steps, in the grammatical cases, aren’t usually visible in the overt structure, but we have reason to think that they are nevertheless real. Thus, for a case like

22. Who did you say that John thinks that Mary will believe that Bill kissed?

there is good reason to think that the left-most expression ‘who’ has been extracted in a series of steps from the most deeply embedded clause so that the structure is more fully revealed in the following:

23. [_{CP} Who₁ did you say [_{CP} who₁ John thinks [_{CP} who₁ Mary will believe [_{CP} who₁ Bill kissed *t*₁

³² See Pietroski (2004) for further discussion.

A long distance relation is composed of an ordered sequence of smaller movements. The chain ending with the left most occurrence ‘who₁’ is anchored in the first clause, and is extracted in a series of movements through each of the intervening CPs. The intervening wh-phrases are rarely, if ever visible:³³ in this case, all but the left most ‘who’ is given overt expression. Nonetheless, there appears to be good evidence for thinking that all extraction proceeds in short, local steps. For example, if one of the intervening points is occupied by another phrase, then extraction isn’t possible:

24. *Who did you say John thinks Mary will believe when Bill kissed.

The presence of ‘when’ at the first step along the way upward to the left edge prevents the long distance extraction. If one could simply extract at will, without looking at intervening structure, such facts would appear puzzling.

The semantic parallel would be this: semantic relations are also established in a point-to-point, local fashion and seemingly long distance relations, such as a binding relation between clauses, are the result of intervening local steps. The intervening links are visible, oftentimes, as complementizers, or at least that is how I want to think of them herein. In the syntactic case, the local relations involved extraction and restructuring of grammatical structure. Binding relations of the kind I am interested in here don’t appear³⁴ to involve extraction but something like the strictly local view of syntactic relations could also be taken as a model of how semantic relations are built up as well. Let us look at an example in a bit more detail to see how talk of ordering clauses with respect to one another might help.

The problem we are addressing is that separating the clauses in the way suggested in Davidson’s text appears to eliminate the syntactic scaffolding which supports the availability of quantificational binding, among other things. Consider again the problematic case and a more conventional representation of the problematic reading:

25. [_{TP} Every boy₁ said [_{CP} that [_{TP} he₁ is pretty smart.]]]

where the subscripts indicate that pronoun is bound by the quantificational expression. As we saw above, the paratactic theory seemingly misses this reading.

Suppose that the reporting clause is an open sentence, the free variable being bound by the quantificational phrase ‘every boy’ and that the sentence final demonstrative understood as a free variable:

25. [Every *x* [boy *x*]] [*x* said *y*]

This sentence is then interpreted relative to sets of sequences of objects in a domain of quantification, in the usual fashion.³⁵ The value of *y* is an utterance of another open sentence, namely:

³³ Though they might be in some cases: see McCloskey (2002).

³⁴ See Kayne (2002) for a different view.

³⁵ See Larson and Segal (1995); Heim and Kratzer (1998).

26. He's pretty smart

where 'he' semantically functions like a variable. Like any other sentence containing a free variable, this sentence is interpreted relative to sets of values or, in more familiar terms, sequences of objects drawn from some universe of discourse. We want the interpretation of this open sentence to be linked, at least potentially, to whatever interpretation we give of the preceding clause.

Let us give a variable for such sequences ' σ ' and assign an index to any variable within the expression whose interpretation is keyed to what is contained within such sequences. A variable having the subscript i is interpreted as denoting, relative to any sequence of objects σ , the i th member of that sequence. Where σ is a sequence of objects $\{a, b, c, \dots\}$, I shall term the n th object in that sequence shall be denoted by ' $\sigma(n)$.' So if $\sigma(n) = a$ say, then a variable x_n is interpreted relative to that σ as denoting a . The subject pronoun in

27. $[[\text{He}_i \text{ is pretty smart}]]^\sigma$

is interpreted by looking at what σ contains in its i th position: if that object is pretty smart, then the sentence is true.

Prima facie, binding would work in the same way for multi-clausal constructions as it does for single clauses. Namely, we would abstract on both the positions occupied by the pronoun as well as that occupied by the quantifier in the subject position of the reporting clause, e.g.,

28. $[[x_i \text{ said that } x_i \text{ is pretty smart}]]^\sigma$

The whole sentence would then be interpreted relative to assignments of certain sequences of objects, each such sequence being an assignment of certain objects to the indexed positions within the sentence.

Let us now suppose that the general contribution of '**THAT**_{C<i>}' (and of subordination generally) is to link the interpretation of the content clause with which it is syntactically associated to that of the reporting clause as follows: whatever sequence σ' of objects interprets the free variables of the content clause is to be understood as a variant of a certain sort of whatever sequence of objects σ interprets the free variables of the reporting clause.

Each indirect discourse report is associated with tokenings of two linguistic expressions, the root clause and the finite indicative clause subordinated by 'say,' and each of these utterances has its usual semantics. In the present context, let us take that to mean that each 'utterance' of an indirect discourse report is evaluated relative to its own context. The reporting clause works like any other unembedded root clause does, e.g., it is evaluated relative to a time of utterance, a speaker and whatever else one includes in a context to evaluate canonical context sensitive expressions. The content clause though has a somewhat different context, one that includes the content clause's *syntactic* predecessor, i.e., the reporting clause, having whatever interpretation it has. If we assume that quantificational structures are interpreted in the way mentioned above, then we can assume that part of what an utterance of such a sentence contributes to its context are just those sets of values

or sequences of objects that interpret it. For the case we're looking at, we have the following pieces:

29. Every boy said that he is pretty smart
 29a. $\sum_{\text{REP}} = [[_{\text{TP}} \text{Every boy said } \mathbf{THAT}_i]]^\sigma$
 b. $\sum_{\text{CON}} = [[_{\text{TP}} \text{He is pretty smart}]]^\sigma$

where \sum_{REP} is the predecessor³⁶ of \sum_{CON} . The effect of ordering clauses in this manner is to link the interpretation of \sum_{CON} to that of \sum_{REP} in the following way: the value assigned to ' \mathbf{THAT}_i ' in 29a by any sequence interpreting \sum_{REP} just is the event of uttering \sum_{CON} : $\sigma(\mathbf{THAT}_i) = u$, where u is the tokening of \sum_{REP} which is immediately preceded by the tokening of \sum_{CON} containing ' \mathbf{THAT}_i .' More specifically, it is \sum_{CON} interpreted via the same sets of values that one uses to interpret the preceding clause. If $\sigma(i) = \mathbf{a}$ in \sum_{REP} , then, for any sequence σ^* interpreting \sum_{CON} , if \sum_{REP} is the syntactic predecessor of \sum_{CON} , then $\sigma^*(i) = \mathbf{a}$.

Intuitively, any value given to the variable subscripted with an i in \sum_{REP} can be 'passed' on to variables marked with i in its syntactic successor, \sum_{CON} . An utterance of a reporting clause makes those interpretations for free variables available and syntactic ordering allows them to be exploited. When the preceding phrase contains a variable that is quantified upon, as it is in the problematic examples motivating this inquiry, and the succeeding phrase contains a pronoun which is bound by that quantifier in turn, then one can use the above idea in the following way: take the index marking the variable in the first clause, the position quantified upon, and pass that index on to the sequences which interpret its successor. Thus to interpret the open sentence

30. $[x_i \text{ said } \mathbf{THAT}_C] \dots$

In the position marked by x_i , one looks at sequences of objects drawn from a domain of discourse. When the variable x is quantified upon by a universal quantifier, as in 31:

31. Every boy x_i $[x_i \text{ said } \mathbf{THAT}_C] \dots$

one looks at every sequence of objects drawn from the domain of discourse that differs at most in what it assigns to the i th position.

To a first approximation, pronouns and variables within the succeeding clause may be indexed with the same index as those within a preceding clause when their containing clauses are ordered in a certain way with respect to each other. Where 31 is evaluated via the assignment of John to x_i : the succeeding content 'He _{i} is happy' is interpreted via an assignment of John as the semantic value of 'He.' The whole is

³⁶ I am using the notion of syntactic predecessor in a way that matches, as close as possible, the linear order in which items appear. It is more common to treat what I am calling the successor clause above as the derivational predecessor, but it is also common to note that the difference between the top down and bottom up strategies within syntax are interchangeable. The top down order seems more useful for semantics.

then true if every such assignment of values to indexed positions is true, that is, if, for every boy, there is an event wherein *he* says this: *he* is happy.

Alternatively put, we can think of the semantics in this way: for every boy x_i , if there's a true utterance of

32. [x_i said **THAT**_C]

then there is a saying of

33. [he_i is happy]

This preserves what needs to be preserved about subordination.

Let's pause for a moment to take stock. The view that I am urging here does not require any special new grammatical trick. I have cast the underlying relation between logical form and grammatical structure in a somewhat different way, so as not to require relations of containment of the kind that would rule out parataxis, virtually by definition. We are, admittedly, at some distance from at least the *rhetoric* of Davidson's theory. In particular, Davidson's talk of an indirect discourse report really being 'two sentences,' while it suggests something like the picture here, also brings invites unwanted implications. The notion of a 'sentence' needs to be refined within the theory of logical form as it already has been in syntactic theory. Further, one can leave the grammar pretty much as it is. The derivation of syntactic structure by itself already affords a natural way of dividing up the chain of linguistic structures so as to generate the right structures for interpretation via relations of precedence and succession.

It is also worth mentioning that while I have not developed here a full theory of binding or scope and a thousand other things, the availability of the bound readings is keyed to precedence relations and the concatenation via the complementizer 'that.' I'll briefly return to the issue of the complementizer below and the motivation for the departure from Davidson's view. But an account like the one I have sketched here doesn't allow just any sort of configuration to permit binding. For example, a sentence like

34. He thinks that everyone will like him

does not have a reading in which the quantifier binds the pronoun in the superordinate clause. And in this case, the inadmissibility of binding is just a reflection of the order of concatenation. The linguistic expression containing the potentially binding quantifier is the successor of the linguistic expression containing its potential bindee. The subordinating complementizer cannot affect the interpretation of the subject pronoun in this case since it does not affect the interpretation of its own containing clause.

Some more complicated cases show that precedence is based upon the structure of a chain of clauses and not merely their overt order. Thus in

35. The man that taught every boy his_i lines said that he_k will be ready for the show.

The quantifier in the relative clause can bind ‘his_{*i*}’ within its minimal containing clause but a bound reading for the pronoun in the content clause seems to be missing. Observe however that the definite description ‘the man’ *can* bind the pronoun in the content clause. If we look at the paratactic parsing though, we can see that there is a reason for why that should be. Here’s a first pass at the paratactic parsing:

- 35a. The man that taught every boy his lines said **THAT_{*n*}**
 b. He will be ready for the show.

It still seems as though the quantifier is in the preceding member of the chain though. But since 35a itself contains a ‘that clause,’ it needs to be given the paratactic parsing:

- 36a. The man **THAT_{*m*}** said **THAT_{*n*}**
 b. x taught every boy his lines

Now the syntactic predecessor of 35b is not 35a; rather, it is 36a, which contains the definite description but *not* the quantifier. The way the chain is parsed paratactically reflects the binding possibilities.

I have just applied the paratactic parsing to a relative clause and not just to a propositional attitude context. It should already be clear that intensionality arises with the use of ‘that clauses’ in more than just the case of speech and attitude *verbs*. It also arises for complements of nouns and adjectives as well:³⁷

- 37a. The idea that John would mislead us is very unsettling.
 b. The idea [_{CP} **THAT** [_{TP} John would mislead us]] [_{VP} is very unsettling]
 38a. It is a surprise that Jennifer broke up with Brad
 b. It is a surprise [_{CP} **THAT** [_{TP} Jennifer broke up with Brad]]
 39a. John is depressed that his yarn is tangled.
 b. John is depressed [_{CP} **THAT** [_{TP} his yarn is tangled]]
 40a. It is disgusting that some people eat insects.
 b. It is disgusting [_{CP} **THAT** [_{TP} some people eat insects]]

Since the only thing that seems to be relevant to the theory offered here is a certain kind of grammatical configuration involving a complementizer, these cases can be treated in the same general way as sentential complements of verbs are treated. One does not need to have a subject of a psychological state in order to have a content imputed via a ‘that clause’ as evidenced by 38 and 40.

Indeed, the fact that we have a purely structural explanation of how intensionality arises in the cases reviewed also helps explain why intensionality seems to arise from the nature of the ‘that clause’ and not purely from the meaning of the

³⁷ As we move further away from indirect speech and reports, the distributional differences between complementizers and demonstratives become more and more obvious, e.g., ‘I am happy that you came’ vs. ‘*I am happy that. You came.’ There are loose paraphrases for what is needed here. For example: I am happy about this: you came. Other paraphrases are possible for the other cases, although there is probably no single way of rephrasing all the cases with parallel nominal demonstratives.

subordinating predicate. And this helps with other cases. For example, some verbs that would not ordinarily seem to afford the possibility of intensionality can be used with sentential complements and can give rise to psychological readings:

41. John **let on** that he wasn't feeling well.
42. It **struck** me that Bill resembles the guy in the FBI photo.
43. I **gather** that you want to meet my wife.

Indeed, even auxiliary verbs can, in the right circumstances, be used with apparent sentential complements:

44. John **has it** on good authority that the army will attack at midnight.
45. I **get** that you don't like me. I just can't figure out why.

And sentential complements often appear as subjects themselves:

46. That there is blood on the carpet is strong evidence that the murder took place here.

Given the appearance of the complementizer 'that' (which differs quite dramatically from that of the demonstrative 'that') in all of these cases, there is a nice explanation for why intensionality seems to arise. If we were to limit Davidson's idea to just a small circle of verbs, we would miss out on an obvious way of encompassing these types of constructions. But we don't want to have to admit two complementizers 'that' – one for indirect speech and propositional attitude verbs and one for every other context. And not having two complementizers 'that' may have some very nice advantages as well.

Returning to the issue of binding, the lessons of a purely structural explanation of intensionality and of ordering between clauses also afford a nice explanation of a case like the following:

- 47a. That every boy didn't show up clearly suggests that he didn't know when the exam was scheduled.
- b. [_{CP} That [_{TP} every boy didn't show up clearly]] [_{VP} suggests [_{CP} that [_{TP} he didn't know when the exam was scheduled]]]

In this case, the clauses are as follows:

- 48a. **THAT_k** clearly suggests **THAT_j**
 - b. Every boy didn't show up
 - c. He didn't know when the exam was scheduled.

Where **THAT_k** refers to 48b and **THAT_j** refers to 48c. The predecessor of 48c, the phrase containing the pronoun, is 48a, which does not contain a potential binder for the pronoun. 48b does contain a quantifier but since 48b is not the predecessor of 48c that quantifier cannot bind the pronoun. If this is right, then there is a way of understanding 'parataxis' which preserves Davidson's core thought, however different in execution it is from Davidson's own proposal. It also has some nice features, once it is suitably generalized to encompass all uses of 'that clauses.'

When outlining the paratactic theory in the opening sections, I mentioned that all that Davidson really needed, apart from one or another version of ‘parataxis,’ was the idea that the argument of ‘say’ was a semantically unstructured singular term. Davidson’s other commitments made it clear that he was not, like Quine, taking the complement clause itself to be a semantically unstructured term. It is very important to Davidson’s view that complement clauses do have structure. More importantly, as the weakened version of the thesis implies, it need not be the case that the complementizer ‘that’ must be utterly parallel with the demonstrative, a good thing, since it very clearly isn’t. Indeed, the only real parallel between complementizers and demonstratives that is required for the theory is that they are both singular terms whose semantic value is liable to change from one tokening to the next.³⁸

But even this weakened version might have advantages. Consider cases where what appears to be an embedded clause is not a finite indicative and could not itself be used as a free-standing assertion. If I am right, there is no need that the referent of the argument of ‘say’ or any linguistic expression delineated by the paratactic parsing needs to be even *capable* of being a root clause. And it does not ultimately matter whether or not it is ‘**THAT_C**’ that plays the subordinating role. Indeed, if ‘**THAT_C**’ encodes the information that the following clause is to be understood as being a finite clause, then the theory predicts that other complementizers appear when their complement clauses are not finite and indicative. Other sorts of complementizers encode different sorts of information. This helps with certain problematic cases. Michael Hand, for example, complains that the paratactic theory is defeated because of nonfinite clauses. Since the content clause of

49. I want to buy a book by Clemens

is the non-finite

50. [_{TP} to buy a book by Clemens]

and since this linguistic expression itself is not, or not usually, used to make a self-standing assertion, it seemingly follows that a speaker cannot make himself a samesayer with its usage.

Whatever the merits of this objection are with respect to Davidson’s *original* view, it gains no traction on the view given here. In more detail, on the view given above, clausal subordination is not a way of making a clause a *subpart* of the logical form of a sentence but rather a way of placing it into a context within which its free variables, whatever they are, may be satisfied by syntactically preceding expressions. The context sensitive expression is the clause itself, so to speak. Its complementizer – its ‘subordinating coordinator,’ in an older terminology – encodes how the clause is to be interpreted in much the same way that person and tense and gender inflection attached to nominals and verbs tell us how we are to interpret them. How

³⁸ For all that Davidson says, demonstratives may be semantically structured or not: no particular view about demonstrative reference is needed.

the values internal to the clause are to be spelled out or satisfied is determined by the properties of the complementizer phrase.

As is well known, different complementizers correspond to different types of clauses with which they associated.³⁹ The examples above considered just the binding of nominal pronouns by a single universal quantifier from one clause to its immediate successor. But, in light of the fact that most clauses have some inflection for tense and that this position is left open, the way in which the morphology of complementizers is linked to tense takes on a special significance. For the case of the infinitives alluded to above, the interpretation of the content clause is keyed to the semantic value of the tense position of the subordinating clause in the following way: the value of the tense position of the content clause can only be a time which is strictly later than whatever time is assigned to the reporting clause. In the case above, the buying must be interpreted as following the wanting. More specifically, if the predecessor of the complement clause is interpreted as having occurred at time t_1 , then the complementizer of the nonfinite clause which follows encode the information that its temporal coordinate is to be assigned a temporal value which is later than t_1 . And that seems right. The CPs of ‘tensed infinitives’⁴⁰ should encode how the tense of their clausal associate is linked to the subordinator. The tenses of embedded indicatives work in a parallel way.⁴¹

Tense and mood are perhaps the default sorts of information encoded by all complementizers.⁴² If that is right, then even though the thesis that complementizers are demonstratives straightforwardly fails, a different understanding of the role of complementizers that does respect the core idea of the paratactic view is possible, one with considerably more empirical appeal than its (theoretical) ancestor.

Returning to the larger argument of this paper: Scaling up the foregoing miniature into a more lifelike semantics for natural language requires a thorough examination of a broad range of data from intensional contexts generally and from other semantic phenomena whose primary interest lies outside the properties of attitude sentences. Needless to say, I have not tried here to give a complete binding theory. Much of the negative data regarding binding, of when certain sorts of binding *aren't* possible, has not been touched on, although I have suggested one way in which one way to think of the negative data. I have also not discussed other quantifiers or the vexed matter of the interpretation of embedded subject positions of infinitivals. What I have tried to do here is show one way in which one might encode dominance relations via an

³⁹ The cases are numerous; I discuss some of them in my 2004. Note that even though Davidson is wrong that the complementizer ‘that’ associated with finite indicative clauses is the same as the nominal demonstrative expression ‘that,’ the great diversity of complementizers in the languages of the world and what they co-vary with – usually, the tense of the clause and the open oppositions within it – tells in favor of the conception of parataxis being urged here.

⁴⁰ Stowell (1982).

⁴¹ See Higginbotham (2002).

⁴² Perhaps then the *only* specific information that ‘**THAT_C**’ encodes pertains to whether or not the clause that follows is interpreted as finite or non-finite.

ordering between clauses without having a subordinated clause form a subsentence of a larger sentence. The negative conditions on binding can be encoded in a similar manner.

What I have done in this section was to propose two ways of understanding the key parts of Davidson's theory. First, I have refined the notion of 'parataxis' in a way allows for binding to occur across a gap, as it were. This involved thinking about the interpretation of multi-clausal constructions in a way that allows for binding but does not entail that a content clause is contained as a proper part in the same structure as the reporting clause. In essence, I have treated subordination between clauses as a matter of ordering their interpretations in a specific way. The second modification I sketched involved thinking about complementizers as encoding information about how the clause with which it is associated is to be interpreted. The result was a structural theory of intensionality, one based upon the ordering between clauses. This has certain advantages in explaining why certain kinds of binding aren't possible and why intensionality arises in some cases where it might not be expected.

If this is right, then the objection with which I started this section really does deliver a decisive blow against a certain package of assumptions, ones that Davidson seems to have had in mind. But the theory can be developed in a different, if not completely obvious way, one which sidesteps the objection.

5 Conclusion

My more general point here has not been that Davidson's original view was right all along. As I have pointed out, that simply isn't the case. But the reasons for the failure, while more or less well understood, leave considerably more options on the table than has been thought. On the view that I have developed here, the structure of natural language merely allows for intensionality to arise in specific constructions. It is a further question why it *does* arise in certain cases and, perhaps more importantly, why it *doesn't* arise in other cases, e.g., why the natural language correlates of propositional connectives often behave in ways that appear to respect a purely extensional semantics. As mentioned above, the answers to these questions will only partly be within the area of natural language semantics. Many of the phenomena that appear to be problematic for a paratactic view involve severing the syntactic bridge between sentences supporting important semantic relations. However, those relations can be reconstructed in a way that both respects the paratactic component of the paratactic theory and does not bring onboard assumptions that are contrary to the spirit of the proposal.

Removing one of the obstacles to the development of the – or least *a* – paratactic view allows a better view of the depth of these issues. If the foregoing reconstruction of Davidson's view can be sustained, then it should be clear that there are respectable ways of understanding natural language in which aspects of the old problem of referential opacity and intensionality are no longer simply problems of the logical form of this or that construction.

Acknowledgments A preliminary version of this material was presented at Carleton University in the fall of 2003. I would like to thank my audience there. The comments and questions from Michael Glanzberg, Robert May, and Kent Bach were of assistance to me as I prepared this version. Drafts of the paper were written while I was a PREA postdoctoral fellow at the University of Western Ontario, support for which I am grateful. Chris Viger's eyes and ears were also helpful, as were Rob Stainton's, again and again. Discussions with Ernie Lepore, Paul Pietroski, Howard Lasnik, and Samuel C. Wheeler III were also particularly helpful.

References

- Blair, D. (2004). *Parataxis and logical form*. Unpublished Ph.D. dissertation, University of Connecticut.
- Botterell, A., & Stainton, R. J. (2005). Quotation: Compositionality and innocence without demonstration. *CRITICA*, 37(110), 3–33.
- Burge, T. (1986). On Davidson's 'On saying that,'. In E. Lepore, (Ed.), *Truth and interpretation*. Blackwell's.
- Büring, D. (2005). *Binding theory*. Cambridge University Press.
- Cappelen, H., & Lepore, E. (2004) *Insensitive semantics*. Basil Blackwell.
- Chomsky, N. (1955/1975). *The logical structure of linguistic theory*. Plenum: New York.
- Chomsky, N. (2000). Minimalist inquiries: The framework. In R. Martin, et al. (Eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik* (pp. 89–155). MIT Press.
- Chomsky, N. (2001). Derivation by phase. In *Ken Hale, A life in language*. MIT Press.
- Chomsky, N., & Halle, M. (1968). *The sound pattern of english*. MIT Press.
- Clapp, L. (2001). Davidson's program and interpreted logical forms. *Linguistics and Philosophy*, 25, 261–297.
- Davidson, D. (1968). On saying that. reprinted in Davidson, 1984.
- Davidson, D. (1979). Moods and performances. reprinted in Davidson, 1984.
- Davidson, D. (1984). *Inquiries into truth and interpretation*. Oxford University Press.
- Hand, M. (1991). On saying that again. *Linguistics and Philosophy*, 14, 349–365.
- Heck, R. (2004). Truth and disquotation. *Synthese*, 142, 317–352.
- Heim, I. and Kratzer, A. (1998). *Semantics in Generative Grammar* Oxford: Blackwell.
- Higginbotham, J. (1986). Linguistic theory and Davidson's program in semantics. In E. Lepore (Ed.), *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson* Oxford: Blackwell.
- Higginbotham, J. (1991). Belief and logical form. *Mind and language*, 6, 344–369.
- Higginbotham, J. (2002). Why is sequence of tense obligatory? In G. Preyer and G. Peter, (eds.) *Logical Form and Language*, Oxford: Clarendon Press, pp. 207–227.
- Higginbotham, J. (2006). Sententialism: The thesis that complement clauses refer to themselves. In *Philosophical Issues, 16, Philosophy of Language*, 101–119.
- Kayne, R. (2002). Pronouns and their antecedents. In S. Epstein, & D. Seely (Eds.), *Derivation and explanation in the minimalist program* (pp. 133–166). Blackwell's.
- Larson, R., den Dikken, M., & Ludlow, P., (forthcoming). Intensional transitive verbs and abstract clausal complementation.
- Larson, R., & Ludlow, P. (1993). Interpreted logical forms. *Synthese*, 95, 305–356.
- Larson, R., & Segal, G. (1995). *Knowledge of meaning*. MIT Press.
- Lepore, E. (Ed.). (1986). *Perspectives into inquiries into truth and interpretation*. Basil Blackwell's.
- Lepore, E., & Cappelen, H. (1997). On an alleged connection between indirect speech and semantic theory. *Mind and Language*, 12, 278–296.
- Lepore, E., & Loewer, B. (1989). What Davidson should have said. In E. Villanueva (Ed.), *Information, semantics and epistemology*, 14, (pp. 3–26). Oxford: Blackwell.
- Lepore, E., & Loewer, B. (1989b). You can say that again. *Midwest Studies in Philosophy*, 338–356.

- Longobardi, G. (2001). The structure of DPs: Some principles, parameters and problems. In M. Baltin, & C. Collins (Eds.), *Handbook of Syntactic Theory* (pp. 562–603). Blackwell's.
- Ludwig, K., & Lepore, E. (2006). *Donald Davidson's Truth-Theoretic Semantics*, Oxford University Press.
- Ludwig, K., & Ray, G. (1998). A semantics for opaque contexts. *Philosophical Perspectives*, 12, 141–166.
- McCloskey, J. (2002). Resumption, successive cyclicity, and the locality of operations. In S. Epstein, & D. Seely (Eds.), *Derivation and explanation in the minimalist program* (pp. 184–226). Blackwell's.
- Pesetsky, D., & Torrego, E. (2001). T to C movement: Causes and consequences. In M. Kenstowicz (Ed.), *Ken Hale: A life in language* (pp. 355–426). MIT Press.
- Pietroski, P. (1996). Fregean innocence. *Mind and Language*, 11, 338–370.
- Pietroski, P. (2000). On explaining that. *Journal of Philosophy*, 97, 655–662.
- Pietroski, P. (2004). *Events and semantic architecture*. Oxford University Press.
- Pollack, J.-Y. (1989). Verb movement, universal grammar and the structure of IP. *Linguistic Inquiry*, 20(3), 365–424.
- Rècanati, F. (2000). *Oratio obliqua, oratio recta*. MIT Press.
- Rizzi, L. (1997). The fine structure of the left periphery. In L. Haegeman (Ed.), *Elements of grammar* (pp. 281–337). Dordrecht: Kluwer.
- Rumfitt, I. (1993). Content and context: The paratactic theory revisited and revised. *Mind*, 102, 429–454.
- Safir, K. (2004). *The syntax of dependence*. Oxford University Press.
- Schiffer, S. (2003). *The things we mean*. Oxford University Press.
- Schlenker, P. (2005) Non-redundancy: Towards a semantic reinterpretation of the binding theory. *Natural Language Semantics*, 13(1), 1–92.
- Segal, G. (1989). A preference for sense and reference. *Journal of Philosophy*, 89, 73–89.
- Segal, G., & Speas, M. (1986). On saying THAT. *Mind and Language*, 1, 124–132.
- Stowell, T. (1982). The tense of infinitives. *Linguistic Inquiry*, 13, 561–570.
- Uriagereka, J. (1999). Multiple spell out. In N. Hornstein (Ed.), *Working Minimalism*. MIT Press.

Part II
Context and “What Is Said”

On the Epistemic Utility of *What is Said*

Sanford C. Goldberg

Abstract Given a speaker who performed an assertoric speech act, what the speaker can be reported to have said (by way of a speech report of the form ‘*S* said that *p*’) depends on the context of the report. Insofar as the notion of *what is said* captures the content dimension of such reports, the result is that this notion is of little use to epistemology, and in particular cannot be used to capture the content dimension of those speech acts that are apt for the testimonial communication of knowledge. To capture such a dimension, we ought to replace the notion of *what was said*, with that of *what was strictly said*. I argue that there are epistemic intuitions, pertaining to speakers’ epistemic responsibilities in producing testimony, that can be used to constrain what counts as the ‘strictly said’ content.

Keywords Testimony · what is said · context-sensitivity

1 Introduction

This paper is structured around a difficulty that arises when we consider speech reports of the form ‘*S* said that *p*’ in connection with the sort of knowledge that one acquires through accepting another speaker’s testimony. The difficulty can be presented in terms of three propositions, each of which enjoys some independent support, but which appear to form an inconsistent triad. The propositions in question are these:

- (1) If *S* said that *p* in her testimony-constituting speech act, then *p* is something that a hearer *H* can come to know¹ through his acceptance of *S*’s testimony.

S.C. Goldberg

Department of Philosophy, Crowe 3-179, 1880 Campus Drive, Northwestern University Evanston IL 60208-2214

¹ ‘can’ come to know: the actual acquisition would depend on the actual acceptance under conditions involving the hearer’s having understood the testimony and the additional satisfaction of various epistemic conditions. I discuss this below. In any case, this qualification should be understood throughout, whenever I discuss (1).

- (2) Whether a speaker *S* counts as having said that *p* – whether her speech contribution on a given occasion can be correctly reported by a speech report of the form ‘*S* said that *p*’ – depends on the context of the reporting.
- (3) What propositional content(s) a hearer *H* can come to know through his acceptance of speaker *S*’s testimony does not depend on the context of reporting (whether by *H* or anyone else) of *S*’s testimony-constituting speech act.

Given (1) (and the satisfaction of the various epistemic conditions), what a speaker said – that is, the proposition that captures what was said – is a content whose truth a hearer can come to know through her acceptance of the speaker’s testimony. Given (2), what a speaker counts as having said depends on the context of the reporting. It follows (given both (1) and (2)) that what a hearer can know through her acceptance of another’s testimony will depend on the context of the reporting of that testimony. But (3) denies this very sort of dependency. Yet each of (1)–(3) appears to be supported by independent considerations – a *prima facie* problem.

A solution to our *prima facie* problem will have to deny one (or more) of (1)–(3). My proposal will be to deny (1). Doing so enables us to accommodate the battery of intuitions that support (2), while simultaneously preserving the key features of testimonial knowledge transmission – features that support (3). In light of this proposed resolution, I suggest that the significance of the foregoing problem is two-fold. First, it suggests that, insofar as we think about language as a vehicle for the testimonial communication of knowledge, we do best to replace the notion of ‘what is said’ by an alternative notion – something akin to ‘what is strictly said’. But second, and more interestingly, the matter of determining what is strictly said (on a given occasion) would itself appear to be constrained by considerations deriving from the pragmatics and epistemology of would-be cases of knowledge communication. In my concluding section I consider what bearing, if any, the present results have for semantics.

2 *What is Said* and Testimonial Knowledge

Let ‘testimonial knowledge’ designate the sort of knowledge acquired through accepting another’s testimony. Among other things, what is distinctive of this sort of knowledge is its epistemic dependence on another’s speech.² More specifically, it is a kind of knowledge whose status as knowledge depends on the fact that the source speaker, in having spoken as she did, has non-accidentally gotten things right.³ We might say that this feature, pertaining to the reliability (or non-accidental truth) of the source speech act, characterizes one aspect of the *epistemic* dimension

² See also Goldberg (2006) and Chapter 1 of Goldberg (2007a) for further discussion.

³ This isn’t quite right: see Goldberg (2005) and Goldberg (2007a) Chapter 1. However, I will disregard the relevant complications, as they do not affect the points made in this paper.

of the process whereby testimonial knowledge is acquired.⁴ The semantic (or, more broadly, representational⁵) dimension of this process comes in when we consider the notion of ‘getting things right,’ as applied to a case of speech. In order to determine whether a speaker ‘got it right’ in speaking as she did, we have to determine how she ‘got it’ or ‘took things to be’ in the first place. Assuming that how she took things to be is itself expressed in her speech act, determining this is a matter of discerning the representational features of that speech act.

In this connection, it would appear that there is a straightforward link between what a speaker *says* in offering a given piece of testimony, and the particular piece of testimonial knowledge which, in testifying as she did, she enables her audience to acquire. This connection is seen in the following idea: what a hearer *H* can come to know in the way distinctive of testimonial knowledge is determined by what the speaker *S* said in her testimony-constituting speech act. If *H*’s would-be testimonial knowledge is through accepting *S*’s testimony, *H* cannot come to know *p* in the manner distinctive of testimonial knowledge unless *p* is something that *S* said in her testimony-constituting utterance. As we might put it: *S*’s having said that *p* (in her testimony-constituting utterance) is a *necessary* condition on *H*’s acquiring the knowledge that *p* through accepting *S*’s testimony. Clearly, it is not sufficient: *S* might have said that *p*, but in so doing she might have said something false (in which case *H*’s testimony-based belief that *p* is not knowledge, for being false); or else, though true, *S* might have arrived at that truth in a way that was too accidental to count as knowledge (in which case *H*’s testimony-based belief that *p* is not knowledge, for inheriting the unreliability of the source testimony⁶); or else, even if *S*’s say-so is both true and reliable (in that she wouldn’t have said so if it had been false), even so *H* (who accepted that say-so and so came to believe that *p*) might have been overly-credulous in his acceptance (in which case his testimony-based belief that *p* is not knowledge, for being unreliable owing to his own credulity).⁷ In any of these cases, *H*’s belief that *p*, formed through accepting *S*’s testimony, would fail to amount to testimonial *knowledge* that *p*. Even so, if *p* is something *S* said in her testimony-constituting utterance, then – so long as *H* understood the testimony, and the various epistemic conditions on testimonial knowledge are satisfied – *p* is something *H* can come to know through accepting that testimony, in the way distinctive of testimonial knowledge.

Some additional support for this last claim comes when we consider the following. When one’s knowledge that *p* is testimonial, ‘Because so-and-so said so’ is an

⁴ ‘One’ aspect: there are others. For example, the hearer’s acceptance of the testimony must be based on a reliable capacity to distinguish reliable from unreliable testimony. See e.g. Goldberg (2007b; forthcoming) for a discussion.

⁵ I will use ‘representational content’ when I want to designate the truth-conditional content of a speech act. I use this, rather than ‘semantic content’, in order to use a designator that will not beg any questions regarding e.g. what pragmatic considerations contribute to the truth-conditional content of a speech act.

⁶ See Goldberg (2007b) for further discussion.

⁷ See Goldberg and Henderson (2006).

appropriate sort of reply to a ‘How do you know?’ question. To be sure, ‘Because so-and-so said so’ answers the ‘How do you know?’ question only given that the speaker *S*’s say-so, and the hearer *H*’s reception of it, are otherwise in order. Since the piece of knowledge *H* vindicates through his reply ‘Because so-and-so said so’ is his knowledge that *p*, this reply would appear to be shorthand for ‘Because so-and-so said that *p*’. (Otherwise *H*’s reply would not be to the point in responding to a challenge to this particular piece of knowledge.)

The foregoing considerations suggest that whenever one has acquired testimonial knowledge that *p*, it will be the case that one’s source can be correctly reported as having said that *p*. But there is a *prima facie* case to be made for thinking that the converse holds as well. In particular, it would seem that whenever one’s source can be correctly reported as having said that *p*, one who understands the speech act is in position (assuming the satisfaction of the remaining epistemic conditions) to acquire the testimonial knowledge that *p*. To see this, suppose (for the sake of *reductio*) that a speaker’s having said that *p*, together with a hearer’s acceptance-with-understanding of that say-so under conditions in which all remaining epistemic conditions on the acquisition of testimonial knowledge were satisfied, do *not* suffice for the hearer’s acquisition of the testimonial knowledge that *p*. In that case, we would get the following result:

(+) A hearer’s reply ‘Because so-and-so said so’ is not an adequate reply to a query regarding how she knows that *p*.

(+) holds, given our supposition, since any reply of the form ‘Because so-and-so said so’ could always be directly met (given our supposition) with something like the following response: ‘Yes so-and-so said so, but this does not suffice (even given the satisfaction of the various epistemic conditions noted above) to establish that you know *p*.’ But this sort of response would appear to be unacceptable. It appears to call into question the very possibility of learning (= coming to know) things through accepting another’s testimony. We are imagining a case in which *S* said that *p*, to a hearer *H* who understood and accepted that testimony (and so who came to believe that *p*), in a situation in which all remaining epistemic conditions on testimonial knowledge were satisfied. If such a situation is not one in which *H* counts as knowing that *p*, it would seem that *S*’s having said that *p* is not a route to *H*’s coming to know that *p* in the first place. Since nothing turns on the details of the case, the result would be general: testimony itself would not be a route to knowledge. But this result is unacceptable, given what most epistemologists recognize is the pervasive and ineliminable role testimony plays in the fabric of our knowledge corpus.⁸ And thus it would seem that we ought to reject our initial assumption; we ought to hold instead that a speaker’s saying that *p* *does* suffice to put an appropriately-situated hearer in a position to know that *p* through his acceptance of that testimony. (1) holds.

⁸ See Fricker (1987), Coady (1992), Sosa (1994), Audi (1997), and Graham (1999), to name just a few.

Indeed, (1) can seem obvious. But obvious or not, the foregoing argument on behalf of (1) is worth making explicit. This is because speech reports of the form ‘*S* said that *p*’ are typically taken to have a feature that, I will argue below, is inconsistent with (1). If this is correct, then something must be wrong with the foregoing arguments on (1)’s behalf.

3 Speech Reports

Many authors have argued that what a speaker can be correctly reported to have said, in a speech report of the form ‘*S* said that *p*,’ will depend on features of the context of the report. Cappelen and Lepore (1997) give several types of example. The argument for this involves cases in which a speech report of the form ‘*S* said that *p*’ is judged intuitively correct in some reporting contexts, but false in others. (Cappelen and Lepore use these examples to suggest that, contrary to what many influential semanticists think, the notion of *what is said*, as captured in speech reports of this form, is of limited use to serious semantic theory.)

In order to show that the truth conditions of a speech report of the form ‘*S* said that *p*’ are determined in part by the context of utterance, Cappelen and Lepore give four types of example: obvious implicature cases, general thrust cases, irony cases, and partial overlap cases.⁹

Consider first an obvious implicature case. You observe me responding to the question whether Smith is having an affair with,

(S1) He sure has been visiting a certain woman in New York City lately.

You take me, quite reasonably, to have implicated that Smith is having an affair. In most normal reporting contexts, then, it would be correct to report my speech with

(R1) Goldberg said that Smith is having an affair.

But now imagine a different context. Suppose that there is a legal proceeding going on, where a good deal hangs on whether Smith is having an affair, and where anyone who speaks falsely or unreliably regarding this topic will receive a legal sanction. In such a context, it would seem that (R1) could not be used to make a correct report of my speech contribution. (I might reasonably deny the report: ‘That’s *not* what I said. I said that he’s been visiting a certain woman lately. . .’) In sum, it would seem that whether one who uses (R1) counts as having made a true report of the speech act I performed with (S1) depends on the context of the report.

⁹ They also suggest cases in which ‘*p*’ in ‘*S* said that *p*’ is replaced by a sentence that results when we substitute one co-referring term for another: as when Jones’ asserts ‘Granny drinks H₂O’ and, knowing that Smith does not know what H₂O is, I report Jones’ speech with ‘Jones said that Granny drinks water’.

Consider next a general thrust case. Smith presents a long-winded, pedantic lecture on the evils of lying. The lecture begins with

(S2) In what, precisely, does the evil of lying consist? In this paper I argue that, *qua* evil, lying differs from various other wrongs; that the harm perpetrated is . . .

Aiming to characterize the general topic of the lecture to Jones (who was not in attendance), McSorley uses

(R2) Smith said that lying is wrong.

Intuitively, such a report would be judged a correct speech report, as it is a way by which McSorley can briefly characterize the general drift of Smith's lecture to someone (namely, Jones) who was interested not in the details, but only the general thrust, of the lecture itself. But we can imagine other contexts in which (S2) would not be seen as a true report of Smith's lecture. Suppose that Smith is a very finicky philosopher who gets irritated whenever his views are mischaracterized, who went to great lengths to characterize the precise evil-making characteristics of lying, and whose main aim was to distinguish the wrong perpetrated by lying from various other wrongs. In that case, Smith might well protest that (S2) fails to capture what he said. It would then seem that, if it is used in a context in which Smith's own perspective on the matter is salient, (R2) is false. Thus whether a report made by way of (R2) is a correct report of Smith's speech will depend on the context of reporting.

Next, consider a case of ironic speech. In the presence of what everyone sees (through the window) is a stormy day, Jones asserts

(S3) It is a beautiful day today.

Then Smith, aiming to let McSorley (who did not hear Jones' utterance) know the upshot of Jones' speech contribution, reports this with

(R3) Jones said that it's a lousy day today.

In such a context the report made by way of (R3) would intuitively be judged true, as a report of Jones' assertion of (S3). But again variations in context are imaginable in which a report made by way of (R3) would be intuitively judged false as a report of Jones' use of (S3). For example, if the report were made in a context in which it is well-known to everyone except Smith that Jones never uses 'lousy' in connection with the weather – he thinks that weather should never be described in such terms – then a report made by way of (R3) would intuitively be regarded false as a report of Jones' use of (S3). (Here, the falsity of the report offers a natural explanation of Jones' negative reaction to it.) Once again, we have a speech report of the relevant form, (R3), which in one reporting context is a true report of Jones' use of (S3), but in another reporting context is a false report of Jones' use of (S3).

Or (to take a final example) consider the category of case which Cappelen and Lepore call 'partial overlap' cases. Suppose that Jones asserts

(S4) Judy bought fourteen Gucci handbags in the mall today.

And suppose that McSorely, who did not hear Jones' assertion, is interested in Judy's recent purchases. Then with McSorely as the intended audience (and with the date unchanged) Smith might correctly report Jones' speech with

(R4) Jones said that Judy bought handbags today.

But if we are in a context in which it is common knowledge that Peterson (who didn't hear Jones' speech) is interested in the number of times people make reference to the Gucci brand, and if – thinking that Jones might have just done so – he asks Smith what Jones said, clearly a report made by way of (R4) will not do as a correct speech report. Once again, whether a report made by way of (R4) is a correct speech report of Jones' use of (S4) depends on the context of the report.

These and other cases suggest the following hypothesis:

- (2) Whether a speaker *S* counts as having said that *p* on a given occasion – whether her speech contribution on that occasion can be correctly reported by a speech report of the form '*S* said that *p*' – depends on the context of the reporting.

Again, this can seem so obvious as not to be worth arguing for.

4 Potential Problem: Testimonial Knowledge and the Context of Speech Reporting

So far I have been arguing for propositions (1) and (2). Rather than arguing for proposition (3) directly, I want to suggest that, by themselves, (1) and (2) appear to raise a problem; it is in reflecting on the nature of the problem that we will come to see that (and why) (3) is true.

I begin by restating the two results so far obtained, the first pertaining to the connection between what was said and testimonial knowledge, and the second pertaining to the context-sensitivity of determinations of what was said. They are as follows:

- (1) If *S* said that *p* in her testimony-constituting speech act, then *p* is something that a hearer *H* can come to know through his acceptance of *S*'s testimony.
- (2) Whether a speaker *S* counts as having said that *p* – whether her speech contribution on a given occasion can be correctly reported by a speech report of the form '*S* said that *p*' – depends on the context of the reporting.

Given (1), any case in which *S* said that *p* (in her testimony-constituting utterance) is a case in which a hearer *H* can come to know *p* through accepting *S*'s testimony. Given (2), whether *S* said that *p* depends on the context of reporting. Putting the two points together, we get the result that whether *H* can come to know *p* through accepting *S*'s testimony will depend on the context in which that testimony is reported. The difficulty is that what one can come to know through acceptance of

another's testimony – the *content* one can come to know in the way characteristic of testimonial knowledge – would not seem to depend on context in this way.

An initial point to make here is this. The contextual features relevant to an assessment of the truth of a given speech report have to do with the interests (or perspectives more generally) of the folks to whom the speech report is addressed. These interests may diverge widely from the speaker's motive for speaking as she did. It seems somewhat curious to say that what can be known through accepting a speaker *S*'s testimony is determined by those interests, since this suggests that what can be known through accepting *S*'s testimony is determined by interests that may have nothing whatsoever to do with *S*'s desire to spread the particular piece of knowledge she expressed in her testimony.

This initial point can be deepened in two ways. The first is by considering what we might call the pragmatics of testifying. To testify that *p* is (among other things) to present oneself as occupying a privileged epistemic position *vis-a-vis* the truth of the content one presents-as-true in the speech act itself. This is in the very nature of testifying: if one is not so presenting oneself, one is thereby not testifying – whatever else one might be doing in speaking as one does. The same point about the pragmatics of testimony can be made by noting the connection between testifying and invoking one's own epistemic authority: to testify is to invoke one's own epistemic authority (if only implicitly) on the matter at hand. That is, one is not merely presenting-as-true the content in question, but is also investing this presentation-as-true with one's own epistemic authority (as in: 'you have my word for it'). As I have argued elsewhere,¹⁰ it is in virtue of the speaker's (implicit) invocation of her own epistemic authority in the act of testifying, that a hearer who is justified in accepting that testimony is *ipso facto* entitled to regard the source speaker as having a non-negligible degree of warrant for the content in question. This consideration points against the hypothesis that what can be known through testimony depends on the context in which the testimony is reported. After all, what can be known in the manner distinctive of 'testimonial knowledge' turns on the content(s) regarding whose truth the speaker invoked her own epistemic authority, and this invocation does not depend on features of the context(s) in which her testimony is reported.

There is a second way to bring out substantially the same (initial) point. It has become popular to ask after the norm of assertion, that standard by which particular assertions (and assertions alone) are assessed as appropriate or inappropriate *qua* speech contribution. Now, whatever one's views on that matter, all parties can and should agree that to assert involves presenting oneself as satisfying the norm in question (whatever that norm is).¹¹ So, on the plausible twin assumptions, first, that all testimony involves assertion, and second, that the norm of assertion is an epistemic

¹⁰ Goldberg (2006).

¹¹ See for example Brandom (1983) and Williamson (1996). Both Brandom and Williamson hold that knowledge is the norm of assertion, and go on to claim (more or less explicitly) that to assert *p* is to present oneself as knowing *p*. Other views hold that the norm is weaker than that of knowledge: justified belief that one knows, for example. See Goldberg (forthcoming) and Chapter 1 of (2007a) for further discussion.

norm of some sort or other,¹² we would get the result that one who testifies presents herself as having the appropriate degree of warrant for the attested content.¹³ We might put the same point in an alternative way by saying that, in testifying, the speaker *renders herself responsible* for having an appropriate degree of warrant for the attested content. However, the content(s) in question are determined by the speaker's testimony-constituting speech act, not by any features of the context in which that speech act was reported.

In light of these points about the pragmatics of testifying, return to the implication of (1) and (2) presently under discussion. This is the implication that

IMP What can be known through accepting a speaker *S*'s testimony is determined by the interests of those who consume reports of *S*'s testimony-constituting speech act.

I submit that IMP is false, and that its falsity can be brought out in connection with the points just made regarding the pragmatics of testifying. In testifying, a speaker *S* presents herself as in an epistemically privileged position *vis-a-vis* the truth of a particular content – *p*, say. But an audience for the report of *S*'s speech may have had no interest in the question whether *p*, and so may have had no interest in the fact that *S* presented herself as in an epistemically privileged position regarding that content. But if IMP is correct, in the relevant reporting context it is *that audience's* interests which go into determining what can be known through an acceptance of *S*'s testimony. And this is to imply that, at least in the relevant reporting contexts, *S*'s (implicit) invocation of her own epistemic authority (in testifying as she did) does not fix what can be known through an audience's acceptance of her testimony. That this is not a happy result can be seen by recalling that *S* bears responsibility for having a warrant for that to which she attested. If that to which she attested is determined by what she said, and if what she said is determined by what she can be correctly reported to have said via a speech report of the form '*S* said that *p*,' we reach the result that *S* bears responsibility for having a warrant for contents regarding whose truth she did not even implicitly invoke her own epistemic authority. I submit that this result is absurd on its face.

The foregoing appears to support the third proposition in our triad of propositions introduced at the outset of this paper:

- (3) What propositional content(s) a hearer *H* can come to know through his acceptance of speaker *S*'s testimony does not depend on the context of reporting (whether by *H* or anyone else) of *S*'s testimony-constituting speech act.

¹² 'Plausible' but not wholly uncontroversial: one might hold that truth is the norm of assertion (as in the rule for assertion: 'Assert *p*, only if *p*'). However, Williamson (1996) presents compelling reasons for doubting that truth is the norm of assertion.

¹³ What degree of warrant is appropriate will depend on what the norm of assertion is; see Chapter 1 of Goldberg (2007a).

The relevant point here is simply that what a hearer can know through accepting another speaker's speech – what a hearer can know in the way characteristic of testimonial knowledge – depends on features of that speech, *period*. That is, it depends on features of that speech, and *not* on the varied interests that might be had by those who report that speech and by the audiences to whom such reports are addressed. It would thus seem that the varied contexts in which a testimony-constituting speech-act might be reported do not affect what can be known through accepting that testimony. Of course, if this is true, then IMP is false – and since IMP was implied by (1) and (2), we reach the result that (1)–(3) form an inconsistent triad. Something has to go.

5 A Dialectical Look at Possible Solutions

There are three 'solutions' to consider: deny (1), (2), or (3). (The 'or' is not exclusive.) In this section I present a dialectical case for thinking that we do best to deny (1).

Consider first a cost-benefit analysis of the move to deny (1). On the cost side, anyone who denies (1) will have to deny the following claim as well:

- (1*) In any case in which the various epistemic conditions on testimonial knowledge have been satisfied *vis-a-vis* testimony in which the speaker said that *p*, a hearer (who understood the testimony) is in a position to know *p* in the way appropriate to 'testimonial knowledge'.

Above, in Section 2, I gave reasons to think that denying (1*) would be costly. But perhaps the cost of doing so is not as great as this reasoning suggested. After all, insofar as we are inclined to think that speech reports of the form '*S* said that *p*' have their truth conditions fixed by contextual matters regarding the interests of the reporter and the audience of the report, to just that extent we should be inclined to think that what is reported in such reports (= the content reported) is not fixed by the source speaker *S*'s invocation of her own epistemic authority, and so is not – or at least is not typically – a content that a hearer can know in the manner of 'testimonial knowledge'.¹⁴ It would seem, then, that once one accepts the context-sensitivity of speech reports, (1*) is vulnerable in precisely the way that (1) was. So the cost associated with the move to deny (1) would not appear to be as great as it might initially have seemed.

What is more, the virtues of resolving our problem by denying (1) would appear to be great. In particular, such a position holds out the prospect of enabling one to

¹⁴ Of course, the knife cuts both ways. Suppose we are confident that what is reported in '*S* said that *p*' is a content that a hearer can know in the manner of 'testimonial knowledge' through the speech being reported. In that case, we might want to conclude that speech reports of this form cannot have truth conditions fixed by parameters relating to the context of the report. See Goldberg (2002) for an argument to that effect. (I now regard this 'solution' to our problem as less than optimal, for reasons I discuss below.)

endorse the strong intuitions elicited by Cappelen and Lepore (1997) and others – and in so doing enabling us to acknowledge that the truth of a speech report is to some important degree a matter of the interests of the audience to whom the report is offered – without having to regard the interests of the audience of speech reports as determining the contents that a given piece of testimony makes available as candidate contents of testimonial knowledge. Such a position will be all the more attractive if we can formulate and motivate a stricter notion of ‘what is said’ – one that will determine (given a piece of testimony) what content(s) are candidate contents of testimonial knowledge through that testimony. (I will return to this below, in Section 6.)

Consider next a cost-benefit analysis of the move to deny (2). The obvious benefit is that by denying (2) we can preserve both (1) and (1*). But it can be wondered how great a benefit this is. For (1*) is a general principle, and so even if it is granted that (1*) is supported by intuition, we might wonder what insight intuition should be thought to have on such a general matter. After all, it is one thing to rely on intuition to yield verdicts in particular cases, quite another to rely on intuition to yield verdicts on general philosophical doctrine. Methodologically, it would seem that we do better – and in any case it would seem more in keeping with the methodology of semantics – to rely on intuition to guide our verdicts in particular cases, but to argue for general claims (such as (1), (2), and (3)) on the basis of their ability to square with the data, make interesting predictions, and so forth. By the lights of this methodology, the cost of denying (2) would appear to be great. For those who deny (2) will have to deny the intuitive verdicts in a host of cases where a report of the form ‘*S* said that *p*’ appears to be true, yet where ‘*p*’ is replaced by a sentence whose propositional content is other than that of the speech being reported. I conclude that the proposal to deny (2) is not as attractive as the proposal to deny (1).

Finally we come to the possibility of denying (3). As in connection with the move to deny (2), the move to deny (3) will have us repudiating the verdicts of intuition in a great range of cases. I repeat (3) here:

- (3) What propositional content(s) a hearer *H* can come to know through his acceptance of speaker *S*’s testimony does not depend on the context of reporting (whether by *H* or anyone else) of *S*’s testimony-constituting speech act

To deny this is to hold that the propositional contents knowable through testimony *do* depend on the context of reporting. But this flies in the face of the conjunction of two points that were defended above: first, that what can be known through testimony is fixed by the source speaker *S*’s (implicit) invocation of her own epistemic authority; and second, that *S*’s invocation of her own epistemic authority does not depend on features of the context in which *S*’s testimony is reported. As a result, one who denies (3) must reject some of the data: cases which, intuitively, we would not describe as cases of testimonial knowledge (because the content known by the hearer is not the content attested to by the source speaker), the present view would be forced to describe as cases of testimonial knowledge after all. Once again, it seems that this move is less cost-effective than the move to reject (1).

The foregoing brief look at our present dialectic supports the proposal to reject (1). This is tantamount to repudiating the epistemic utility of the notion of *what is said*. Below I will be suggesting that, if our aim is to have a semantic notion up to the task of characterizing the content made available for uptake in cases of the testimonial transmission of knowledge, we would do well to replace the notion of *what is said* with that of *what is strictly said*; and I will suggest how the latter notion might be characterized in testimony-theoretic terms. But before I do, it is worth asking whether there are any modifications of (1) that might preserve the epistemic utility of *what is said*.¹⁵

In asking this question, we are asking whether there is any condition whose obtaining, together with the truth of a report of the form ‘*S* said that *p*’, ensures that *p* is available as something that a hearer can come to know through an acceptance of *S*’s testimony. It might be thought that we have such a condition if we insist that speaker *S* knows that *p*. The proposal would be that *S*’s knowing that *p*, together with *S*’s having said that *p* in her testimony-constituting speech act, suffices (given the satisfaction of the epistemic conditions on testimonial knowledge) to make *p* available as something a hearer can come to know through *S*’s testimony. But this proposal will not do, independent of present considerations: there are reasons to think that testimony not only transmits pre-existing knowledge, but it can also generate knowledge for the hearer that the source speaker lacks (see Lackey, 1999; Graham, 1999; Goldberg, 2005, (2007b), and Chapter 1 of (2007a)).¹⁶ For this reason, the present proposal is too restrictive. Perhaps a second proposal will be tried: perhaps it will be thought that the epistemic utility of *what is said* can be preserved by changing the required condition. Instead of insisting that *S* knows that *p*, perhaps we can insist instead that *S*’s saying that *p* must be part of a speech act that constitutes *testimony that p*. Here the idea is that a saying that *p* which is also a case of testifying that *p* is a case where the ‘said’ content is available as something that can be known in the way appropriate to testimonial knowledge. If it is to succeed, this proposal needs an independent characterization of (the semantic dimension of) the notion of *testimony that p* – a characterization which itself does not rely on a

¹⁵ I thank an anonymous referee for this suggestion, and for the particular proposals I consider in the next paragraph.

¹⁶ I also think that there are uses of ‘say’ that would show that the proposed fix fails in its aim. Consider the following case, involving a class of students of psychoanalysis, watching a video of an interaction between an analyst and a self-deceived patient. The patient responds to a query to clarify his comments about his mother by asserting, ‘She was truly wonderful. She really was. I loved her deeply. Really!’ The teacher then stops the video and says to the students, ‘Now class, what the patient is really saying here is that he was deeply conflicted about his mother.’ I submit that even if the students come to know that the patient was deeply conflicted about his mother, this knowledge is not knowledge acquired through the patient’s testimony. And yet the example can be expanded so that all of the proposed conditions are satisfied. Suppose that the patient himself, deep down, perhaps unconsciously, *knows* that he is deeply conflicted about his mother. Then we can say the following (let *p* be the proposition that he is deeply conflicted about his mother): the patient is correctly reported as having said that *p*; the patient knows that *p*; the audience comes to know that *p* through accepting what she took the patient’s testimony to ‘say’; the audience’s acceptance is based on the patient’s having ‘said so’; and yet this is not a case of testimonial knowledge.

prior notion of saying that *p*. The difficulty is that, to the extent that the semantic dimension of testifying is characterized at all, the characterizations typically appeal to a prior notion of *saying* (see e.g. Fricker, 1987; Sosa, 1994). In advance of an independent characterization of *testifying that p*, then, this proposal is a non-starter. (Below I will say what I think might provide such an independent characterization; but my proposal in effect abandons the epistemic utility of the notion of *what is said*.)

Although the foregoing discussion of attempts to preserve the epistemic utility of the notion of *what is said* is far from exhaustive, the failure of these attempts does suggest that we might do better to reject (1), and with it the epistemic utility of *what is said*, and replace it with some other semantic notion of more direct epistemic utility. Can we do so? As noted, the move to deny (1) is made less costly if we can find a suitable replacement for (1*). Above I argued against the epistemic utility of the notion of *what is said* on the grounds that what a speaker can be correctly described as having said is something that is fixed in part by contextual parameters (reporter and audience interests) that have little or nothing to do with the source speaker's (implicit) invocation of her own epistemic authority. Is there an alternative notion, one that is not so easily detached from the speaker's perspective, and in particular from her (implicit) invoking of her own epistemic authority? Consider the notion of *what is strictly said*, where this is some function from (i) the meanings of the words used by the speaker, (ii) the assignment of references to any context-sensitive expressions, as guided by the meanings of the expressions themselves and the relevant contextual-supplied values (for this occasion of use), and (iii) any further 'widely' pragmatic processing (if any) needed to recover the truth conditions of the speech contribution.¹⁷ It should not need arguing that *what was strictly said* can serve the role of the content that the speech act makes available as a candidate for testimonial knowledge. More specifically, the hypothesis which is to replace (1*) is this:

- (1**) In any case in which the various epistemic conditions on testimonial knowledge have been satisfied *vis-a-vis* testimony in which the speaker *strictly* said that *p*, a hearer is in a position to know *p* in the way appropriate to 'testimonial knowledge'.

Rather than try to argue for (1**), which may seem obvious in any case, I want to suggest how what is strictly said on a given occasion can itself be investigated by appeal to intuitions pertaining to the knowledge-transmitting use of language.

6 Epistemic Intuitions and *What is Strictly Said*

Suppose that we accept (2), and so regard a single speech act as susceptible to multiple correct but distinct speech reports of the form '*S* said that *p*', where correctness is determined in part by audience interests. Then what a speaker counts as

¹⁷ I am borrowing the notion of 'wide' pragmatic processing from Bach (1997).

having said is something that permits of various correct characterizations, depending on who is interested, and why; but what a speaker counts as having attested to does not permit of such varied characterizations, and in any case does not depend on who is interested, or why.¹⁸ In that case we ought to deny that what is said (as identified in the content portion of these speech reports) = what is attested. We have an epistemic motive, it would seem, for wanting a dimension of assertoric speech on which there is one and only one correct characterization of the representational content of the assertion itself. ‘What is attested’ is a label for this content; I suggested above that we identify this content with what is strictly said.

Interestingly, epistemology provides us not only with a motive for postulating a notion of *what is strictly said*, but also with a set of tools for constraining what can count, on a given occasion, as part of the ‘strictly said’ content. Return again to the pragmatics of testifying. I said above that to testify that *p* is (among other things) to *present oneself* as occupying a privileged position *vis-a-vis* the truth of the attested content. From the speaker’s perspective, testifying thus involves the presenting of oneself as conforming to certain norms – the norms governing appropriate testifying. From the hearer’s perspective, accepting testimony (at least in the paradigmatic type of case) involves relying on one’s interlocutor to have succeeded in conforming to those norms. But it is worth noting that these norms are characterized in terms that are *representational through and through*: one ought to present things as being a certain way (in one’s testimony), only when one occupies a suitably privileged epistemic position regarding the claim that things *are* that way. On the present proposal to identify what is attested with what is literally said, we get the following result: what one literally says = how one is presenting things as being, insofar as one is (in one’s speech) presenting oneself as occupying a suitably privileged epistemic position regarding how things are. Speaking more loosely but more intuitively, we can say that the notion of *what is literally said* is bound up with notions of a speaker’s epistemic responsibilities *qua* potential testifier. To wit:

- (*) A speaker *S* literally said that *p* only if: given a hearer *H* who (i) observed and understood the testimony, (ii) was epistemically entitled to accept the testimony, and (iii) as a result formed the belief that *p* on the basis of *H*’s trust in that testimony, *H* would be entitled to hold *S* epistemically responsible for the falsity or unwarrantedness of *H*’s own belief that *p*, in the eventuality that this belief turns out to be false or otherwise unwarranted.

¹⁸ This is not quite right. Consider attestations made by way of assertoric uses of ‘It was large.’ What is attested in this way may well depend on audience interests: in a context in which it is mutually known that the audience is interested in the size of a nearby asteroid, a speaker who attests through the use of this sentence will have attested to a different content, than that attested to by a speaker who uses the same sentence-type but in a context in which it is mutually known that the audience is interested in the size of an insect observed exclusively by the speaker. However, I take it that this sort of dependence on audience interests is different in kind than the sort at issue in speech reports.

(*) is based on the following line of reason. To testify that p is to present oneself as occupying a privileged epistemic position regarding the truth of p ; but to occupy a privileged epistemic position regarding p is to have warrant for p (where this amounts to having an amount of truth-conducive support for p above some threshold of adequacy¹⁹); so one who testifies to p , and so who has presented herself in such a way that she is behaving properly only if she has a warrant for p , can be held responsible (by a hearer satisfying (i)–(iii)) to have a warrant for p .

The foregoing characterization of *what is literally said* is restrictive in application: (*) holds only for those speech acts that constitute, or that are potential candidates for constituting, testimony. A more general characterization would have to connect up with speech acts that do not fall in this purview. What is more, (*) states only a necessary condition. A more general characterization might try to state a sufficient condition as well. Rather than discussing how such a characterization meeting these two *desiderata* might be achieved, however, I want to suggest how the foregoing (admittedly partial) characterization in (*) can be used to derive interesting results in our attempts to determine what a speaker strictly said on a given occasion. (*) implies that *S's strictly saying that p* goes hand-in-hand with *S's* bearing a certain kind of epistemic responsibility pertaining to having a warrant for p , such that if a hearer is not entitled to regard *S* (in virtue of her speech contribution) as having such a warrant for the claim that p , then *S* did not count as having strictly said that p . Now I submit that the matter of *S's* epistemic responsibilities *qua* speaker is a matter regarding which we will (sometimes? always?) have relevant intuitions. If I am correct about this, then these intuitions, which in the first instance have an epistemic content, nevertheless can be exploited in the attempt to determine the precise representational content of the testimony itself.

An illustration may help. Let us return to the case from Section 3 involving obvious implicature. You ask me (in an ordinary context) whether Smith is having an affair, and I reply by asserting,

(S1) He sure has been visiting a certain woman in New York City lately.

We noted above that in most normal contexts it would be correct to take me to have implicated that Smith is having an affair, and that as a result it would be correct (in most normal contexts) to report my speech with

(R1) Goldberg said that Smith is having an affair.

Now suppose that, on the basis of understanding and accepting my testimony (through a process that is suitably discriminating), you come to believe something whose content is *that Smith is having an affair*. Now suppose that what you believe turns out to be false: Smith's visits with the woman in question were perfectly innocent after all. In this case, would I be responsible for failing to have had the sort of

¹⁹ As above, the precise threshold will depend on one's views about the norm of testimony. Here I remain neutral on that. See Goldberg (forthcoming).

warrant that would render your belief knowledge (had your belief been true)?²⁰ To be clear about this, let us allow (what is no doubt true in any case) that, in speaking as I did, I gave you a reason to believe as you did. This does not settle our question. For our question is not whether, in speaking as I did, I gave you a reason to believe as you did. Our question is rather whether, in speaking as I did, I am appropriately blamed for having wrongly presented myself as in an epistemically privileged position *vis-a-vis* the truth of what you came to believe on the basis of accepting my testimony. On this score the intuitive verdict is that I am not appropriately blamed in this way. For to be appropriately blamed in this way is to have presented myself in this way, and I did not so present myself.

It is worth noting that this verdict can be defended by more than its intuitiveness. To begin, consider that, in having spoken as I did, I gave you the very evidence I have for thinking that Smith is having an affair. The result is that, when it comes to whether Smith is having an affair, you no longer need my word for it – you have my evidence.²¹ So it would seem that, in speaking as I did, I did *not* present myself as in an epistemically privileged position *vis-a-vis* the truth of what you came to believe. This verdict, together with (*), entails that I did not strictly say that Smith is having an affair – a happy result.

Now I do not pretend that this particular result is very exciting: no one would have regarded the implicated content as part of what I strictly said. But in terms of my overall dialectic, the obviousness of this result a good thing. I am proposing that there is a stock of intuitions – intuitions that in the first instance pertain to epistemic matters – which nevertheless can be brought to bear on claims in the philosophy of language (by way of (*)). It can now be said in defense of this proposal that it does not lead us astray with respect to at least one simple case. Although this is far from a soundness proof (!), it does give us some minimal confidence in the soundness of the proposal to use these epistemic intuitions to evaluate claims in the philosophy of language. And some minimal confidence is what we need if we are going to take the proposal seriously, and begin to develop it – and continue to test it – with the attention I think it deserves.

I leave for later the task of presenting a more substantial argument on behalf of the soundness of (*), as a tool for approaching matters pertaining to the representational (truth-conditional) content of utterances. In what remains, I only want to make clear why I think that this is a matter worth pursuing. I do not contend that the relevant epistemic intuitions will always, or even ever, enable us to *determine* the representational (truth-conditional) content of a speech act. (It may well be that in a good many controversial cases, our epistemic intuitions are no clearer than the

²⁰ In asking this question, I am assuming that knowledge is the norm of testimony. For those who think that this is too strong, it can be weakened: where *e* is the relevant (epistemic) norm, the question becomes, am I responsible for failing to have had the sort of warrant that obtains in cases involving beliefs possessing *e*?

²¹ Of course, you need to take my word regarding the evidence, in the sense that you are relying on me to have been in a privileged position regarding the truth of the claim that expresses my evidence; but that is another matter.

sort of semantic intuitions traditionally used in the philosophy of language – the sort elicited when we ask whether a particular term applies or not across a range of cases.) My claim is weaker: these epistemic intuitions can be used to *constrain* hypotheses regarding the representational content of a speech act: if p is to count as what was strictly said by S on O , then this ascription must gibe with our intuitions about the epistemic responsibilities S imposed on herself merely in virtue of having made the speech contribution she did. I submit that in this way we might bring epistemic considerations to bear on several vexed issues in the philosophy of language – for example, whether the meaning of a name is equivalent to the meaning of some definite description, or whether there are uses of definite descriptions on which their meaning is equivalent to that of some referring term. Such hypotheses have been addressed in a variety of ways; it would be interesting to see what results (if any) we would get on such matters if we approached them from the perspective of a speaker's epistemic responsibilities in testimony cases.

7 Conclusion

The aim of this paper has been to resolve a problem that derives from the linkages between the semantics and the epistemology of testimony. After arguing that three propositions, each enjoying some independent attractiveness, form an inconsistent triad, I argued that we resolve the problem by jettisoning (1), a thesis about the role of the notion of *what is said* in identifying what a hearer can come to know through accepting another's say-so. To wit: what a speaker counts as having said on a given occasion depends on the interests of the audience to whom speech reports are offered, and this renders the notion of what is said inappropriate for capturing the content testimony makes available for being believed in the way appropriate to testimonial knowledge. In bringing this out we saw that epistemologists interested in testimony ought to be interested instead in another notion – that of what the speaker *strictly* said. What is more, I suggested that epistemology here may give us some tools with which to discern what is strictly said on a given occasion. Although this suggestion was tentative, I pointed out reasons for thinking that it is worth pursuing in greater detail than I have done here.²²

References

- Audi, R. (1997). The place of testimony in the fabric of knowledge and justification. *American Philosophical Quarterly*, 34(4), 405–422.
- Bach, K. (1997). The semantics-pragmatics distinction: what it is and why it matters. *Linguistische Berichte*, 8, 33–50.
- Brandom, R. (1983). Asserting. *Noûs*, 17, 637–650.

²² I would like to thank Rob Stainton, Claire Horisk, and two anonymous referees for comments on earlier versions of this paper.

- Cappelen, H., & Lepore, E. (1997). On an alleged connection between indirect speech and the theory of meaning. *Mind and Language*, 12, 278–296.
- Coady, C. (1992). *Testimony: A philosophical study*. Oxford: Oxford University Press.
- Faulkner, P. (2000). The social character of testimonial knowledge. *Journal of Philosophy*, 97(11), 581–601.
- Fricker, E. (1987). The epistemology of testimony. *Proceedings of the Aristotelian Society*, Supplemental Vol. 61, 57–83.
- Goldberg, S. (2002). Reported speech and the epistemology of testimony. *Protosociology*, 17, 57–75.
- Goldberg, S. (2005). Testimonial knowledge from unsafe testimony. *Analysis*, 65(4), 302–311.
- Goldberg, S. (2006). Reductionism and the distinctiveness of testimonial knowledge. In J. Lackey, & E. Sosa (Eds.), *The epistemology of testimony* (pp. 127–144). Oxford: Oxford University Press.
- Goldberg, S. (2007a). *Anti-Individualism: Mind and language, knowledge and justification*. Cambridge: Cambridge University Press.
- Goldberg, S. (2007b). How lucky can you get? *Synthese*, 158, 315–327.
- Goldberg, S. (forthcoming). Williamson on assertion and the conditions on testimonial knowledge. In D. Pritchard, & P. Greenough (Eds.), *Williamson on knowledge*. . Oxford: Oxford University Press (in preparation).
- Goldberg, S., & Henderson, D. (2006). Monitoring and anti-reductionism in the epistemology of testimony. *Philosophy and Phenomenological Research*, 72(3), 600–617.
- Graham, P. (1999). Transferring Knowledge. *Noûs*, 34, 131–152.
- Lackey, J. (1999). Testimonial knowledge and transmission. *The Philosophical Quarterly*, 49(197), 471–490.
- Sosa, E. (1994). Testimony and coherence. In B. K. Matilal, & A. Chakrabarti (Eds.), *Knowing from words* (pp. 59–67). Amsterdam: Kluwer Academic Publishers.
- Williamson, T. (1996). Knowing and asserting. *The Philosophical Review*, 105(4), 489–523.

In Defense of Context Shifting Arguments

Lenny Clapp

Abstract In a series of recent papers and a book Cappelen and Lepore advance the negative view that the “Context Shifting Arguments” (CSAs) utilized by proponents of radical pragmatics are unsound. And they advance the positive view that (at least some of) the problematic phenomena invoked in CSAs can be accounted for – or at least set aside – by distinguishing between the truth conditions of a *sentence*, and “what is said” by an *utterance*. My primary goal in this paper is to defend radical pragmatics from Cappelen and Lepore’s negative view. I argue that Cappelen and Lepore’s positive view – with the exception of one glitch – is an instance of the sort of view that the reformers endorse. Moreover, I argue that – though they may not be aware of it – what really compels Cappelen and Lepore to endorse their positive view is CSAs, the very arguments they claim to be unsound. Thus there is a considerable amount of tension in their views. My goal here is to resolve this tension by defending CSAs from Cappelen and Lepore’s criticism, and to illustrate why one should endorse these arguments.

Keywords Semantic minimalism · radical pragmatics · context shifting · contextualism · truth conditional semantics · speech acts · indexicals · context sensitivity · pragmatics · semantic content · utterance content

1 Introduction

A vivacious debate is currently taking place in semantics, pragmatics, and analytic philosophy of language. The foundational model of meaning and communication – a model endorsed and developed by such luminaries as Frege, Russell, Montague, Grice and Davidson – is being questioned, and rejected, by a growing number of philosophers and linguists. These reformers – supporters of “radical pragmatics” – claim that many phenomena are problematic for the foundational model. But other

L. Clapp

Investigador, Instituto Investigaciones Filosóficas, Circuito Mario de la Cueva S/N, Zona Cultural, Ciudad Universitaria, La UNAM, CP 04510, Col. Coyoacán, Del. Coyoacán, El D.F., Mexico
e-mail: lclapp@minerva.filosoficas.unam.mx

philosophers and linguists – supporters of “truth conditional semantics” – defend the adequacy of the foundational model. The reformers include such theorists as Sperber and Wilson (1986), Travis (2000), Carston (2002), Searle (1980), Recanati (2004), Bezuidenhout (2002), Stainton (2005), Pietroski (2005) and myself. The defenders include such theorists as Soames (2002), Stanley and Szabó (2000), Bach (1994) (though he is a special case), and, the focus of the present paper, Lepore (2004). In a series of recent papers (2003a, 2003b, 2004, 2005b) and in a recent book (2005a) Cappelen and Lepore have criticized the general sort of argument advanced by the reformers, and they have advanced a strategy for accounting for at least some of the problematic phenomena. More specifically, they advance the negative view that the “Context Shifting Arguments” (CSAs) utilized by the reformers are unsound. And they advance the positive view that (at least some of) the problematic phenomena invoked in CSAs can be accounted for – or at least set aside – by distinguishing between the truth conditions of a *sentence*, and “what is said” by an *utterance*.¹ In Cappelen and Lepore’s (2005a) terminology, the positive view falls under the headings “Semantic Minimalism” (SM) and “Speech Act Pluralism” (SPAP).

My primary goal here is to defend the reformers from Cappelen and Lepore’s negative view. After a quick review of what is at stake in the debate between the defenders and reformers of traditional truth conditional semantics, I will argue that Cappelen and Lepore’s positive view – with the exception of one glitch – is an instance of the sort of view that the reformers endorse. Moreover, I will argue that – though they may not be aware of it – what really compels Cappelen and Lepore to endorse their positive view is CSAs, the very arguments they claim to be unsound. Thus there is a considerable amount of tension in the views endorsed by Cappelen and Lepore: On the one hand they utilize CSAs to motivate their positive view, yet on the other hand they argue that *all* CSAs are unsound. My goal here is to resolve the tension in Cappelen and Lepore’s position by defending CSAs from their criticism, and illustrating why Cappelen and Lepore can and should endorse these arguments.

2 The Debate Between Radical Pragmatics and Truth Conditional Semantics

The issue that divides radical pragmatics and truth conditional semantics concerns semantic compositionality: how what is said by an uttered sentence is determined by the relevant syntactic structure of the sentence, the meanings of the words in the sentence, and the context in which the sentence is uttered and interpreted.² Here’s a simple model which an unreflective person might suppose applies to natural languages. (As is the custom, I will consider only declarative sentences, and I will ignore tense.)

The Simple “Assign and Combine” Model (SACM): The truth conditions expressed by, i.e. what is said by, an utterance of S is a *function of only* (i) S’s logical form and (ii) the meanings of the words in S.³

According to this very simple model what competent interpreters do when they come to understand what is expressed by an utterance is, first, determine the logical form of the utterance, and what words occur in it. Then, relying on their lexical knowledge, they assign semantic values to these words. Finally, they apply their knowledge of the semantic theory for their language to the logical form and the meanings of the words, and thereby arrive at truth conditions of, or equivalently what is said by, the utterance. A semantic theory according to this model consists of a finite number of rules stating how lexical meanings are combined, in accordance with the structure of the logical form, to determine truth conditions. A semantic theory thus expresses a sort of function *from* the logical form of a sentence and the meanings of words in the sentence *to* the truth conditions of the sentence. The tasks for a semantic theorist are thus to figure out what the logical forms are (a task performed, or at least constrained, by theories of syntax), what the semantic values of words are, and what the rules for combining the semantic values of words – as directed by logical forms – to arrive at truth conditions. Thus, the following pseudo-formalism can represent the Simple “Assign and Combine” Model:

SACM

(LF of S + Values of words in S) \Rightarrow TCs of utterance of S.

SACM is obviously too simple to apply to natural languages. Note that nowhere does the model mention *context*. It assumes that the relevant meaning, the *semantic value*, of every lexical item is *invariant*; i.e. it assumes that every instance, or token, of a word type has the same semantic value. Thus, when it is applied to natural languages, the simple model has the following prediction, and/or entailment:

Entailment #1: For all sentences S, every utterance of S has the same truth conditions.

But, as we all know, this prediction does not accord with our intuitive judgments. For there are words whose semantic values seem to vary across contexts, and thus there are sentences, e.g. ‘I am hungry’, such that different utterances of them seem to express different truth conditions.

This is hardly news, but note what I have just done: I have just sketched what Cappelen and Lepore (2005a) call a “Context Shifting Argument” (CSA) against SACM. What exactly did I do? I found a sentence (type), viz. ‘I am hungry’, such that, according to the model all the factors that are supposed to be relevant for determining the truth conditions of utterances of this sentence remain constant across contexts, but nonetheless the truth conditions seem to vary across contexts. The sentence thus serves as a *counterexample* to the claim that the truth conditions of utterances of the sentence are a *function of* just those features of the sentence.

Note that to provide a counterexample in this way it is not necessary to provide just one sentence type which seems to have different truth conditions in different contexts; it would suffice to provide two *relevantly similar* sentences that have distinct truth conditions in different contexts. Sentences S and S’ are *relevantly similar* just case they do not differ in any way that matters to the model – just in case the model applied to the sentences will assign the same semantic values and

combine them in the same way. So, for example, ‘Twain smokes’ and ‘Clemens smokes’ are *relevantly similar*. (At least they are if we ignore tense and assume that ‘Twain’ and ‘Clemens’ are assigned the same semantic value, and that these sentences have the same logical form.) So, alleged counterexamples to the SACM need not be CSAs; being a CSA is sufficient, but not necessary, for being a counterexample. The following defines what constitutes a counterexample to the SACM:

Sentences *S* and *S'* (which may be identical) constitute a *counterexample* to SACM iff (a) *S* and *S'* are *relevantly similar*, and (b) there are contexts *C* and *C'* such that the truth conditions of *S* in *C* are not the same as the truth conditions of *S'* in *C'*. (CSAs are alleged counterexamples in which *S* is identical to *S'*.)

I think, and hope, that thus far everything I have said is familiar, if somewhat vague. (I have intentionally glossed over distinctions one might want to draw between the content expressed by sentences relative to contexts, and the content of speech acts performed by uttering sentences, as well as between various conceptions of context. Such distinctions are typically drawn *in response* to apparent counterexamples; my concern thus far has been to illustrate what motivates one to draw such distinctions.) I am taking such pains to review this familiar ground because by doing so I hope to make the following less familiar claim obvious, or at least very plausible: The *only* (reasonable) way to refute SACM is by formulating a counterexample, as defined above. For according to SACM, the truth conditions of an utterance are a *function* of various relevant properties of that utterance. How else could this model be refuted, other than by providing cases where the relevant properties remain constant, yet the truth conditions diverge?

How might a defender of SACM respond to the counterexample, indeed, the CSA, provided by “I am hungry”? There are two strategies of reply: “Grice’s Strategy” and “Kaplan’s Strategy.” The defender of SACM *could* respond by denying that clause (b) in the above definition of a counterexample is satisfied. Inspired by Grice, he could claim that *really*, i.e. *semantically*, every occurrence of “I am hungry” expresses the very same truth conditions and our judgments to the contrary are due to our inability to discern the genuine semantic facts, which are not “visible to intuition,” and distinguish them from mere pragmatic facts, which are the subject of our intuitions concerning meaning and truth. The defender might make much ado over distinctions between, say, “what is said by speakers in making utterances” and “the truth conditions of sentences.” In response to *this* counterexample, however, Grice’s Strategy has not proved to be very popular. Its unpopularity is probably due to the fact that the other response seems much more plausible.

The more plausible response on behalf of SACM is to reject the simplistic assumption that *all* words have an invariant semantic value and instead allow that *some* words have different semantic values in different contexts. The paradigmatic examples of such special words are indexicals, demonstratives, tensed verbs, and “contextuals” such as ‘enemy’. Indeed, Cappelen and Lepore maintain that the set of such context sensitive words, which they refer to as the “basic set,” contains nothing beyond such paradigmatic examples (2005a, 2). I’m not going to review Kaplan’s ideas here, as I assume that they are familiar, but the important thing to notice is that what Kaplan’s Strategy succeeds in doing is defanging this alleged

counterexample by making only a minor amendment to the assign and combine model. (Note that to appeal to context sensitive words in response to a CSA is in effect to deny that clause (a) in the definition of a counterexample is satisfied.⁴) That is, if one utilizes Kaplan's Strategy to defang alleged counterexamples involving "basic" context sensitive words, then one endorses what I will call the "Amended Assign and Combine Model" (AACM):⁵

$$\begin{array}{c} \text{AACM} \\ (\text{LF of } S + (\text{SVs of "basic" CS words in } S + \text{SVs of non-CS words in } S)) \Rightarrow \text{TCs of } S \text{ in } C.{}^6 \\ \uparrow \\ \text{Context } C \end{array}$$

The only difference between SACM and AACM is that AACM concedes that *some* special words have semantic values only relative to contexts of utterance. But other than this invoking of context to help determine the semantic value of such special words, the model is unchanged: Once the semantic values of the special context sensitive words are fixed, the assign and combine model applies as before.

Because AACM is only slightly different from SACM, it makes a very similar prediction, or has a very similar entailment:

Entailment #2: For all sentences *S* that *do not contain* "basic" context sensitive words, every utterance of *S* has the same truth conditions.

As a consequence of entailment #2, AACM is also susceptible to refutation via counterexamples (some of which might be CSAs⁷) though to allow for context sensitive words the relation of *relevant similarity* must now be relativized to contexts:

Sentences *S* and *S'* (which may be identical) constitute a *counterexample* to AACM iff there are contexts *C* and *C'* such that (a) *S* in *C* and *S'* in *C'* are *relevantly similar*, and (b) the truth conditions of *S* in *C* are not the same as the truth conditions of *S'* in *C'*.⁸

The debate between the proponents of radical pragmatics and the defenders of truth conditional semantics concerns the adequacy of AACM for natural language. The radical pragmatists maintain that there are counterexamples against even AACM, whereas the defenders of truth conditional semantics maintain that all such alleged counterexamples can be defanged by some combination of Grice's and/or Kaplan's strategies.

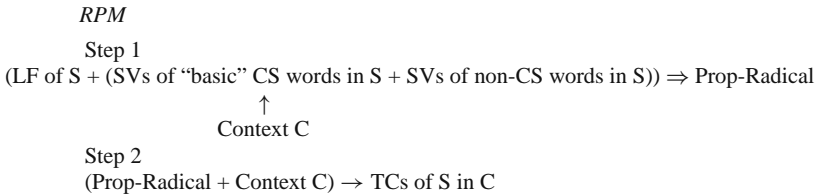
For the sake of illustration, let's look at a familiar sort of alleged counterexample to AACM: the opacity of attitude ascriptions. Let *S* be 'Mary thinks Twain smokes' and *S'* be 'Mary thinks Clemens smokes'. If a number of very plausible assumptions are made, two utterances of these sentences are *relevantly similar* in their respective contexts, yet express different truth conditions. So, we seem to have a counterexample to AACM. How do the defenders of AACM respond?

Some, Salmon (1986) for example, respond by utilizing Grice's Strategy. They claim that clause (b) in the definition of a counterexample is not satisfied, and our initial judgment to the contrary was a result of our confusing semantically encoded truth conditions (which accord with AACM) with information that is merely pragmatically imparted. Such additional information does not fall within the explanatory

domain of AACM, because, despite our naïve intuitions, it does not *really* concern truth conditions, and thus is *not really semantic*.

Others, Crimmins (1992) for example, respond by utilizing Kaplan’s Strategy. They claim that clause (a) in the definition of a counterexample is not satisfied because the logical forms for attitude ascriptions contain so-called “hidden indexicals” which are of course assigned different semantic values in different contexts.

What sort of model do the proponents of radical pragmatics want to put in place of AACM? Though proponents of radical pragmatics endorse different ideas, there is at least a core of agreement. The following I suggest represents this common core of the Radical Pragmatic Model (RPM):



The idea is roughly this. Interpreting an utterance is a two step process. The first step is identical to the process described by AACM, except that according to RPM the output of this process is not propositional; it is not truth conditions, or “what is said.” Rather it is, in Carston’s (2002) terminology “sub-propositional.” Here I will borrow Bach’s (1994) suggestive terminology, and call the output of the first step a *propositional radical*. The second step consists of an *additional* at least partly heuristic interaction between the propositional radical produced in the first step, and context. Different radical pragmaticians think of this second step in different ways: Carston (2002) speaks of the propositional radical as being in need of “completion” and/or “enrichment” – heuristic processes which yield something fully propositional. I prefer to conceive of the second step within the framework of *dynamic semantics*, so that that instead of thinking of context being utilized to “fill in” or “develop” the propositional radical, one conceives of the radical as being like a set of instructions which are *applied to* the context of utterance, to yield a new context. Though these differences among the views of radical pragmaticians are significant, they will not be my concern here. Here all that is relevant is that all radical pragmaticians agree that context plays a much larger role in determining truth conditions of utterances than is claimed in the AACM. All radical pragmaticians agree that the role of context is not *limited* to fixing (or helping to fix) the semantic values of “basic” context sensitive words.

3 Cappelen and Lepore’s Positive Views and the Case of “Incomplete Definite Descriptions”

The core claim of Cappelen and Lepore’s positive views is that the problematic phenomena invoked in CSAs can be explained away by distinguishing between the truth conditions of a *sentence*, and “what is said” by an *utterance*. In this section I am

going to consider a particular problematic phenomenon that is addressed in detail by Lepore (2004). I will argue that what really motivates Lepore's view concerning incomplete definite descriptions is counterexamples – indeed, CSAs – against AACM. Lepore's view concerning the problem of incomplete definite descriptions, however, is merely an instance of the *general* positive view advocated by Cappelen and Lepore (2005a) and (2005b); this more general view is a combination of what Cappelen and Lepore call Semantic Minimalism (SM) and Speech Act Pluralism (SPAP). The upshot is then that what supports Cappelen and Lepore's positive view are CSAs. But this creates a significant tension because Cappelen and Lepore's negative view is that all CSAs are unsound.⁹

Why exactly do “incomplete definite descriptions” pose a problem, and for whom, or what, are they a problem? Lepore introduces the issue as if all that were at stake were Russell's quantificational theory of definite descriptions: “Critics and champions alike have fussed and fretted . . . about whether Russell's treatment is compatible with certain alleged acceptable uses of incomplete definite descriptions” (2004, 41). But he later, correctly in my view, suggests that what motivates the fussing and fretting is not merely an allegiance to Russell's quantificational view, but rather something more foundational:

Giving up Russell's achievements has proven enormously difficult. The definite article 'the' behaves grammatically and in at least some cases uncontroversially semantically like standard quantifier expressions, so much so that it's hard to see how a *semantic theory* of complex noun- phrases could *proceed systematically* were we to deny definite descriptions quantificational status. (2004, 43, my emphasis).

In this passage, the fate of a possible “systematic semantic theory” as a whole is tied to the fate of Russell's quantificational theory of definite descriptions. I think it is clear from the context of the above citation that what Lepore has in mind when he writes of a “semantic theory” that “proceed[s] systematically” just is a theory that fits AACM. Thus, *what* incomplete definite descriptions pose a problem for is AACM. But now *why* are they problematic for AACM? Why does so much ride on the ability of theorists to defend Russell's theory? Why is anything more than Russell's theory of the definite article itself at stake? Lepore is not very explicit about this, but nonetheless these questions are easily answered.¹⁰ For recall the very plausible claim I previously took pains to highlight: The only (reasonable) way to refute AACM is by invoking counterexamples. Thus, the reason incomplete definite descriptions pose a problem for AACM is that they can be utilized to formulate counterexamples, indeed CSAs, against AACM.

Consider the sentence

(1) The table is broken.

This sentence permits of problematic “attributively incomplete” uses. That is, speakers can use this sentence to say something true, even though we all know that there are lots of tables out there. Such uses are, or at least seem to be, incompatible with Russell's quantificational theory of definite descriptions. But why would such “attributively incomplete” uses threaten the possibility of a “systematic semantic theory”? The reason is that sentence (1) gives rise to CSAs against AACM. That is,

it is *also* at least *possible* that somebody use (1) in a way that is *compatible* with Russell's theory, i.e. there are also "attributively complete" uses of (1). So, though typically (1) is used in an attributively incomplete way, one can also imagine it being used in an attributively complete way. This is problematic because the truth conditions corresponding to the two uses differ. In other words, many sentences containing definite descriptions give rise to CSAs, and thus counterexamples, to ACCM.¹¹

For the sake of illustration, let's use (1) to formulate a CSA against AACM. First, imagine a world much like the actual world, in which you and I know that there are many tables. We enter my living room, and you politely compliment my furnishings. Not being one to stand on pretense, I utter (1), informing you of the fact that the most prominent table in my living room is broken. Contrary to what is predicted by Russell's theory you would *not* judge my utterance to be untrue (much less false) on the grounds that there are many tables out there aside from the most prominent one in my living room. Second, imagine a world very different from the actual world; furniture is very rare here. Indeed, I believe, and I take you to believe, that for each type of furniture piece, there is only one piece of that type; e.g. one chair, one table, etc. We are discussing the status of the furniture when I utter (1). As it turns out, there are really two tables, and you know this. Here, as predicted by Russell's theory, you would judge my utterance to be untrue (though perhaps not false, but this is not relevant). In response to my utterance of (1), you might say something such as, "Um, hold on, that's not right. There are actually two tables." Thus we have a CSA against AACM: my stories about imaginary contexts illustrate that different utterances of sentence (1) intuitively have different truth conditions; some utterances of (1) have the truth conditions Russell's theory predicts them to have, while other utterances do not. Since (1) contains no "basic" context sensitive expressions, my stories involving utterances of (1) constitute a counterexample, a CSA, against AACM. *That's the problem.*

How does Lepore respond to the problem? Given passages such as the one cited above, in which Lepore seems to endorse Russell's quantifier analysis of definite descriptions and more importantly the sort of "systematic semantics" Russell endorsed, one might expect him to launch a traditional defense of AACM. That is, one might expect him to formulate some version of either Kaplan's Strategy, or Grice's Strategy, in an attempt to defang the alleged counterexamples (CSAs) involving incomplete definite descriptions. But this is not what Lepore does. Indeed, the first two thirds of the paper on the problem of incomplete definite descriptions (2004) is concerned with *rejecting* "solutions" to the problem which utilize these strategies. (Lepore calls Kaplanesque strategies "semantic strategies" and Gricean strategies "pragmatic strategies.") I am not going to review the arguments presented against Kaplan's and Grice's strategy, though I think Lepore is right to reject such attempts to rescue AACM. What is important for my purposes is that thus far Lepore is marching in lock-step with the radical pragmaticians: he considers a phenomenon that is problematic for AACM because it gives rise to CSAs and then he rejects as inadequate all the utilizations of Kaplan's and Grice's strategies that attempt to

rescue AACM from the problem. This is hardly what one expects from a theorist who maintains that all CSAs are unsound.

How then does Lepore propose to defend AACM from CSAs that arise from the phenomenon of “incomplete definite descriptions”? Lepore’s positive proposal is founded upon the important observation that our judgments concerning “what is said” by an utterance – i.e. what we judge the *truth conditions* of an *utterance* to be – are influenced by *both* the context of utterance, and by *the context of judgment*. He formulates this point in terms of *reporting* what is said by an utterance, instead of in terms of judging what is said by an utterance, but he clearly takes these to be equivalent:

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 [or make judgments about] what’s said. (2004, 65)

Let’s call the view of “what is said” being advanced in the passage the *relativity of what is said*. This is a rather radical doctrine, reminiscent, I think, of the later Wittgenstein and his followers. For according to this doctrine, we have no context-independent, God’s-eye access to “what is said” by an utterance, real or imaginary. And since what is said by an utterance just is what, in a particular context of judgment, we judge (or “intuit”), to be said by the utterance, it follows that there is *no objective, i.e. context-of-judgment-invariant, fact concerning what is said by an utterance*. This is pretty radical, and, ironically, it is heartily endorsed by Travis, perhaps the most radical of the radical pragmaticians. Compare the above passage from Lepore with the following passage from Travis:

On the present conception of a thought [i.e. what is said], Max counts as expressing the *thought* Sid did just where he counts as saying things to be the way Sid did. On the new occasion, he would so count, and would fail to, in ways he would not have done on the occasion on which Sid spoke. We need not take this to mean that on different occasions Sid will count as having expressed different thoughts (though there may sometime be cause to speak that way). We need only say that, on different occasions, different things would count as expressions of the thought that Sid expressed in saying what he did. No purpose would be served here by insisting on one occasion-insensitive criterion for ‘same-thought’ [i.e. for the identity of “what is said”]. (Travis, 2000, 160)

How does this doctrine of the relativity of what is said help Lepore explain away the phenomenon of incomplete definite descriptions? Recall what the problem is and for whom it is a problem. The problem is that we judge different utterances of, e.g., (1) to have distinct truth conditions. (Indeed, if the doctrine of the relativity of what is said is correct, we might in different contexts of judgment even judge the very same utterance to express different truth conditions.) Thus, such utterances of (1) seem to constitute a counterexample to AACM. Now, Lepore has claimed that there is nothing to what is said by an utterance beyond what we, on a particular occasion, judge to be said, and moreover that such judgments depend upon “a number of . . . ‘non-semantic’ features of the context of utterance” (2004, 61). So, in other words, Lepore’s response to the problem is simply to reject the idea that

what is said by an utterance is determined solely by “semantic” factors. It certainly seems as if Lepore is endorsing CSAs involving incomplete definite descriptions and advocating rejection of AACM in favor of some version of RPM.

But things are not that simple. Lepore clearly takes himself to be defending Russell’s theory, and “systematic semantics” along with it: In summarizing his position Lepore writes, “Incomplete definite descriptions are non-denoting, just as Russell taught us,” (2004, 65) and “Sentences have truth conditions, as determined by their meaning” (2003, 61). What’s going on? How can Lepore, like Travis, endorse the radical thesis of the relativity of what is said, and at the same time take himself to be defending Russell’s theory, and “systematic semantics” along with it? Here is how we are to make sense of things: Lepore wants to drive a wedge between, on the one hand, what is said by and/or the truth conditions of *utterances*, and on the other hand, the truth conditions of *sentences*. The following four points are intended to drive home this wedge:

- (a) The truth conditions of a sentence *S* need not correspond to what’s said or stated by an utterance of *S*.
- (b) What’s said by an utterance of *S* can be true, even though the truth conditions for *S* aren’t satisfied (and *vice versa*).
- (c) What’s said by an utterance of *S* can be ‘about’ something (e.g. NYC) even though the truth conditions for *S* makes no reference (to that thing).
- (d) Because of (a)–(c), intuitions about utterances of sentences can in no simple and direct way be used as guides to the truth conditions for those sentences. (2004, 61)

So what Lepore is doing is this: He agrees with radical pragmatics that CSAs succeed in demonstrating that AACM is inadequate with regard to the truth conditions of *utterances*; according to Lepore, what is said by an *utterance* is “no more systematic than determining whether two items are similar” (2004, 66). But, unlike the radical pragmatist, Lepore maintains that AACM is true of sentences relativized to satisfiers for the “basic” context sensitive items. (Perhaps he even takes SACM to deliver truth conditions for sentence types.¹²) It is only with respect to the latter claim then that Lepore’s view differs from that of radical pragmatics: Lepore maintains that *sentences* systematically encode minimal semantic truth conditions in the way described by AACM, whereas radical pragmatics maintains that sentences encode only propositional radicals.

Lepore’s views concerning the problem of incomplete definite descriptions are an instance of the general positive view advanced in Cappelen and Lepore (2005a), and this positive view is composed of Semantic Minimalism (SM), and Speech Act Pluralism (SPAP). SM is, in essence, the view that AACM holds *not* for what is said by utterances – for utterance content is subject to the relativity of what is said – but rather to the semantic content of *sentences*.¹³ Thus, Lepore’s claim that Russell’s theory of definite descriptions is correct as applied to the semantic content of sentence (1) is one instance of SM. SPAP is, in essence, the complementary view that what is said by utterances – but not the semantic content of sentences – is subject to the relativity of what is said.¹⁴ Thus, Lepore’s admission that Russell’s theory of

definite descriptions does *not* apply to what is said by all utterances of (1) – nor even all judgments as to what was said by one utterance of (1) – is one instance of SPAP. Moreover, just as Lepore’s motivation for his views concerning incomplete definite descriptions is provided by CSAs involving different utterances of sentences such as (1), so Cappelen and Lepore’s motivation for endorsing SM and SPAP generally must be provided by CSAs. For recall the plausible claim introduced earlier: the *only* sort of argument that could persuade one to abandon AACM *as applied to what is said by utterances* are CSAs. Hence, just as it is CSAs that force Lepore to sharply distinguish between the semantic content of sentence (1) and the intuitive content of utterances of (1) and to concede that different utterances of (1) will express different intuitive content to different interpreters, so it is CSAs generally that compel Cappelen and Lepore to endorse SM and SPAP.¹⁵

The tension in Cappelen and Lepore’s view arises because, in an attempt to refute radical pragmatics, they *also* formulate a global criticism of all CSAs; they argue that *all* such arguments are unsound. But, if all CSAs are unsound, then there is no motivation for Lepore’s positive view concerning incomplete definite descriptions, nor, more generally, for endorsing the complementary views of SM and SPAP. There is then a considerable amount of tension in Cappelen and Lepore’s overall view: They – perhaps unknowingly – rely on CSAs to motivate their positive view, but they also reject *all* such arguments.¹⁶ And thus the stage is set for the next section, in which I resolve the tension in Cappelen and Lepore’s views by showing why their main criticism of CSAs fails to demonstrate that all CSAs are unsound.

4 Cappelen and Lepore’s Main Argument Against CSAs

Cappelen and Lepore (2003a, 2005a) argue that all CSAs against AACM (i.e. CSAs that do not involve “basic” context sensitive words) are unsound. The essence of their argument is as follows:

1. If there is a sound ICSA against AACM, then this CSA must be capable of being transformed into an intuitively true “real context shifting argument” (RCSA).
2. There are no intuitively true RCSAs against AACM.
Therefore,
3. There are no sound CSAs against AACM.

I will show that this argument is unsound, because the second premise is false. But first some clarification is in order concerning premise 1, and the distinction between “ICSAs” and “RCSAs.”

Cappelen and Lepore offer a buttressing argument in support of premise. First, they claim that a sentence *S* is context sensitive “only if there is a true utterance of an instance of the following schema for Inter-Contextual Disquotation . . .

(ICD) There are (or *can be*) false utterances of $\ulcorner S \urcorner$ even though *S*.

Unless [S] passes this ICD Test, it is *not* context sensitive” (2005a, 105). Second, they maintain that the only way to demonstrate that S passes the ICD test is to tell an intuitively true story in which one utters an instance of ICD; they claim that “it is only through such stories that we should be convinced that [a sentence S] passes the ICD test” (2005a, 108). Now to tell such a story, a story in which one utters an intuitively true instance of ICD and thereby *uses* an instance of S, is to construct a “Real Context Shifting Argument” (RCSA) involving S. In contrast to RCSAs, Impoverished Context Shifting Arguments (ICSAs) are not sufficient to establish that S is context sensitive. In an ICSA one only *mentions* S, and thus one does not utter an instance of ICD; and thus, in turn, one does not establish the context sensitivity of S.

For example, consider again my above stories involving the sentence

(1) The table is broken.

In Cappelen and Lepore’s terminology, these stories constitute merely an ICSA because in telling the stories, I never *used* (1) (or its negation) – I merely mentioned (1). Now Cappelen and Lepore claim – in premise 1 – that if this ICSA is sound, then one ought to be able to transform it into a “real context shifting argument” (RCSA), i.e. a CSA in which I *use*, and do not merely mention, sentence (1). That is, in an RCSA, one must utter a sentence such as

(1a) There are false utterances of ‘The table is broken’ even though the table is broken.

So much for clarifying the first premise. The second premise simply claims that there are no RCSAs in support of radical pragmatics; that is, any attempt to formulate an RCSA for a sentence S which contains no “basic” context sensitive words results in making utterances that are intuitively false, or at least not true. If both premises are true, it follows that there are *no* sound CSAs against AACM. That is, it follows that even seductive ICSAs invoked to refute AACM must contain some hidden flaw.

5 Vagueness and a Reductio of the Main Argument

Before delving into the details of Cappelen and Lepore’s main argument, I want to persuade you that something must be wrong with it by presenting a quick *reductio* of the assumption that it is sound: If Cappelen and Lepore’s main argument were sound, then there would be no non-“basic” vague expressions, or at least we would have no good reason for thinking any non-“basic” expression was vague. A vague expression is standardly defined as one that “admits of multiple precisifications.”¹⁷ To say that ‘bald’ “admits of multiple precisifications” is to say that the satisfaction conditions interpreters associate with ‘bald’ vary from context to context. In some contexts, ‘bald’ is precisified in a such a way that a certain individual, say Jim, satisfies ‘bald’, and in others it is precisified in such a way that he does not. Thus, since ‘bald’ admits of these multiple precisifications, the intuitive truth conditions

of utterances of sentences containing ‘bald’ vary from context to context. And this is simply to point out that what philosophers and semanticists have called “semantic vagueness” is a sort of ubiquitous, non-“basic,” context sensitivity. But according to Cappelen and Lepore’s main argument a sentence such as “Jim is bald” is context sensitive only if it is possible to construct an RCSA involving the sentence. That is, by premise 1, we must be able to tell a story in which there occurs an intuitively true utterance of “There are false utterances of ‘Jim is bald’ even though ‘Jim is bald’ . But, since ‘Jim is bald’ contains no (relevant) “basic” context sensitive words, by premise 2, it is not possible to tell such a story. And hence we are led to the absurd conclusion that ‘Jim is bald’ is not context sensitive, and thus that ‘bald’ does not admit of “multiple precisifications.” Hence, if Cappelen and Lepore’s main argument in support of the claim that there are no non-“basic” context sensitive words were sound, it would follow that there are no non-“basic” vague words. But, obviously, there are such words. So, clearly Cappelen and Lepore’s main argument is unsound.¹⁸ The task is now to locate where in the argument things go awry.

6 Why the Second Premise of the Main Argument Against CSAs Is False

In order to demonstrate that premise 2 of the above argument is false, I must present an intuitively true RCSA.¹⁹ But as preparation for constructing such an RCSA, I would first like to consider a sort of puzzle raised by premise 1. Cappelen and Lepore concede that the radical pragmaticians have succeeded in presenting intuitively true ICSAs; in Chapter 1 of (2005a) they describe seven different *kinds* of intuitively true ICSAs. But Cappelen and Lepore deny that such ICSAs can be transformed into intuitively true RCSAs. And this raises a puzzling question, a question that both radical pragmatics and Cappelen and Lepore need to address: Why would it be more difficult to construct an intuitively true RCSA for a sentence *S* than an intuitively true ISCA for *S* (where *S* contains no “basic” context words)?

The radical pragmatician can answer this question. Indeed, if radical pragmatics is true, then one would *expect* there to some difference in the ease with which intuitively true RCSAs and ICSAs can be formulated. Suppose, as I claim the defender of AACM must admit, the noun ‘home’ is context sensitive, and as a result so are VPs containing the word, e.g., ‘is at home’.²⁰ Suppose a speaker is trying to explain the meaning, and thus the context sensitivity, of the word to a (very) foreign visitor. Obviously, the speaker is going to *mention* the word in giving her explanation, and this will proceed unproblematically:

‘Home’ Lecture

‘home’ refers to different places in different contexts. For example, if John utters ‘I am at home’ then, typically anyway, what he says is true if he, John, is at his permanent place of residence at the time of the utterance. And if, in another

context, Mary utters this same sentence, this will typically mean that Mary is at *her* place of permanent residence, at the time of her utterance. But the context sensitivity of ‘home’ does not end there – ‘home’ does not have a hard-edged “character” as does, e.g., the indexical ‘I’. So, for example, if Mary is visiting John, and John is at his office, and Mary calls John from John’s house, in response to John’s question ‘Where are you?’ Mary might respond by saying ‘I’m at home’ and thereby say that she is at *John’s house*, and not her own house. Moreover, people sometimes use ‘home’ to refer not to their *current* place of residence, but rather to the house in which they grew up. And sometimes ‘home’ refers not to a house, apartment, or other domicile, but rather to a town, state, country or region. So you see, ‘home’ is a very slippery context sensitive word.

So far, so good. This all sounds intuitively true, that is, the above lecture is at least the beginnings of a plausible ICSA for ‘home’. But suppose that speaker of the above CSA now *uses* ‘home’, boldly declaring, in the middle of her lecture,

(2) John went home.

followed by

(2*) There are false utterances of ‘John’ went home’ even though John went home.

I think we all agree that there’s something very odd, and problematic, about uttering either (2) or (2*) – thereby using ‘home’ – in a context in which one has made very clear the context sensitivity of the word. But it is not a deep mystery as to what, roughly anyway, is going on here: ‘home’, *because it is a very slippery context sensitive word*, requires a sufficiently helpful context for unproblematic usage. The context must provide sufficient clues to enable the interpreter to arrive a unique (unique enough for the utterance) interpretation of ‘home’. The context created by the lecture on the slippery context sensitivity of ‘home’ fails to satisfy this adequacy condition. But this is an understatement: in taking pains to make very clear the slippery nature of the word ‘home’, the above lecture *undermines* the possibility of unproblematic interpretation. Generalizing now, contexts in which the slippery context sensitivity of a term is intentionally made manifest undermine the potential for unproblematic *use* of the term. I will call this general phenomenon “the phenomenon of Contextual Muddying.” So, one reason it is more difficult to formulate plausible RCSAs than corresponding ICSAs is because of this phenomenon of Contextual Muddying: in presenting an ICSA one cannot help but muddy the context.

Another, complementary, reason one might be able to formulate an ICSA yet not a corresponding RCSA concerns what Cappelen and Lepore call “Contextual Salience Absorption”: “As soon as you think about a context C, the standards of C affect the standards of the context you’re in (the context in which you do your thinking). Just thinking about C changes the context you are in so that there is no longer a clear distinction in contextual standards between the Storytelling Context and the Target context” (2005a, 114). In terms of formulating an RCSA to demonstrate the

context sensitivity of ‘home’, the problem of Contextual Salience Absorption is this: Suppose the context of the RCSA, the “storytelling context,” is sufficient for an unproblematic *use* of ‘home’. If the storytelling context is sufficient to make salient some semantic value for uses of ‘home’ – if the context is rich enough to “nail down” the semantic values of *uses* of ‘home’ – then this richness and salience is going to bleed into the contexts for the *mentionings* of ‘home’, i.e. the “target contexts.” For example, suppose we are in a context rich enough so that we can unproblematically interpret (2) as saying that John went to Cleveland; in the storytelling context, Cleveland is salient. But if this is so, it makes acceptance of ‘home’ mentionings, such as,

- (3) ‘home’ can be used to refer to cities other than Cleveland, e.g., it can be used to refer to Seattle.

problematic. One can imagine a frustrated interpreter complaining, “Look, *here* and *now* it is clear that home is Cleveland, darn it. So why is it being claimed that ‘home’ could refer to Seattle? Are you suggesting that Cleveland is not home afterall?”

Contextual Muddying and Contextual Salience Absorption are *complementary* reasons for why it is problematic to transform an intuitively true ICSA into an intuitively true RCSA. If the storytelling context is sufficiently rich to unproblematically nail down a use of ‘home’, then – due to Context Salience Absorption – it is going to be *at least somewhat* difficult to make just plain obvious by mentioning ‘home’ that in other utterances the word gets nailed down differently. Conversely, if one makes it obvious by mentionings of “home” that in different utterances the word gets nailed down differently, then – due to Contextual Muddying – it is going to be *at least somewhat* difficult to *use* ‘home’ in such a way that it is unproblematically nailed down. The *general* reason that it is not a trivial matter to transform an intuitively true ICSA into an intuitively true RCSA is that Contextual Muddiness and Contextual Salience Absorption are complementary phenomena: in amending the context so as to lessen the effects of one phenomenon, one automatically augments the effects of the other.²¹

This is not of course to concede that such a transformation cannot be achieved. I maintain that, given an intuitively true ICSA, if one takes pains to overcome the complementary phenomena of Contextual Muddying and Contextual Salience Absorption, one can transform the ICSA into an intuitively true RCSA. That is, premise 2 is false. Cappelen and Lepore express concern that when it comes to premise 2, the debate is in danger of turning into a mere clash of intuitions. Their concern is well-motivated – my intuitions differ from the intuitions Cappelen and Lepore claim to have. In an attempt to persuade you that premise 2 is false, I’m going to do two things: First, I’m going to try to do slightly better than a *mere* appeal to intuitions by formulating an argument in support of the claim that my intuitions are the right ones, thus supporting my claim that premise 2 is false. After presenting this argument, I’m going to stoop to mere intuition mongering: I’m going to formulate an RCSA and ask you to consult your intuitions about it. I think you, like me, will judge the RCSA I formulate to be intuitively true.

Cappelen and Lepore claim that to formulate an RCSA corresponding to the above ICSA involving sentence (1), I must be able to “hear” an utterance of

(1a*) Some utterances of ‘The table is broken’ are true even though it is not the case that the table is broken.

as true. I insist that though (1a*) sounds a little funny – especially in isolation – I “hear it as true.” In an attempt to shake my intuitions, Cappelen and Lepore present several stories about a pretended familiar character called “Rupert.” These stories are supposed to be sincere (I’m sure!), albeit obviously failed, attempts to formulate intuitively true RCSAs. I’m going to argue that my intuitions must be right as follows: I am going to repeat one of Cappelen and Lepore’s (2003a) Rupert stories, “Rich Rupert,” and then I am going to mimic this story as closely as possible, only instead of the controversial context sensitive adjective ‘rich’, my mimicked story, “Home Rupert,” will involve the obviously context sensitive noun ‘home’.²² When these stories are compared, you will have the intuition that my “Home Rupert” story is no *more* plausible than is their story, “Rich Rupert.” That is, ‘home’ and ‘rich’ stand or fall together; if “Rich Rupert” shows that ‘rich’ is not context sensitive, then “Home Rupert” shows that ‘home’ is not context sensitive. But, ‘home’ obviously is context sensitive.²³ Consequently, since “Home Rupert” *cannot* show that ‘home’ is not context sensitive, neither does “Rich Rupert” show that ‘rich’ is not context sensitive.

I here repeat verbatim “Rich Rupert”:

Rich Rupert

Rupert isn’t rich. Anyone who is rich must make more money than 92% of Americans. That’s necessary for being rich. It is impossible to be rich without satisfying this condition. Rupert doesn’t. His income is average. However, one could still utter ‘Rupert is rich’ and express a true proposition, but not because Rupert makes more money relative to other Americans, but rather because in this possible context of utterance, some other comparison class has been rendered salient. This other comparison class, however, is of course irrelevant to whether or not Rupert is rich; again, a person is rich just in case he makes more than 92% of Americans. Still, somehow or other, the salience of this other comparison class (the comparison class that’s irrelevant to whether or not [Rupert] is rich) makes a difference. (2003a, p. 15)

(Note that “Rich Rupert” is allegedly attempting to provide an intuitively true utterance of “There are true utterances of ‘Rupert is rich’ even though Rupert is not Rich.” Thus “Rich Rupert” and the following CSAs utilize a slightly different form of ICD.) And here’s the above, allegedly sincere, attempt at formulating an RCSA for ‘rich’ mimicked for the obviously context sensitive ‘home’:

Home Rupert

Rupert is not at home. Anyone who is at home must be at 703 E Front St., Bloomington IL. That’s necessary for being at home. It is impossible to be at home without satisfying this condition. Rupert doesn’t. He is at 333 Maple Lane,

in Cleveland OH. However, one could still utter ‘Rupert is at home’ and express a true proposition, but not because Rupert is at 703 E. Front St. in Bloomington IL, but rather because in this possible context of utterance, some other location has been rendered salient. This other location, however, is of course irrelevant to whether or not Rupert is at home; again, a person is at home just in case he is at 703 E Front St., Bloomington IL. Still, somehow or other, the salience of this other location (the location that is irrelevant to whether not Rupert is at home) makes a difference.

I think that, though both “Rich Rupert” and “Home Rupert” are odd, and somewhat strained, “Home Rupert” is no less odd sounding than is “Rich Rupert.” It follows that “Rich Rupert” does not demonstrate that “rich” is not context sensitive.

Let me clear about what is going on. I think both “Rich Rupert” and “Home Rupert” sound odd, a bit strained – I want you to have this intuition too. Suppose you do. What follows? We know that ‘home’ is context sensitive. Cappelen and Lepore claim that if an expression is really context sensitive, then there must be an intuitively true, RSCA, a Rupert Story, involving that term. They claim to have demonstrated that ‘rich’ is not context sensitive on the grounds that “Rich Rupert” is not intuitively true. But, since you think that Rich Rupert is no less plausible, or intuitively true, than Home Rupert, it would follow that your intuitive unease concerning “Home Rupert” shows that ‘home’ is not context sensitive, just as your intuitive unease concerning “Rich Rupert” shows that ‘rich’ is not context sensitive. But this result is bogus, since ‘home’ obviously *is* context sensitive.

What are we to make of all this? It important to keep in mind that in order for premise 2 to be true, it must be *impossible* to construct an intuitively true RSCA. I grant that the Rupert stories sound a bit strained. But this only shows that some half-hearted attempts at constructing RSCAs might fail. The funniness of “Home Rupert” does not show that ‘home’ is not context sensitive, and the funniness of “Rich Rupert” does not show that ‘rich’ is not context sensitive. What we must do, on pain of denying the obvious context sensitivity of ‘home’, is do a better job telling the Rich Rupert story, making it plausible and intuitively true. Here goes:

Rich Rupert II

The adjective ‘rich’ is context sensitive, not in an obvious way like ‘I’ or ‘now’, but rather in a more subtle way. In some contexts, say in a sociology seminar discussion concerning the economic status of members of an Amish community, ‘rich’ might apply to a person if they have several healthy horses, and some fertile farm-land. But in a discussion concerning how much Hollywood stars are paid, a correct application of this predicate to someone might require that the person have much more than horses and some fertile land. So, the predicate ‘is rich’ is context sensitive; in different contexts it has different satisfaction conditions. (It “admits of different precisifications.”) Given this context sensitivity, every sentence of the form ‘NP is rich’ is context sensitive: That is, in different contexts it will express different truth conditions. In particular, in different contexts utterances of ‘Rupert is rich’ will express different truth conditions, and thus it is

at least possible that in some contexts an utterance of ‘Rupert is rich’ says something true, and in other contexts an utterance of this sentence says something false. Let us give the sentence type ‘Rupert is rich’ the name, ‘ α ’. So, our conclusion thus far is that some utterances of α are true, while other utterances of α are false. Now, let’s just stipulate that here and now, i.e. in *this* context, one can apply ‘is rich’ to a person only if they have as much money as Dick Cheney has right now – this criterion is arbitrary, I know, but given that I have just been talking about the context sensitivity of ‘is rich’, if I am going to *use* ‘is rich’ to say something sufficiently clear, I’m going to have to be a bit heavy handed. (After all, I must work to overcome the effects of Contextual Muddying, and Contextual Saliency Absorption.) Indeed, given the somewhat odd linguistic task I am performing right now, everything is going to sound a bit heavy-handed. Ok, so I am about to use ‘is rich’ to say something false – if you don’t interpret my forthcoming utterance in such a way that it says something false, then I insist that you are misinterpreting me. Oh, one more thing – just believe me when I tell you that Rupert has less money than Dick Cheney. Here it is: Rupert is rich. That, *partially by stipulation*, is false. But you agreed above, remember, that ‘rich’ is context sensitive – and it follows that some utterances of α are false, while other utterances of α are true. So, it follows from what you’ve agreed to that in some other contexts, wherein ‘rich’ has other satisfaction conditions, an utterance of α is true. Yet, since I have stipulated and demanded that in *this* context ‘rich’ applies to a person only if they have as much money as Dick Cheney, and poor Rupert does not have this much money, I speak falsely when I say, here and now, that Rupert is rich. Thus, there are true utterances of α , even though, and I insist on this, Rupert is not rich.

As clumsy and heavy handed as this RSCA is, it is intuitively true.²⁴

At this point the debate has devolved into an appeal to intuitions, and I cannot *argue* against Cappelen and Lepore that my intuitions are correct. I can, however, argue that Cappelen and Lepore are not compelled to reject my intuition – accepting it would be fully compatible with their positive view. Moreover, given that Cappelen and Lepore utilize CSAs to motivate their positive view, they ought to *want* my intuitions to be correct.

If Cappelen and Lepore did share my intuitions concerning Rich Rupert II, and thus admit that there can be intuitively true RCSAs against AACM, *this admission would not require them to amend their positive views in any way whatsoever*. For both SPAP and SM are fully compatible with conceding the intuitive truth of some RCSAs. SPAP, which incorporates the radical doctrine of the relativity of what is said, is certainly compatible with the intuitive truth of some RCSAs. And SM is fully compatible with the intuitive truth of some RCSAs because thesis (7) of SM states that “intuitions about, and other evidence for, speech act content are not direct evidence for semantic content: an intuition to the effect that an utterance *u* said that *p* is not even prima facie evidence that *p* is the proposition semantically expressed by *u*” (2005a, 145).²⁵ So, Cappelen and Lepore can accept my intuitions concerning the *utterances* in Rich Rupert II and still maintain – in keeping with SM

– that the *sentence* ‘Rupert is not rich’ invariantly encodes a minimal proposition (ignoring tense). Generalizing now, Cappelen and Lepore’s positive views are fully compatible with the intuitive truth of RCSAs.

But, more significantly, Cappelen and Lepore should want to agree with my intuitions, and they should want to accept the general claim that it is possible to formulate intuitively true RCSAs. For, as we have seen, they utilize CSAs to motivate their positive view, the combination of SPAP and SM. Another way to see this point is to appreciate that if SPAP were true, then one would fully expect there to be intuitively true RCSAs. If one endorses SPAP, then one maintains that the satisfaction conditions for what is said by utterances of S vary depending upon “non-semantic” factors concerning both contexts of utterance and contexts of interpretation. If this is granted then it seems that one has no choice but to agree that there are *some* intuitively true utterances of the form

(ICD*) There are utterances of ‘ $\lceil S \rceil$ ’ involving other contexts of utterance and other contexts of interpretation which are, relative to those contexts, correctly interpreted as being false, even though, given our context of utterance and context of interpretation, it is true that S.

In other words, SPAP seems to entail the possibility of intuitively true RCSAs.

7 Conclusion: Cappelen and Lepore Should Endorse CSAs

It is now clear how the tension between Lepore and Cappelen’s positive and negative views ought to be resolved: Cappelen and Lepore should abandon their negative view that CSAs are unsound. For first, the intuitive truth of some RCSAs is fully compatible with SM, for SM concerns the semantic content of sentences and this rather abstract theoretical content is not the subject of our intuitive judgments concerning what speakers say in making utterances. And second, CSAs provide the reasons for endorsing SPAP, and, conversely, SPAP seems to require the intuitive truth of some RCSAs.

But this proposed resolution leaves one crucial question unanswered: If CSAs are not only compatible with Cappelen and Lepore’s positive views, but moreover are the real motivation for their positive views, then why do Cappelen and Lepore feel compelled to criticize CSAs *at all*?

I think I can answer this question. Consider the following passage:

Nothing that [radical pragmatics] has so far argued shows, for example, that (5*) is not true.

(5*) ‘Smith weighs 80 kg’ is true iff Smith weighs 80 kg.

... consider whether alleged context shifting shows that (5*) is not true. (5*) is not true just in case it is either false or lacks a truth value. With regards to falsity, context shifting doesn’t establish that for any utterance of (5*), its LHS and RHS can disagree in truth value. If Smith weighs 80 kg, [‘Smith weighs 80 kg’] is true; and if he doesn’t, [‘Smith weighs 80 kg’] is false. ... If your intuition is

that [*'Smith weighs 80 kg'*] is false, your intuition will also be that he doesn't weigh 80 kg. It's exceedingly hard for us to see how to pull these intuitions apart. (2005b, 56)

This passage suggests that CSAs and truth conditional semantics are incompatible. According to the version of truth conditional semantics favored by Cappelen and Lepore, a semantic theory is to take the form of a Davidsonian T-theory, and such theories yield theorems such as (5*). Hence, Cappelen and Lepore are compelled to defend the truth of (5*). But, a sound CSA involving "Smith weighs 80 kg" would demonstrate that utterances of "Smith weighs 80 kg" have different truth values in different contexts. And thus it would seem to follow from such a sound CSA that (5*) is not necessarily true, that is, that some utterances of (5*) are not true. So, first, it appears that if truth conditional semantics is correct, then CSAs must be unsound. Moreover, second, Cappelen and Lepore suggest that it is "exceedingly hard" to see how an utterance of (5*) could not be true. So, it appears that CSAs and truth conditional semantics are incompatible, and, moreover, that truth conditional semantics must be correct.

But the appearance is an illusion. First and most importantly, note that even if Cappelen and Lepore could bring themselves to "pull these intuitions apart," and thus intuitively judge of an utterance of (5*) that it was not true, this would *not* require them to reject truth conditional semantics. Because Cappelen and Lepore endorse truth conditional semantics, they are committed to the truth of the *semantic content* of sentence (5*). But, since Cappelen and Lepore maintain that "an intuition to the effect that an utterance *u* said that *p* is not even prima facie evidence that *p* is the proposition semantically expressed by *u*" (2005a, 145), this commitment to the truth of the semantic content of (5*) does *not* commit them to the intuitive truth of every *utterance* of (5*). Cappelen and Lepore fail to appreciate that by sharply distinguishing between intuitive utterance content and abstract sentence content, they insulate their semantic theory from refutation by CSAs. And hence they have no reason to reject CSAs. Second, the above passage suggests that in order to formulate a CSA involving '*Smith weighs 80 kg*', one must create an utterance of (5*) that is intuitively false. But this is not correct. Sentence (5*) applies the predicate 'is true' to a sentence, and not to an utterance; but to formulate a CSA involving '*Smith weighs 80 kg*' is to elicit the intuitive judgment that utterances of the sentence in different contexts will express different truth conditions. Hence, to formulate a CSA involving '*Smith weighs 80 kg*' it is not necessary to elicit the intuition that an utterance of (5*) is not true; rather, it is only necessary to elicit the intuition that an utterance of

(5**) Every utterance of '*Smith weighs 80 kg*' is true iff Smith weighs 80 kg.

is not true. And, to repeat, since Cappelen and Lepore endorse SPAP, it seems they *must* agree that there are intuitively untrue utterances of (5**). For according to SPAP, "No one thing is said (or asserted, or claimed, or . . .) by any utterance: rather, indefinitely many propositions are said, asserted, claimed, stated, etc." (2005a, 199). According to SPAP then, not only do different utterances of '*Smith weighs 80 kg*'

express different intuitive truth conditions, but moreover *one utterance* will express different intuitive truth conditions to different interpreters.

Suppose then that Cappelen and Lepore accept my proposal for resolving the tension in their views, and thus they reject their negative claim that all CSAs are unsound. What then would be the difference between Cappelen and Lepore's version of truth conditional semantics, and radical pragmatics? Both sides now accept that CSAs succeed in undermining AACM *as applied to the intuitive truth conditions of utterances*. That is, both sides accept the radical context sensitivity of what is intuitively said by utterances that is encompassed by SPAP. Moreover, both sides accept that AACM is adequate for providing a semantic core, a linguistically encoded minimal representation that provides interpreters with a crucial "starting point" (2005a, 185) for determining what is intuitively said by utterances. Both sides agree that such contextually invariant encoded information serves as the "shared fallback content" that is further enriched and developed with context specific information to yield what is said by an utterance. Thus both sides agree that the task of a pure semantic theory is to delineate the purely semantic *constraints* that the invariant meaning of a *sentence* places on the intuitive truth conditions expressed by, or what is said by, utterances of that sentence. The only difference concerns the nature of this minimal semantic core: Radical pragmatics, which endorses RPM, denies that sentences (even sentences relative to contexts) encode truth conditions, and instead maintains that sentences encode what I earlier called, following Bach, "propositional radicals." Cappelen and Lepore, in contrast, maintain that sentences (relative to contexts) encode minimal semantic truth conditions. However, given that these minimal semantic truth conditions are neither (a) immediately accessible to the intuitions of interpreters, nor (b) identical to the intuitive truth conditions of what is said, it is not clear that there is a significant disagreement here.

Notes

1. A very similar positive view is advanced in Soames (2002) and rightly, in my view, rejected in Soames (2005).
2. Cappelen and Lepore resist the suggestion that "deep down it is all about compositionality" ((2005a, 12). There may be no fact of the matter as to what it is "all about deep down," but there can be no doubt that semantic compositionality is central to the debate: Truth conditional semantic compositionality for *sentences* is essential to Cappelen and Lepore's Semantic Minimalism (2005a, 144–5), and the core idea of their Speech Act Pluralism (2005a, 199–204) is that such semantic compositionality does not apply to what is said by *utterances*. Moreover, as will be demonstrated, the arguments in support of radical pragmatics that Cappelen and Lepore reject, viz. CSAs, are arguments against truth conditional semantic compositionality for utterances.
3. I borrow the suggestive phrase 'assign and combine' from Reimer (2002).
4. One might suggest that Grice's Strategy and Kaplan's Strategy do not exhaust the possible strategies of response. In particular, one might deny that clause (a) is satisfied on the grounds that S and S' are not "relevantly similar" because they do not have the same relevant syntactic structure (LF). Indeed, one might point out that this response is favored for sentences exhibiting quantifier-scope ambiguities. That is, because of scope ambiguities, 'Everybody loves somebody' seems to express different truth conditions in different contexts, and so it

seemingly gives rise to CSAs. The standard response in defense of AACM is to defang the alleged counterexample by positing two LFs, each of which corresponds to the “sentence” ‘Everybody loves somebody.’ The suggestion assumes that it is possible to have one “sentence” associated with two LFs. For the sake of clarity I hereby stipulate that this is impossible: if you have distinct LFs, you have distinct sentences.

5. In the pseudo-formal representations of SACM and AACM, ‘ \rightarrow ’ represents a heuristic, informal, sort of mental process, whereas ‘ \Rightarrow ’ represents an algorithmic, formal, sort of mental process. For example, according to AACM in interpreting an utterance of ‘John is here’ the process whereby one determines the proper interpretation of ‘here’ is somewhat “fuzzy,” requiring inductive inferences based upon contextual “clues.” But once the LF of the sentence, and the semantic values are – perhaps by heuristic processes – fixed, it is automatically *determined* by the semantic rules what the truth conditions of the utterance are. I do not intend this to be precise, but some such distinction seems to be essential to the semantic minimalism defended by Cappelen and Lepore, for in the end it is by appeal to some such distinction between kinds of mental processing that the pragmatic content of utterances is to be distinguished from the semantic content of sentences. Indeed, Borg (2004) argues that *only* “formal” processes can be involved in an interpreter’s determining the semantic content of a sentence.
6. Cappelen and Lepore use the phrase ‘basic context sensitive expressions’ to designate those context sensitive expression that they concede are really, “semantically,” context sensitive. These include only the pronouns, indexicals and demonstratives with which Kaplan was concerned, and perhaps some other terms such as ‘local’ that are widely recognized as context sensitive. (See Cappelen & Lepore, 2005, p. 2.) I will need a term to designate the class of expressions that Cappelen and Lepore concede are context sensitive, so I will follow them in this use of ‘basic’; but I will always use the term in scare-quotes to indicate that I do not necessarily think these are the only context sensitive terms.
7. Cappelen and Lepore distinguish between CSAs and what they call “Incompleteness Arguments” (2005a, 33–8). Incompleteness Arguments, however, are sub-species of CSAs, which in turn are a species of counterexamples against AACM.
8. So for example, ‘John is here’ in C is *relevantly similar* to ‘He is there’ in C’ just in case the sentences have the same LF (or their LFs have same structure), and the words in the sentences, including the context sensitive words, are assigned the same semantic values in both C and C’. So, for example, ‘John is here’ in C is relevantly similar to ‘He is there’ in C; only if the semantic value of ‘John’ in C is identical to the semantic value of ‘he’ in C’, and the semantic value of ‘here’ in C is identical to the semantic value of ‘there’ in C’. (I continue to ignore tense.)
9. Two scholarly caveats are in order: First, Lepore (2004) is *not* co-authored by Cappelen. Nonetheless this paper presents a clear application of the general theses advanced by works that are co-authored by Cappelen and Lepore (including 2005a), and thus I will treat the paper as if it were co-authored. Second, in the last footnote of Lepore (2004) he states that in the time intervening between his writing the paper and its publication he has “in places changed his views” and he suggests that Cappelen and Lepore (2005a) represents his current views. However, I can detect no differences between the views espoused in Lepore (2004), and those espoused in Cappelen and Lepore (2005a) that are relevant to my purposes.
10. That the problem is really that incomplete definite descriptions give rise to CSAs is just barely discernible in Lepore’s brief sketch of the problem:

In short, unexceptional uses of incomplete definite descriptions that are not singular terms are endemic. But the traditional view that all definite descriptions are quantifiers contributing only general uniqueness conditions to propositions expressed by their use is jeopardized by incomplete ones *if* they on occasion denote. So, what are we theorists of language to do in the face of these commonplace linguistic facts? (2004, 43)

Note that, in the terms of Cappelen and Lepore (2005a), this passage, at best, sketches an “ICSA” against AACM; the passage falls far short of constituting a full-blown “RCSA.” Nonetheless, Lepores take the “problem” presented in the passage as the central motivation for his positive, and radical, view concerning incomplete definite descriptions.

11. To see that the problem essentially involves CSAs, suppose it was *impossible* to use (1) such that (1) had the truth conditions predicted by Russell’s theory. If this were the case, then, first, we (as defenders of “systematic semantics”) would simply *reject* Russell’s theory. And consequently, second, we would not consider incomplete uses of definite descriptions to be threats against “systematic semantic theories,” i.e. against AACM.
12. It is not clear whether or not Lepore endorses the adequacy of SACM for the truth conditions of sentences. It seems initially unlikely that he would, but he does write as if *sentences* have truth conditions wholly independently of context. For example, Lepore states that “Sentences have truth conditions, as determined by their meanings” (2004, 61) And this statement is made apparently in contrast to Soames’ statement that “the meaning of a sentence can be thought of as a function from contexts to what is said by the sentences in those contexts” (Soames, 1989, 394). The lack of clarity here may result from a concern that there is no theoretical motivation for endorsing AACM *as a model for the truth conditions of sentences*, yet denying SACM *as a model for the truth conditions of sentences*. For recall that our intuitions about what is said by utterances of, e.g., ‘I am hungry’ are not directly relevant to the truth conditions of the sentence ‘I am hungry’. So, why not claim that the sentence ‘I am hungry’ expresses the same minimal “proposition” in every context, just as – as Cappelen and Lepore (2005a, 2005b) suggest – the sentence ‘It’s raining’ expresses the same minimal “proposition” in every context (ignoring tense)?
13. In (2005a), pp. 144–5, Cappelen and Lepore define SM in terms of seven theses. Thesis (5) is an explicit endorsement of AACM (as applied to sentences).
14. In (2005a), pp. 199–204, Cappelen and Lepore define SPAP in terms of eight theses. Thesis (1) and (4) constitute an explicit endorsement of the contextual relativity of what is said.
15. In (2005a), pp. 190–6, Cappelen and Lepore develop a sophisticated CSA involving Nixon’s Watergate tapes. They demonstrate that not only are there sentences such that we intuitively judge that different relevantly similar utterances express different truth conditions, but moreover that – because of the relativity of what is said – there are *single utterances* such that in different contexts we intuitively judge of the same utterance that it expresses different truth conditions.
16. In a very revealing passage Cappelen and Lepore write:

If there really were (or had to be) a close connection between speech act content and semantic content, then all the data we think support Speech Act Pluralism would also serve to undermine Semantic Minimalism. That’s how some of the most clearheaded [radical pragmaticians] argue. Our strategy is to endorse the data they invoke, but undermine their assumption that this data has semantic implications. (2005, 4)

Cappelen and Lepore are here admitting that it is the intuitions invoked by CSAs that support SPAP. But they fail to take the next step: If their attempt to undermine all CSAs is successful, then they thereby undermine all their support for SPAP. (And the passage begs the question against the clearheaded radical pragmatician: The clearheaded radical pragmatician neither accepts nor rejects the claim that “there [is] a close connection between speech act content and semantic content” (MA); rather, the clearheaded radical pragmatician denies that there is any such thing as “semantic content,” as Cappelen and Lepore utilize the term; i.e. the clearheaded radical pragmatician denies that *sentences* encode truth conditions.)

17. That Cappelen and Lepore accept this standard definition of semantic vagueness is suggested by their, to my mind extremely puzzling, remark that “the effects of context on assigning truth conditions to an utterance are limited to fixing the values of context sensitive linguistic items, precisifying vague terms, and disambiguating ambiguous strings in the sentence uttered” (2003b, 70).

18. It is telling that Cappelen and Lepore attempt to set vagueness aside, as if vagueness were irrelevant to the debate between the radical pragmaticians and the defenders of AACM. They claim, or at least suggest, that the sentences put forth as counterexamples to AACM must be “unambiguous, non-vague, [and] non-elliptical.” (See 2003a, pp. 1, 4.) But this request to set vagueness aside rests on confusion. For example, Travis’ examples concerning color words are designed to show that AACM is incompatible with the *vagueness* of such words. One might go so far as to say that the radical pragmaticians reject the AACM *because* it is incompatible with the ubiquitous and multifaceted phenomenon of semantic vagueness. Bach (2003) at least hints at this sort of reductio.
19. I think the first premise is also flawed, but I am granting it here for the sake of argument. Premise 1 is flawed because it presupposes that the only way to refute AACM is to utter a true instance of ICD thereby showing that some sentence *S*, (where contains no “basic” context sensitive words) is context sensitive. But this assumption is false. For AACM entails claims of the following form:

What is said by every utterance of $\ulcorner S \urcorner$ is the same.

Since such claims only mention and do not use $\ulcorner S \urcorner$ there is no reason to suppose that intuitively true ICAs that mention but do not use $\ulcorner S \urcorner$ would not suffice to establish their falsity.

20. I am assuming that, upon reflection, Cappelen and Lepore would admit ‘home’ into the set of “basic” context sensitive words. As with Stainton (2003), an underlying theme of the present paper is that Cappelen and Lepore, as with other defenders of the AACM, ignore words that – at least upon reflection – are *obviously* context sensitive, but are not included in the “basic” context sensitive indexicals and demonstratives discussed by Kaplan.
21. Cappelen and Lepore maintain that the radical pragmatician cannot appeal to “Contextual Salience Absorption” to explain why it is possible to construct a plausible ICSA, yet not possible to construct a plausible RSCA:

Context Shifting Arguments are supposed to convince [defenders of AACM] that their position is flawed. The evidence takes the form of a counterexample: intuitive evidence that there is at least one utterance *u* of *S* that semantically expresses a different proposition or has different truth conditions . . . than another utterance of *S* in the Storytelling Context. But to endorse Contextual Salience Absorption is in effect to grant that no intuitive counterexample is forthcoming. (2005a, 115)

This point, however, would apply only to a radical pragmatician who accepted both premises of Cappelen and Lepore’s argument.

22. “Rich Rupert” does not appear in Cappelen and Lepore’s more recent (2005a). My point could have been made with “Known Rupert,” “Red Rupert,” or “Skinny Rupert,” which do appear in (2005a), but for reasons that are not important “Rich Rupert” is more convenient.
23. Note that ‘home’ passes Cappelen and Lepore’s tests for context sensitivity. For example, if John utters ‘Somebody is at home’ in a context *C* and in a different context $\ulcorner C$ Mary reports John’s utterance with ‘John said that somebody is at home’ then it “will be just an *accident*” (2005a, 89) if they refer to the same place with their uses of ‘home’. But “home” is not context sensitive in a Kaplanesque way, not in the same way that, e.g., ‘I’ is context sensitive. For one thing, ‘home’ seems to be in some sense governable by “monsters”: Compare, e.g., ‘Home is not a happy place’ and ‘For John, home is not a happy place.’
24. Stainton also points out that there are many resources useful for constructing intuitively true RCSAs that are not utilized in Cappelen and Lepore’s Rupert stories. Stainton points out that “Caveats, ‘Pssts’ and other devices thus afford [construction of intuitively true RCSAs]” (2003, pp. 10–1). Stainton seems to concede to Cappelen and Lepore – incorrectly in my view – that all such devices are “monsters,” or what Stainton calls “pseudo-operators.” The issue of whether or not there are “monsters” in natural language is very much relevant to the soundness of CSAs, but I do not have the space to address it here.
25. Perhaps this point can be made more perspicuously as follows: Cappelen and Lepore reject “the mistaken assumption” (MA) that “a theory of semantic content is adequate just in case it

accounts for all or most of the intuitions speakers have about speech act content” (2005a, 53). Because they reject MA, appeals to intuitions concerning what was said by utterances cannot be used to refute their semantic theory, viz. SM. So, there is no reason for them to reject CSAs, since CSAs crucially involve intuitions concerning what is said.

References

- Bach, K. (1994). Conversational implicature. *Mind and Language*, 9, 124–62.
- Bach, K. (2003). Comments on Herman Cappelen and Ernie Lepore, ‘Context Shifting Arguments’. Central APA, Cleveland OH, April 26, 2003.
- Bezuidenhout, A. (2002). Truth conditional pragmatics. In J. Tomberlin (Ed.), *Philosophical perspectives, 16, language and mind*.
- Borg, E. (2004). *Minimal semantics*. Oxford: Oxford University Press.
- Cappelen, H., & Lepore, E. (2003a). Context shifting arguments. Forthcoming in G. Preyer & G. Peters (Eds.), *Contextualism and philosophy*. Oxford: Oxford University Press.
- Cappelen, H., & Lepore, E. (2003b). *Unarticulated constituents and hidden indexicals: An abuse of context in semantics*. Rutgers University Manuscript.
- Cappelen, H., & Lepore, E. (2005a). *Insensitive semantics: A defense of semantic minimalism and speech act pluralism*. Oxford: Blackwell Publishing.
- Cappelen, H., & Lepore, E. (2005b). Radical and moderate pragmatics: Does meaning determine truth conditions? In Z. Gendler Szabó (Ed.), *Semantics versus pragmatics*. Oxford: Oxford University Press.
- Carston, R. (2002). *Thoughts and utterances: The pragmatics of explicit communication*. Oxford: Blackwell Publishing
- Crimmins, M. (1992). *Talk about beliefs*. Cambridge: MIT Press.
- Lepore, E. (2004). An abuse of context in semantics: The case of incomplete definite descriptions. In A. Bezuidenhout & M. Reimer (Eds.), *Descriptions and beyond: An interdisciplinary collection of essays on definite and indefinite descriptions and other related phenomena*. Oxford: Oxford University Press.
- Pietroski, P. (2005). Meaning before truth. In G. Preyer & G. Peters (Eds.), *Contextualism and philosophy*. Oxford: Oxford University Press.
- Recanati, F. (2004). *Literal meaning*. Cambridge: Cambridge University Press.
- Reimer, M. (2002). Do adjectives conform to compositionality? In J. Tomberlin (Ed.), *Philosophical perspectives, 16, language and mind*.
- Salmon, N. (1886). *Frege’s puzzle*. Cambridge: MIT Press.
- Searle, J. (1980). The background of meaning. In J. Searle, F. Kiefer, & M. Bierwisch (Eds.), *Speech act theory and pragmatics*. Dordrecht: Reidel.
- Soames, S. (1989). Direct reference, propositional attitudes, and semantic content. *Philosophical Topics*, 15, 47–87.
- Soames, S. (2002). *Beyond rigidity: The unfinished semantic agenda of naming and necessity*. Oxford: Oxford University Press.
- Soames, S. (2005). Naming and asserting. In Z. Gendler Szabó (Ed.), *Semantics versus pragmatics*. Oxford: Oxford University Press.
- Sperber, D. & Wilson, D. (1986). *Relevance*. Oxford: Oxford University Press.
- Stainton, R. (2003). Comments on Cappelen and Lepore’s context shifting arguments. Central APA, Cleveland OH, April 26, 2003.
- Stainton, R. (2005). In defense of non-sentential assertion. In Z. Gendler Szabó (Ed.), *Semantics versus pragmatics*. Oxford: Oxford University Press.
- Stanley, J. & Szabó, Z. (2000). On quantifier domain restriction. *Mind and Language*, 15/2, 219–61.
- Travis, C. (2000). *Unshadowed thought: Representation in thought and language*. Cambridge: Harvard University Press.

Contextualism, Skepticism and Objectivity

David Hunter

Abstract In this paper, I try to make sense of the idea that true knowledge attributions characterize something that is more valuable than true belief and that survives even if, as Contextualism implies, contextual changes make it no longer identifiable by a knowledge attribution. I begin by sketching a familiar, pragmatic picture of assertion that helps us to understand and predict how the words “S knows that P” can be used to draw different epistemic distinctions in different contexts. I then argue that the examples provided by Cohen and DeRose meant to illustrate Contextualism fail to do so, and I construct an example that does. I conclude by considering the response that an objective assessment of skepticism depends, not on what we might use sentences of the form “S knows that P” to say, but on what such sentences themselves say—on their literal, context-invariant meanings. I argue that there is little reason to believe that our words have such context-invariant meanings, and I suggest that the pragmatic picture of assertion can secure a rich enough conception of objectivity to address the skeptic.

Keywords Knowledge · assertion · contextualism · objectivity · skepticism

Contextualism in epistemology is the view that the truth conditions of knowledge attributions depend on context because the truth conditional contribution of “knows” in sentences like “S knows that P” depends on context. Contextualists disagree about which contextual factors are relevant. According to some, the truth conditions of a knowledge attribution depend on the standards of evaluation in the context of attribution, while others say it depends on which possibilities or contrasts are presupposed in that context.¹ Contextualists also disagree about what linguistic mechanisms explain the context dependence. Some posit underlying syntactic or semantic complexity in the word “knows”, while others explain the context-dependence in terms of pragmatic mechanisms like those involved in

D. Hunter

Philosophy Department, Ryerson University, 350 Victoria Street, Toronto, ON, M5B 2K3, Canada

conversational implicature.² But all Contextualists agree that even if many ordinary knowledge attributions are true there are some very special contexts where very few if any will be true. In those special contexts, exemplified by the skeptic's fantastic doubts, the truth conditional contribution of the word "knows" makes true knowledge attributions virtually impossible. In fact, some Contextualists hold that merely mentioning these possibilities has this same result. As David Lewis put it in his defense of Contextualism, knowledge is elusive: doing epistemology with a skeptic seems to make our knowledge go away.³

But one might wonder how we could have, and why we should value, knowledge that is elusive in this way. If a person's grip on the facts were really of the right kind and strength for knowledge, then how could merely humoring the skeptic who mentions the bare possibility of her being tricked by, say, an evil demon weaken this grip? Just what was her relation to the facts, and how strong could it have been, if it is so easily loosened? And why should we value knowledge if it is so easily lost? True belief, presumably, is not altered or weakened merely by humoring the skeptic. Nothing the skeptic says can force a change in a person's beliefs let alone affect their truth. So why not prefer stable true belief to elusive knowledge? Why value knowledge at all, if it must be elusive?

Part of the Contextualist reply to these questions is that they involve a subtle use-mention confusion.⁴ What changes when the skeptic raises her fantastic possibilities, the Contextualist says, is not our grip on reality, but only our means for describing that grip. The grip remains wholly unchanged in nature and value, but the linguistic resources we once had for identifying and characterizing it, resources that included the word "knows", are no longer of the same use once the skeptic's doubts are raised. The grip remains knowledge in every respect but name. Contextualism is a thesis, not about our cognitive relations to the world, but about our resources for identifying those relations, and context-dependence affects, not our relations to the world, but only our means for saying what those relations are.

The aim of this paper is to explore this Contextualist reply. How should we make sense of the idea that true knowledge attributions characterize something that is more valuable than true belief and that survives even when it can no longer be identified by a knowledge attribution? Contextualists typically address this by providing examples of the kinds of contextual changes they have in mind. But I want to approach the issue at a more abstract level. I begin by sketching a somewhat idealized picture of context-dependence and then a minimalist account of what makes knowledge more valuable than true belief. I will then describe two uncontroversial respects in which knowledge attributions are context-dependent. Seeing these will allow us to understand the more radical kind of context-dependence that Contextualists claim is involved when the skeptic raises her challenge. I will argue that this claim is more subtle and more interesting than a simple charge of confusing use with mention.⁵ For it involves also challenging a certain view of what an objective conception of our cognitive relations to the world would be like.

1 Assertion and Context

The picture of context-dependence I will describe is a familiar one that fits assertion into a broader conception of inquiry.⁶ On this picture, the goal of inquiry is to identify which of the many ways things might be is the way they are. Progress towards that goal requires drawing a distinction among ways things might be and locating the actual world on one side of the divide. Picturesquely, we can think of this as drawing a distinction among possible worlds, where each world represents a way things might be. In judging or asserting something one marks off some of those worlds from the rest. To judge or assert that P is to judge or assert that the set of P worlds includes the actual world.

One might identify the content of a judgment or assertion with the ways things would be if the judgment or assertion were true, or, more accurately, with the ways things would be if they were to be as they were judged or asserted to be. Abstractly, this is to identify the contents with a set of possible worlds: those where things are as they were judged or asserted to be. On this view, equivalent judgments or assertions would have the same content. Many philosophers, though, prefer to individuate contents more finely so that different contents could have the very same truth conditions. For our purposes, though, we can identify contents with sets of worlds. For on any account, the content of an assertion or judgement determines truth conditions and thereby a division among ways things might be. So, on any account of judgement or assertion, to judge or assert is at least in part to draw a distinction among possibilities. More importantly, though, Contextualism is at heart a thesis about the truth conditions of knowledge attributions: about the role context plays in determining truth conditions. So while a more complete account of such attributions might need to individuate their contents more finely than by truth conditions alone, I doubt this will be needed for my purposes here.

I said that to judge or assert something is to draw a distinction among a set of possibilities. Which set? A natural idea, and one that introduces context dependence, is that it is the set of worlds left open at that point in the inquiry.⁷ At all but the very last stages of an inquiry, some questions have been settled and others remain open. We can represent this state of the inquiry by a set of worlds. What the worlds in the set have in common represents answers to settled questions. If all the worlds are ones where P is the case then this represents the fact that the inquirer(s) has decided, at least provisionally, that P is true. Differences among the worlds represent questions that have yet to be answered. If only some are Q worlds then this represents the fact that the inquirer(s) has not yet decided whether Q is true. The goal of inquiry is to answer those open questions.

As the inquiry proceeds and questions are answered the set of worlds changes. One kind of change is that some worlds may be eliminated from the set. After judging that P, for instance, the new set of worlds contains only P worlds. As a consequence, the set of worlds among which the participant's judgments and assertions distinguish changes as well. Any subsequent judgment will distinguish among the worlds in that set, all of which are P worlds. The effect of an *assertion* that P is more complex when it occurs as part of a communal inquiry since the other

participants might not accept the assertion. But when an assertion that P is accepted, the new set contains only P worlds and any subsequent assertion will distinguish among the worlds in that set, all of which are P worlds.

There are also more subtle changes to the set of worlds. For the agreement among the inquirers concerns the way the world is, and so concerns their inquiry and, more specifically, the linguistic resources they have for pursuing it. After an assertion that P is accepted, the initial set of worlds is changed by eliminating all the not-P worlds. Since all sides also agree that that assertion was accepted, every world in the changed set will be one where that assertion occurred and was accepted. And if the sentence S was used in asserting that P, then every world in the new set will also be a world where S is (or could be used to assert something that is) true in all the worlds. As the inquiry proceeds, the set of worlds changes to reflect changes in the participants' shared attitudes about the world and about their relations to it, including their linguistic relations to it.

I have described the elimination of worlds from the set. But the set of worlds can also change through the addition of a world. Such a change would occur when the inquirers decide that a question needs to be re-opened or that some question has been overlooked. If they decide that they were hasty when they agreed that P, they may decide to re-introduce not-P worlds into the set. The more an added world differs from the worlds already in the set, the greater the change to the participants' shared views. And as with the elimination of worlds, adding a world reflects a change in attitudes both about the world and about the inquirer's linguistic relations to it. Accepting an assertion changes not just what the participants agree on about the facts, but also what they agree on about how they can use their language to state the facts.

This picture suggests that some kinds of assertions will seem odd. To assert something that is true in all of the worlds in the set will seem odd because one will have not ruled out any possibility. Likewise, to assert something that is false in all of the worlds will seem odd because one will have ruled out all the open possibilities. Given that the goal of inquiry is to rule out some worlds until only one remains, both kinds of assertion frustrate the goal. This is already a familiar feature of this account of assertion. It is also familiar that participants in an inquiry can and typically do re-interpret such speech acts and that speakers can exploit this fact. The subtle dynamics of assertion and accommodation have been the subject of much study.⁸

A related detail is worth noting. If all of the worlds in the set are P worlds, then an assertion made in that context will be true only if P is true. Does this mean that the speaker will have asserted that P? In that case, she will have asserted something that is true in all the worlds and, for reasons just sketched, this is typically an odd thing to do. One might instead distinguish what a speaker presupposes to be the case from what she asserts (or judges) to be the case. If all the worlds in the set are P worlds, then the speaker is *presupposing* that P is the case, and in subsequent assertions not *asserting* that it is, even though what she asserts will be true only if P is true. In general, one does not assert what one presupposes.⁹

This picture of assertion and inquiry predicts that what is asserted using given linguistic resources will be context-dependent. To see this, suppose that some of the

worlds in the set are P worlds while others are not. Asserting that P will mark off the P worlds from the rest. Suppose further that some other fact, Q, also obtains in all and only the P worlds. The set of P worlds is the set of Q worlds, even though P and Q are different facts. In this case, asserting that Q will mark off both the Q worlds and the P worlds from the rest. The effect of asserting that Q is, in this case, the same as that of asserting that P, even if different linguistic resources would be used to make each assertion. In general, sentences that could be used to mark the same distinction among a set of worlds in one context may, in a different context, be used to mark different distinctions. As the inquiry proceeds, which distinctions a given sentence will mark may change. This picture of assertion predicts that this kind of context-dependence will be widespread.

This picture involves considerable idealization. It idealizes away from unrecognized disagreements and inconsistencies among participants in an inquiry, as well as from vagueness and indeterminacy in the distinctions given resources could be used to draw among the possible worlds. It also assumes an unrealistically high level of awareness among the participants about what they believe. In particular, it implies that all sides agree (and believe they agree) on which distinctions among the possibilities their words could be used to draw. I will return to this idealization later. In practice, inquiries and participants are not nearly so well behaved. One virtue of the picture, though, is that it helps us to predict and understand some of this “misbehavior”. This picture of assertion also abstracts away from the specific resources, linguistic and non-linguistic alike, that we have for making assertions. It focuses instead on the rational nature of assertion, viewing it as a kind of rational action to be explained in the ways rational action in general are best explained. Is this amount of idealization and abstraction acceptable?

An idealization is acceptable if it provides clearer insight into some elements of a phenomenon than could a picture that is more detailed and precise. I think that this picture of assertion, while highly idealized and abstract, helps to reveal subtle features of knowledge attribution, features that can help us to understand and perhaps even see the appeal of Contextualism. More specifically, though, this picture helps to reveal the structure of the skeptic’s challenge. When she raises a skeptical possibility during a discussion of whether S knows that P, the skeptic is proposing that a possibility be added to the set of worlds under discussion, a possibility that, she believes, is incompatible with S’s knowing that P. Understanding the skeptic’s challenge requires understanding just what the world she proposes adding to the set is like. As I will describe in Section 4, there are many kinds of challenges that a skeptic might make, and this picture can, I think, help us to distinguish them. To see all of this, though, it will be helpful to have a picture of knowledge in mind.

2 Knowledge and Belief

Contextualism is a thesis about knowledge attributions, not about knowledge itself, and should not be tied too closely to any specific account of knowledge. So I want to sketch a relatively minimalist account of knowledge, but one that indicates why knowing that P is more valuable than merely believing truly that P. The picture will

(or so I hope) be minimalist in the sense that it could be incorporated into otherwise very different accounts of knowledge and belief. But nothing I will say about Contextualism later will depend on this picture. It will serve purely as a helpful picture. I begin with a rough first pass, and then add details and qualifications.

What are we saying about Jones when we say that she knows that P? A natural idea is that we are saying in part that she has exercised a capacity to tell whether it is the case that P: a capacity to distinguish the actual world from worlds where it is not the case that P. By contrast, one can believe that P without having exercised a capacity to tell whether it is the case that P. Believing that P, even believing it truly, requires only a disposition to treat the world as if it were the case that P. Before discussing why this difference makes knowing that P more valuable than merely believing truly that P, let me note some details and qualifications.

I said that knowing that P requires a capacity to tell whether it is the case that P. Typically, there are different ways to tell. One can on occasion tell whether it is raining by listening to radio reports, or by asking a friend who is well-positioned, or by looking out the window, or by listening for the sounds of rain. Some ways to tell whether P may be more *extensive* than others. A maximally extensive capacity would distinguish a P world from any possible not-P world, whereas a minimally extensive one would distinguish a P world only from the not – P world or worlds closest to it. Perhaps listening for the sound of rain is less extensive, in this sense, than is listening to the weather report, since there are a lot of things that sound like rain, but few things that sound like a news report of rain. Perhaps there are no maximally extensive capacities for telling whether it is raining, or for telling anything else. Just how extensive a capacity one must have to count as knowing that P is an issue I will return to when I discuss the skeptic's challenge.

There are two reasons for thinking that the difference between knowing whether P and merely believing that P involves the exercise of a capacity and not merely its possession. First, having a capacity to tell whether P might be insufficient for believing that P, and so insufficient for knowing whether P. For it may be possible to have a capacity without having exercised it, and it may even be possible to have a capacity to tell whether P even though one is not in a position to exercise it. Perhaps Jones is able to tell by looking whether it is raining even though she has not checked and is nowhere near a window. If so, then Jones might not know whether it is raining, even though she can, in a sense, tell whether it is. That is one reason to think that knowledge requires having exercised some capacity, and not merely possessing it.

A second reason is that having a capacity to tell whether P might be necessary for believing that P. Perhaps one could not believe that it is raining without having some way to tell whether it is. If so, then a form of verificationism about the limits of belief would be true: one could believe that P only if one could know whether it is the case that P.¹⁰ These reasons indicate that the difference between knowing that P and merely believing (even truly) that P might involve having exercised a capacity to tell whether P, and not just having such a capacity. For, whatever else it might involve, believing that P does not require having exercised a capacity to tell whether P, whereas knowing that P does.

Does knowing that P require believing that P? It seems to me that if one has exercised a capacity to tell whether P, and it is the case that P, then one will be disposed to treat the world as if it were the case that P. If it is raining and Jones can tell whether it is raining, and has checked whether it is, then, it seems to me, Jones will be disposed to treat the world as if it were raining. Is this enough for Jones to believe that it is raining? It seems clear that believing that P requires being disposed to treat the world as if it were the case that P, but perhaps more is needed. I want to set this aside, since what I want to focus on is the fact that believing that P, whatever it might require, does not require having exercised a capacity to tell whether it is the case that P.

The value of knowing that P over believing truly that P comes clear when we contrast knowing that P with being a lucky guesser. If Jones knows that P, then she has exercised a capacity to tell whether it is the case that P. This means that, had it been the case that not-P, she would have found that out. If Jones knows that it is raining, then had it not been raining Jones would have known that instead. True belief is not like this. A lucky guesser or a wishful thinker might believe that P, even though she would have believed that P even had it not been the case, or might have not believed that P even if it were the case. A desperate farmer's wish for rain might lead to belief whether it is raining or not. Believing that P does not require that one would have not believed that P had it not been the case that P. Not even true belief requires that. The capacity involved in knowing that P tracks the facts in a way that even true belief does not. Plausibly, this is at least part of what makes knowing that P more valuable than merely believing truly that P.

According to this picture of knowledge, to attribute knowledge to someone is to say something about her capacities. It might be that we are saying more about her than just this. Perhaps we are also saying something about which capacity she has, or about how she has exercised that capacity, or about how extensive her capacity is, or about how she has acquired that capacity. If so, then a complete account of knowledge attributions would need to explain what more it is we are saying. But since my present aim is just to locate a fundamental difference between knowledge and true belief, I want to set all of this aside. However one fills out the details, what I am calling a fundamental difference between knowledge and belief will, I think, remain.

Assuming this picture of knowledge, we can now say more precisely what Contextualism claims. In fact, it makes two claims. The first claim concerns the context dependence of knowledge attributions in general: which capacity a phrase of the form "knows that P" can be used to identify may vary from one context to another. I think this kind of context-dependence is relatively uncontroversial. I will say why in the next section. Contextualism's second claim concerns what happens when the skeptic raises her challenge: a capacity once identifiable using a phrase of the form "knows that P" may, after the skeptic raises her challenge, no longer be so identifiable. This is what makes knowledge seem elusive when the skeptic joins in the discussion. The picture of assertion I sketched can help us to see what these kinds of context-dependence would be like.

3 Two Uncontroversial Kinds of Context Dependence

On the picture of assertion I sketched, the source of context-dependence is that what is presupposed can vary from one context to another. This gives rise to two kinds of context-dependence in knowledge attributions that should be uncontroversial. Seeing them will help us see the more controversial kind of kind of context-dependence that Contextualism claims is involved when the skeptic raises her challenge.

First, there will be context-dependence if the “that”-clause of a sentence of the form “S knows that P” contains a sentence that is itself context-dependent. In such a case, which capacity a use of it may identify will depend on context. This is clear when that sentence contains an indexical. Suppose Suzie and Sarah each make an assertion using “Jones knows that I am here.” Suzie and Sarah said different things, and what Suzie said might be true even if what Sarah said were false. Suzie ascribed a different capacity to Jones than did Sarah, even though they used the same words. And Jones might have the former capacity without having the latter one. This shows that which capacity a sentence of the form “S knows that P” can be used to identify can depend on context.

This example involves an indexical and indexicals are obviously context-dependent. But the picture of assertion I sketched entails that many, perhaps even all, sentences will be context-dependent—even those that do not contain indexicals. For, according to this picture, what distinction a given sentence can be used to mark among possibilities will depend on what is presupposed when it is used, and in different contexts a given sentence may mark different distinctions. If this is right, then since knowledge attributions involve a “that”-clause, such attributions will also be context-dependent. Which capacity a sentence of the form “S knows that P” could be used to identify, even when the “that”-clause contains no indexicals or demonstratives, would then depend on context.

Illustrating this would, I think, take us too far afield.¹¹ The indexical example is enough to make the general point. And, in any event, this kind of context-dependence should not be controversial. For it involves contextual variation in what a sentence says someone to know, not what it takes for someone to know that. It reveals nothing special about the nature or value of the grip involved in knowing a fact. And, most importantly, it does not entail the kind of elusiveness that Contextualists claim the skeptic’s challenge generates.

There is a second kind of uncontroversial context-dependence. The capacity identified by a knowledge attribution might, depending on the context, also be identified without reference to knowledge. I will illustrate this kind of context-dependence using three cases.

Case 1: suppose that all sides in an inquiry agree that P and that S believes that P, but disagree on whether S just got lucky. What is in question is whether S has exercised a capacity to tell whether it is the case that P. We can represent this state of the inquiry using a set of P-worlds in all of which S believes that P. Some of the worlds are ones where S got lucky in believing that P. Call these the “L” worlds, to suggest some degree of luck. Other worlds are ones where S did exercise a capacity

to tell whether P. Call these the “K” worlds, to suggest that they are worlds where S (assuming that nothing else is needed for knowledge) knows that P. Suppose, further, that in all and only the K worlds, S read the New York Times that morning.

The worlds in the set differ in one linguistic respect. In the K worlds, an assertion using “S knows that P” would be asserting something true, whereas in the L worlds an assertion using it would be asserting something false. But the worlds are linguistically similar in an important respect: in all of them an assertion using “S knows that P” would mark a distinction between the K worlds and the L worlds. This reflects the fact that while it is an open question whether an assertion using that sentence would say something true, it is not an open question what such an assertion would say. The inquirers agree on how to state what the facts might be, even while they disagree on what the facts are.

Any assertion would, in this case, presuppose both that P and that S believes that P. The content of an assertion using “S knows that P” would (let us assume) be that the actual world is one of the K worlds. Notice that this would also be the content of an assertion using “S read the NY Times today”. In this case, such assertions would simply be different ways of marking off the K worlds from the rest, different ways of determining the very same truth conditions.

Case 2: suppose that all sides agree that P but wonder whether S believes that P. We can represent this state of agreement by a set of P worlds that includes some in which S believes that P and some in which S does not. Suppose also that some of the worlds where S believes that P are K worlds while the rest are L worlds. This reflects the fact that the participants also wonder whether, even if S does believe that P, she also knows that P. Finally, suppose that in none of the worlds did S read *The New York Times* that day.

An assertion using “S knows that P” would, in this case, mark a very different distinction than in case 1. For in case 2 it is not being presupposed that S believes that P. In asserting that the actual world is a K world one would also be asserting that it is a world where S believes that P. But notice that, in this case, this effect would not be achieved using “S read *The New York Times*.” An assertion using this sentence would not mark off all and only the K worlds from the rest. Indeed, such an assertion might not even mark off all of the worlds where S believes that P from the rest, since it might be true in some of them that S would not believe that P if not P. However, the very same distinction marked using “S knows that P” could also be marked using “S has exercised a capacity to tell whether P.”

Case 3: suppose that it is an open question whether P is the case, whether S believes that P, and even whether S can tell whether it is. To represent this state of inquiry, we need a set of worlds where some are P worlds and others are not; where some but not all of the P worlds are ones where S believes that P; and where some but not all of these are K worlds. An assertion using “S knows that P” would, in this case, be saying even more than was said in case 2. For in this case, it is not even presupposed that P. Intuitively, the content of this assertion is much richer than the assertions in the first two cases. And notice that an assertion using “S has exercised a capacity to tell whether P” would not necessarily mark off the K-worlds from the rest, since this might be true even in some of the not-P worlds in the set.

I suspect that case 1 exemplifies the typical case when a knowledge attribution is made (i.e., when a knowledge attributing sentence like “S knows that P” is used to make an assertion). The use of “S knows that P” in case 3 strikes me as very odd. For, in effect, it is in that case a way to assert that P. Too much is being ruled out in one assertion. A more reasonable strategy, it seems to me, would be first to assert that P is the case and then that S believes it, and only then, as in case 1, to assert that the actual world is a K-world. Case 2 strikes me as not quite as odd as case 3, but still less natural than case 1. I do not mean to suggest that cases 2 and 3 are somehow defective or impossible, only that they are probably not examples of typical knowledge attributions.

These cases illustrate that what is said in a knowledge attribution might also be said, depending on context, using other words. One way to describe this phenomenon might be to say that the content of a knowledge attribution varies from context to context. But notice that in all of these cases, the effect of the knowledge attribution is to mark off the K worlds from the rest. What varies is only the membership in the set of excluded worlds, not the membership in the K set. What is more, in all of these cases, the set of K worlds can be identified using a knowledge attributing sentence, even though it can also be identified in other ways. This kind of context-dependence is, I think, relatively uncontroversial.

In any event, it is not the kind of context-dependence that makes Contextualism controversial. For everyone can agree that which capacity a knowledge attributing sentence can be used to identify can vary from one context to another. That is, which set of worlds a knowledge attributing sentence can be used to identify can vary from context to another. All sides can also agree that, depending on context, a given capacity may be identified without explicit reference to knowledge. That is, the set of worlds a knowledge attributing sentence can be used to identify may, depending on context, also be identifiable in other ways. None of this, I think, should be controversial. Contextualism is controversial, because it claims that a set of worlds, once identifiable in a knowledge attribution might, after a change in context, not be so identifiable. This is the alleged elusiveness of knowledge. According to Contextualism, this is what happens when the skeptic raises her fantastic doubts.

4 Contextualism and the Skeptic

To describe this kind of context-dependence, I want to consider a challenge to a knowledge attribution made in case 1, from above. One reason for focusing on this case is that, as I indicated, I think it is the most representative kind of knowledge attribution. But another reason is simply a concern to keep the presentation as simple as possible. So, to recall that case, an assertion is made using “S knows that P” in a context where all sides agree that P and that S believes that P. Let us suppose that the assertion is accepted by all sides, so that all the worlds in the set after the assertion is accepted are K-worlds. But then the skeptic raises her challenge: “But what if Q?” The force of the skeptic’s challenge is that some possibility, Q, that she believes

to be incompatible with S's knowing that P, be added to the set. The nature of the challenge depends on what that possibility is. We need to consider three cases.

Truth Skepticism: In one kind of case, the skeptic raises a possibility that is incompatible with P. This skeptic is questioning whether S knows that P because she is questioning whether it is the case that P. Perhaps the skeptic has some evidence that it is not the case that P, and wishes to consider it. In that extreme case, she is proposing that all the P worlds be replaced by Q worlds. Or, more moderately, perhaps she simply proposes reopening the question as to P. In that case, her proposal is that Q worlds be added to the set.

In practice, I think, the onus would be on the skeptic, whether moderate or extreme, to justify adding this world to the set. In practice, I think, merely mentioning a possibility is not sufficient for it to be included. But let us suppose the skeptic has the more moderate challenge in mind and that it is accepted. (This requires only that the participants accept for the purposes of the inquiry that it is an open question whether P, and not that they give up their belief that P.)

What is this Q world like? Given that it is not a P world, it is a world where S has not exercised a capacity to tell whether P. (This skeptic might agree, though, that S has such a capacity.) There is an important linguistic similarity between Q worlds and K worlds. In both kinds of worlds, an assertion using "S knows that P" would say something that is false in the Q worlds and true in the K worlds. This means that, as in case 3, the K worlds can still be identified using "S knows that P". This would be (to my ear) a bit odd, since it would be asserting a great deal more than was originally asserted using those words in case 1. But the important point, for present purposes, is that the K worlds identified using "S knows that P" could, even after the truth-skeptic's doubts are accepted, still be identified that way.

Capacity-Skepticism: A second kind of skeptic agrees that P, and that S believes that P, but questions whether S has the capacity to tell whether P. She agrees that if the actual world were a K world then S would know that P. That is, she agrees that in those worlds S has (exercised) the capacity required for knowing whether P. But she doubts whether S in fact has (exercised) such a capacity. She wonders whether S really just got lucky. As before, this challenge might be either extreme or moderate. A moderate capacity skeptic is proposing that an L world be added to the set, one where although S believes truly that P, S did not exercise a capacity to tell whether P. I think that the onus would be on the skeptic to justify this addition. But suppose that it is a moderate challenge and that it is accepted.

As before, the skeptic's challenge has produced a change in the set of worlds under discussion. The set is now like that in case 1, above. And, as in case 1, the K worlds can still be identified using the sentence "S knows that P." For, as with the truth-skeptic, in every world in the set an assertion using this sentence would say something that is true in all the K worlds and false in the rest.¹² So the capacity identified using "S knows that P" before the skeptic raised her challenge could still be so identified even after the skeptic's challenge is accepted.

Standards-Skepticism: A third kind of skeptic agrees that the actual world is a K world, but denies that it is a world where S knows that P. It remains, as she would put it, an open question whether S knows that P, though she is happy to agree (at

least for the sake of the discussion) that the actual world is a K world. This skeptic is raising a question about what it takes for S to know whether P, about the standards that must be met for this. Her contention might be that S's capacity to tell whether P is not extensive enough for knowledge: there are possibilities where P is not the case that S's capacity could not distinguish from the actual world and which it must, she contends, for S to know that P. She is not claiming that one of those worlds is the actual world. Her claim is that knowing that P requires having a capacity that would distinguish such possibilities from the actual world, and that S has not exercised such a capacity.

This kind of challenge is hard to classify because while the skeptic is proposing that some world be added to the set, it is not immediately clear what this world is supposed to be like. It is not one where P is not the case, nor is it one where S does not believe that P, nor even is it one where S lacks a capacity to tell whether P. For this skeptic may agree that, as she would put it, the actual world is such a world. Her claim, as she would put it, is that in none of these K worlds does S know that P.¹³ Let us call the kind of world she is proposing be added to the set an "NK" world (to suggest that it is a newly introduced). How do NK worlds differ from K worlds?

One clear difference is linguistic. While everyone, including the skeptic, agrees that the actual world is a K world, they disagree about whether the sentence "S knows that P" is true. So, at the very least, the skeptic is proposing to add a world that is just like the K worlds, in that it is one where S truly believes that P and has a capacity to tell whether P, but where the sentence "S knows that P" is false (or at least not true). But there is a more important linguistic difference between K and NK worlds. An assertion using "S knows that P" would mark a different distinction in K worlds than in NK ones. An assertion using it in a K world would assert something true in all the worlds, even the NK worlds; whereas an assertion of it in an NK world would assert something false in all of the worlds, even the NK worlds. In other words, this skeptic's challenge is that it is an open question what an assertion using "S knows that P" would say, and not just whether it would say something true. The standards-skeptic's proposal is that it is (at least) an open question what the truth conditions of an assertion using "S knows that P" would be. This marks off this kind of skeptical challenge from both the truth skeptic and the capacity skeptic. For their challenges concerned, not what the content of an assertion using "S knows that P" is, but only the truth of that content.

Given all of this, it might be tempting to say that the standards-skeptic's challenge concerns not the facts but rather the linguistic or conceptual resources for thinking about or describing the facts. It is, one might think, solely a challenge about meaning and not about truth: what she questions is not what the facts are, but rather which words can be used to state those facts. However tempting this description of the skeptic's challenge might be it is best resisted. For a disagreement about the linguistic facts is still a disagreement about the facts. The fact that a phrase like "knows that P" means what it does derives from facts about speakers' intentions, conventions, and past usage. So K worlds differ from NK worlds not just in what an assertion using that sentence would say but also with respect to these other kinds of facts. And it seems right that the standards-skeptic's claim is, in part anyway, that

our past usage and intentions concerning “knows that P” are such that it rarely if ever is used to say something true. This is why it is an unsatisfactory response to such a skeptic merely to say: Well, maybe that’s what you mean by “knowledge”, but it is not what I mean. For it is, or may be, part of her claim that she is the one being faithful to ordinary language.

There is a second reason to resist this tempting response. The standards-skeptic’s challenge may also concern what kind of relation to the facts is most valuable. According to an extreme standards skeptic, perhaps, only the strongest kind of relation—perhaps one impervious to demonic trickery—is worthwhile. Anything less is, in her eyes, no better than a lucky guess. It is harder to know how to resolve this kind of disagreement. Again, though, it is no good to respond simply by saying: Well, maybe that’s what you want, but it is not what I want. For part of her claim is that we in fact do, or perhaps should, want the kind of cognitive relation she finds most valuable.

In any event, once the standards-skeptic’s challenge has been raised, there are three possible responses. The other participants might decide to make no change to the set of worlds, perhaps because they are unconvinced by the reasons the skeptic offers for changing it. Or the participants might decide that it is an open question whether the skeptic is right. Finally, they might agree that she is right in her challenge. Some Contextualists claim that any challenge to the standards for knowledge invariably succeeds. I doubt this. There is no obvious reason why the skeptic’s challenge cannot simply be resisted. In any event, since the skeptic’s challenge is accepted in only the last two cases, I will focus on these.

First, suppose the participants agree to re-open the question by adding an NK world to the set. And suppose that one of the participants wants to claim that S has the capacity earlier identified using “S knows that P.” That is, she wants to mark off the K worlds from the NK ones. Making an assertion using “S knows that P” will not achieve this. For what would be said using that sentence depends on whether the actual world is a K world, and it is an open question whether it is. And this means that it would not be clear what distinction among the worlds the speaker would intend to mark. What is clear to all in the inquiry is that if the actual world is a K world, then S has a certain capacity and an assertion using “S knows that P” would say something true, whereas if the actual world is an NK world, then S has that very same capacity and an assertion using that sentence would say something false. But since it is an open question which world in the set is the actual world, it would be unclear which kind of world one who uttered that sentence would be saying is the actual world. For this reason, the K worlds cannot be identified in that way, even though they were identified in that way before the skeptic’s challenge. After the skeptic’s challenge is accepted, S’s capacity eludes identification by “S knows that P.”

But notice what is agreed to by all sides: in all the worlds, S has a capacity to tell whether P. All sides can agree that S’s grip on the world, identified at one time using the words “S knows that P”, remains just as it was in nature after the skeptic has raised her challenge. All sides can also agree that it remains more valuable than mere true belief, even if they disagree about its ultimate worth. Finally, all sides

also agree that after the skeptic's challenge that grip can no longer be identified by "S knows that P." All sides, including the skeptic, agree on the nature of S's grip on the facts and on its value over mere true belief. But they differ on whether it is properly identified by "knowledge that P." The grip remains unchanged in every respect but name.

Suppose, next, that the participants decide that the skeptic is right. That is, they agree with the standards-skeptic that the actual world is an NK world. To reflect this acceptance, an assertion using "S knows that P" would say the same thing in all the worlds in the set. And in all the worlds it would say something that is false in all those worlds. This reflects the fact that the participants have accepted the skeptic's claim that S's capacities, however impressive, do not measure up to the standards of knowledge. But notice that in all of the worlds it remains the case that S has that capacity. All sides, in other words, agree that S has a capacity that it is more valuable than mere true belief. However, they also agree that it is a capacity that cannot be identified using "S knows that P". As before, the capacity remains the same in nature and value, but not in name.

I have been discussing an extreme version of a standards-skeptic's challenge to a knowledge claim. She claims that the standards for knowledge are so high that, as she would put it, in none of the K worlds does S know that P. But more moderate challenges are possible. A more moderate challenger might claim, instead, that the standards for knowledge are met in only a subset of the K worlds, and question whether that subset includes the actual world. Such a challenge might again rest on a view about just how extensive a capacity must be for knowledge. She might allow that in some of the worlds S has a capacity that is extensive enough while in others she has a different, and insufficiently extensive, capacity. And she may wonder whether the actual world is one of latter. But in the case of all such challenges, whether by an extreme standards skeptic or a moderate one, all sides agree about S's capacities. The dispute concerns, not what S is capable of doing, but only whether her capacities meet the standards for knowledge. And in all such cases, once the skeptic's challenge is accepted, S's capacity cannot be identified using "S knows that P."

I have discussed Contextualism without discussing the specific mechanisms involved. As I noted at the outset, some Contextualists posit underlying syntactic or semantic complexity in the word "knows", while others explain the context-dependence in terms of pragmatic mechanisms like those involved in conversational implicature. Just what the mechanisms in fact are is an important empirical issue. It may turn out that several different kinds of mechanism are at work. But the picture of assertion I sketched suggests that this kind of context-dependence will occur regardless of the specific resources used for making knowledge attributions. For it suggests that, in general, as conversations proceed the distinctions that given linguistic resources can be used to draw among possible worlds will change too. There is no reason to expect that knowledge attributions would be an exception. In my view, Contextualism should be tied, not to a specific account of the mechanisms involved, but rather to a more abstract account of the rational nature of assertion.

5 Some Examples

My discussion of Contextualism has been rather abstract. It would be helpful to consider a concrete example. Two leading exponents of Contextualism, Stewart Cohen and Keith DeRose, offer now familiar examples which they claim illustrate how different contexts can involve different standards of knowledge. But it is not clear to me that their examples succeed. Since the examples have played an important role in the literature, I will consider them at some length.¹⁴ I will then offer what may be a better example.

Here is an example Cohen uses to motivate his version of Contextualism.

Mary and John are at the L.A. airport contemplating taking a certain flight to New York. They want to know whether the flight has a layover in Chicago. They overhear someone ask a passenger Smith if he knows whether the flight stops in Chicago. Smith looks at the flight itinerary he got from the travel agent and responds, "Yes I know—it does stop in Chicago." It turns out that Mary and John have a very important business contact to make at the Chicago airport. Mary says, "How reliable is that itinerary? It could contain a misprint. They could have changed the schedule at the last minute." Mary and John agree that Smith doesn't really know that the plan will stop in Chicago. They decide to check with the airline agent. (Cohen, 1999, 58)

In commenting on this example, Cohen claims that Smith and Mary have different standards as to what counts as knowing whether the plane will land in Chicago. Smith's standard, Cohen claims, is weaker than Mary's. Cohen considers whether we should say that one or the other or neither standard is correct. He argues that saying any of these things would be unacceptable. Instead, he suggests, the more reasonable view is that

[N]either standard is simply correct or simply incorrect. Rather, context determines which standard is correct. Since the standards for knowledge ascriptions can vary across context, each claim, Smith's as well as Mary and John's, can be correct in the context in which it was made. When Smith says "I know. . .", what he says is true given the weaker standard operating in that context. When Mary and John say "Smith does not know. . .", what they say is true given the stricter standard operating in their context. *And there is no context independent correct standard.* (Cohen, 1999, 59; italics in original)

Plainly, Cohen thinks that this example involves a disagreement about standards for knowledge.

But it is not clear from Cohen's description of the example that this is what Mary and Smith really disagree about. Mary's challenge to Smith's claim to know is complex, but also under-described. Clearly, she questions whether he has exercised an appropriate capacity to tell whether the plane will stop in Chicago. But her reason for this, it seems, has to do with questions she has about whether the plane will in fact stop in Chicago. It is not that she agrees that the plane will land in Chicago, but doubts whether Smith knows this. She doubts whether the plane will land in Chicago. She and Smith disagree about this: he believes it will, she is not convinced. Her challenge to his claim to know is thus, at least in part, that of a truth-skeptic.

Since they disagree about whether he is right, Mary and Smith must also disagree about whether he has exercised an appropriate capacity. But there are different ways

they might disagree on this. They might disagree about which capacity he used. Perhaps Mary believes he used the inappropriate capacity C while Smith believes he used the appropriate capacity C*. This disagreement would be like that of the capacity-skeptic. Or perhaps Smith and Mary agree about which capacity he used, but disagree about whether it is appropriate. This would be like the case of the standards-skeptic. Unfortunately, there is not enough detail in Cohen's example to decide just where they are disagreeing here.

It is not even clear from the example just which capacity Smith has exercised. We are told that Smith looks at the itinerary he got from the travel agent. But that itinerary might be minutes old, or it might be 2 weeks old. And Smith might be relying on other evidence he has. Nor is it clear whether Smith and Mary agree on which capacity he has exercised. Given that she hardly knows him, Mary probably does not really know what capacity Smith has exercised. In any event, there is not enough detail in the story to make it clear that the example does involve a genuine disagreement about standards. So it is not clear that this example really does support Contextualism.

Here is an example Keith DeRose uses in an exposition and defense of Contextualism. It involves contrasting cases.

Bank Case A: My wife and I are driving home on a Friday afternoon. We plan to stop at the bank on the way home to deposit our paychecks. But as we drive past the bank, we notice that the lines are very long, as they often are on Friday afternoons. Although we generally like to deposit our paychecks as soon as possible, it is not especially important in this case that they be deposited right away, so I suggest that we drive straight home and deposit our paychecks on Saturday morning. My wife says, "Maybe the bank won't be open tomorrow. Lots of banks are closed on Saturdays." I reply, "No, I know it'll be open. I was just there two weeks ago on Saturday. It's open until noon."

Bank Case B: My wife and I drive past the bank on a Friday afternoon, as in Case A, and notice the long lines. I again suggest that we deposit our paychecks on Saturday morning, explaining that I was at the bank on Saturday morning only two weeks ago and discovered that it was open until noon. But in this case, we have just written a very large and important check. If our paychecks are not deposited into our checking account before Monday morning, the important check we wrote will bounce, leaving us in a *very* bad situation. And, of course, the bank is not open on Sunday. My wife reminds me of these facts. She then says, "Banks do change their hours. Do you know the bank will be open tomorrow?" Remaining as confident as I was before that the bank will be open then, still, I reply, "Well, no. I'd better go in and make sure." (DeRose, 1992)

DeRose comments that the "contexts of my utterance in the two cases make it easier for a knowledge attribution to be true in Case A than in Case B." (DeRose, 1992, 110.) He suggests that the relevant contextual difference concerns the importance of being right. In Case A, less hangs on whether he is right about the bank's hours than in Case B. Consequently, he claims, the standards for knowing whether the bank will be open are lower in Case A than in Case B.

Again, though, it is not clear that the difference in the cases is a difference in standards. The wife in Case B seems to think that is an open question whether the bank will be open on that Saturday. It is not that she agrees with her husband that it will be open but questions whether he can tell that it will. She worries that it might be closed. So, at least part of her reason for challenging his claim to know

is that she has doubts about whether he is right. Her challenge is, to this extent, that of a truth-skeptic. Of course, this means that she also questions whether her husband has exercised a capacity to tell whether P. So she is also raising a capacity-skeptic's challenge. And perhaps she even questions whether he has exercised a sufficiently extensive capacity. So perhaps she is even raising a standards challenge to her husband's claim to know. There is, I think, not enough detail in DeRose's example to decide this.

More specifically, it is not clear from the way DeRose describes the examples why the wife does not have the same concern in Case A as in Case B. In Case B, her concern is that the bank might have changed its hours in the last two weeks and will be closed on that Saturday. The fact that it does not matter as much in Case A whether the actual world is such a world is perhaps reason not to care as much whether her husband is right. But it is not reason to think that he is right. The reason she has in Case B for questioning his claim to know ought, it seems to me, be a reason for her to challenge his claim in Case A. For that challenge rests, not on the standards for knowledge, but on whether the husband's belief is true. And whether it is true is just as relevant to whether he knows in Case A as it is in Case B. So it is not clear to me that DeRose's example is the right kind to support Contextualism.

What is needed is an example where all sides agree that the subject's belief is true. Let me try to construct one.

Ernie is sitting in a Buffalo garden watching a blackbird. He, in turn, is being watched by Suzie and Sarah, from one corner, and by Frank and Sam from another. Suzie and Sarah are good friends of Ernie's. Frank and Sam are ornithologists who have been studying a recently discovered species of pigeon found only in the tropics that has the same outward appearance as Buffalo area blackbirds. Here are their conversations.

Conversation 1:

Suzie: See that blackbird out there?

Sarah: Oh, yes. Pretty.

Suzie: I wonder what kind of bird Ernie thinks it is.

Sarah: He knows it is a blackbird.

Suzie: Well, I am sure he *thinks* it is a blackbird. But can he tell the difference between a blackbird and a raven? They are very similar.

Sarah: Yes, he knows it is a blackbird. He's an amateur ornithologist.

Suzie: Oh, really, I did not know that.

Conversation 2:

Frank: Look, there's Ernie studying that blackbird.

Sam: Oh, yeah. Isn't it amazing how much it resembles those pigeons.

Frank: Yes, it really is. Can Ernie tell the difference?

Sam: Yes, he knows that's a blackbird. Remember, we told him about the pigeons, and how to identify them. And look how he is studying the blackbird. He knows what to look for, and he always wants to be right about things like that.

Frank: No, you are thinking of Henry. We taught Henry how to distinguish pigeons and blackbirds. I am not so sure Ernie can tell that it is a blackbird. I mean, I am sure he thinks that it is. He's a pretty good amateur. He doesn't think it is a raven or anything.

Sam: Oh, yeah, you are right. I guess Ernie doesn't know. We should show him.

Here is my commentary on these conversations. Sarah and Sam each used the words "He knows that it is a blackbird." to say something about Ernie. Sarah said something true while Sam did not. They each ascribed a capacity to him. Ernie has the one Sarah said he has, but lacks the one Sam said he has. The capacities are both capacities to tell whether something is a blackbird. But they differ in their extensiveness. The capacity Sam ascribed to Ernie would distinguish blackbirds from those tropical pigeon look-alikes. The one Sarah ascribed would not. The truth conditions of utterances of "He knows that it is a blackbird" are thus different in the two conversations.

But perhaps this case involves a difference in what is known, and not a difference in what it takes to know that. Earlier I said that different uses of "S knows that I am here" involve a difference in the fact known, and not necessarily a difference in what it takes to know something. Perhaps the fact Sarah said that Ernie knows is not the one Sam said he knows. I find this not very plausible. For it seems natural to think that the belief Suzie ascribes to Henry is the very belief Frank ascribes to him. They agree, it seems, that Sam believes that it is a blackbird. They agree, it seems right to say, about which fact would make that belief true. Where they disagree, it seems to me, is over what relation to that fact "knows that it is a blackbird" identifies.

But suppose the case does involve a difference in the fact known. This would still imply that the truth conditions of the two utterances of "He knows that it is a blackbird" differ. It would, though, trace this variation to one in the contribution of "it is a blackbird". On this view, uses of this sentence state different facts in different contexts, even when the same bird is being referred to. Perhaps the view is that uses of "is a blackbird" can identify different properties in different contexts.¹⁵ What is more, if the fact Sarah identified is not the one Sam identified, then the capacity Sarah ascribed to Ernie is not the one Sam ascribed to him. So this view even grants that which capacity a phrase of the form "knows that P" can be used to identify can depend on context. It seems to me that granting all of this is as good as agreeing with the Contextualist.

6 Contextualism and Objectivity

The picture of assertion I have been relying on suggests that context-dependence will be common and predictable. But it does not entail that it will be necessary. The context-dependence derives, on this picture, from the context-dependence of what is presupposed. Because what is presupposed can differ from one context to another, and because it can change as inquiry proceeds, given linguistic resources

could be used to mark different distinctions among possibilities in different contexts and as inquiry proceeds.

But couldn't one eliminate this source of context-dependence? Couldn't one make a knowledge attribution in a context where nothing was presupposed, where all possibilities were left open, or at any rate where all those we can recognize or distinguish among were left open? And wouldn't such a context in fact provide a more objective standpoint or perspective on the facts, one less tied to our contingent interests and purposes. And wouldn't it allow us better to identify just which distinctions our concepts, and in particular our concept of knowledge, mark among the possibilities.¹⁶ And wouldn't this be the proper perspective from which to take on the skeptic?

I want to conclude this paper by discussing these questions, though what I will say will be pretty impressionistic. I will mostly pose questions about the idea of a presupposition-free context. I think it is not so clear what such a context would be like or how we could come to occupy one. But I will begin by arguing that, in an important sense, a presupposition-free context would not provide a more objective perspective on the facts.

The context-dependence I have been discussing derives from the context-dependence of what is presupposed in an inquiry or a conversation. It is true that what is presupposed can depend on the interests and purposes of those engaged in the inquiry or conversation. Changing those interests and purposes can affect what is presupposed by changing which questions are considered settled and which are considered still open. How given resources, whether linguistic or not, can be used to distinguish among possible worlds as we try to locate the actual world thus depends on our interests and purposes. But whether we succeed in locating the actual world does not. At least, nothing in the picture of assertion and inquiry I have sketched entails that truth is somehow dependent on our interests and purposes. In this sense, even a context where much is presupposed can permit a kind of objectivity worth having.

Still, a presupposition-free context might seem to promise a different kind of objectivity. In such a context, it might seem, our words and concepts would draw distinctions among the ways things might be free of interference from our changeable interests and purposes. We could, from such a context, identify just which distinctions our words and concepts mark, their "literal" meaning. This kind of "conceptual" objectivity might seem just as desirable as an objectivity about truth. For, it might seem, a debate with the skeptic should concern, not which distinction we might, on some occasion, use our word "knowledge" to draw, but rather which distinction that word itself, or that concept, draws. Only in a presupposition-free context, it might seem, could such a debate take place. So we should consider whether the idea of our occupying such a context is a coherent one.

For reasons we have already noted, no context for inquiry or assertion could be wholly free of presupposition. There have to be some similarities among the worlds in the set. More specifically, there have to be linguistic similarities. A speech act can succeed only if it is determinate what its force and content would be. What the content of an assertion would be cannot depend on which of the worlds in the set

is the actual world. For, in that case, we could not know what we were asserting until we knew which world we were in. To avoid this, an act using given resources would have to have the same force and content in every world in the set. Likewise, we would have to have the very same concepts in every world in the set. For if which concepts we had depended on which world in the set were the actual world, then we could not know which way we were conceiving of things until we knew how things were.¹⁷ So a context for assertion or judgment cannot be wholly free of presupposition.

But, at least in the abstract, this amount of presupposition might not seem too limiting. It would require only settling all questions of meaning while leaving open all questions of fact. More specifically, it would require only that each sentence mark, in each of the worlds in the set, the same distinction among the worlds in that set. With questions about how to mark distinctions among possible worlds out of the way, we could then focus on locating the actual world. In the dispute with the skeptic, once we settle questions about what our words say, or about how our concepts distinguish among the ways things might be, we could focus on determining what, if anything, we know. Couldn't we come to occupy such a context?

I have already noted one obstacle. The fact that our words mean what they do depends on facts about past usage and speaker intentions. So agreeing on what our words mean would require agreeing on such facts too. In other words, the worlds in the set must also agree in matters concerning past usage and speaker intentions. Agreeing with the skeptic about the meaning of "knows" would thus require agreeing with her about such facts. As with other historical questions, disputes about these facts might be practically impossible to settle. In that case, the only step forward would be for one side to concede in the other's usage. This would, of course, be to agree on a new language, and risks the charge that, instead of settling the dispute, it has merely changed the topic.

There is an even more serious difficulty. Agreement on what our words mean is not independent of agreement on what the non-linguistic facts might be. We cannot determine how our words distinguish among the possibilities without determining what possibilities there are. A disagreement about what is possible will lead to a disagreement about whether two sentences say the same thing. So we cannot agree on the meanings of our words without agreeing on how things might be. So agreeing with the skeptic on the meaning of "knows" requires agreeing with her about what is possible. But part of the dispute concerns what is possible. In particular, it concerns whether it is possible for us to have the capacities we all think we have and yet for us to fail to know anything.

The point is not that there is a merely practical bar to reaching these agreements. It is not just that reaching an agreement on our language requires a prior agreement on a different language, with a threatening regress. For the very same difficulty is present when one tries to identify one's own concepts. To figure out what I mean or what I will decide to mean by "knowledge", I have to decide how it distinguishes (or is to distinguish) among the possibilities, and to do this I need to figure out what the possibilities are. But these include possibilities concerning knowledge. And how could I figure out these without knowing what I mean by "knowledge"? The point

is that knowing what our words mean, and knowing what concepts we have to think with, is not independent of knowing what is possible.¹⁸

All of this suggests that attaining a context that would provide what I called conceptual objectivity is not a realistic goal. What follows from this about the nature of our concepts and, in particular, about our concept of knowledge? I think there are two broad responses. One might respond that, even though we cannot occupy such a context, our concepts are nonetheless suited for one. According to this response, we may not be able to understand just which distinctions our concepts draw among the possibilities, because we are incapable of occupying the kind of context such understanding requires. It is not that we cannot ourselves mark those distinctions: it is rather that we cannot know them to be the ones that our concepts and words mark. For, if such a context is unattainable, then we are never in a position to separate the distinctions the concepts themselves draw from those they draw *given our presuppositions*. On this response, our presuppositions about how things are prevent us from seeing how our concepts carve up the space of possibilities.

I am inclined to view this response as a kind of Platonism about our concepts, in that it views the question of what distinctions we use our concepts to draw as independent of the question what distinctions those concepts themselves draw. In fact, it sees the concepts, or better, the fact that our words express these concepts, as more fundamental than the fact that we use these words to mark distinctions. The fact that we cannot avoid presupposition is seen as an epistemic barrier, blocking us from a complete understanding of our conception of the world.

One who finds this Platonism appealing will likely not be impressed by a Contextualist response to skepticism. For it will seem to her that the skeptical debate concerns our concept of knowledge, and not just those distinctions that we, with our shifting presuppositions, might use that concept to draw. The debate, as she sees it, concerns whether we have the kind of cognitive grip that our concept of knowledge is a concept of. From her point of view, it is of only limited relevance whether we have the kind of cognitive grip that we might from one context to another use that concept to identify.

One might, though, draw a different lesson from the fact, if it is one, that we are unable to occupy a presupposition-free context. The proper lesson, one might think, is that because we cannot occupy such a context, our concepts are not suited for one. It is not that we are barred from understanding the distinctions our concepts mark among possibilities, but rather that our concepts mark no distinctions on their own. The lesson, on this response, is that our concepts are suited for the uses we make of them, for the distinctions we draw with them. And if these uses are invariably against a background of presuppositions, then our concepts are suited for such use and would not serve without such a background. It is a mistake, on this view, to think that our concepts have a semantic life of their own, a "literal" meaning, independently of the uses we make of them.

I think of this response as a kind of Pragmatism about our concepts, in that it views the use we make of our words and concepts as being more fundamental than any semantic properties our words or concepts may have. We can, from a given perspective, draw distinctions among possibilities, and we can use our concepts

and words to do this. In this sense, we can conceptualize the world. But different perspectives may yield different conceptualizations. And we should not think of such conceptualizations as second-rate versions of a genuinely objective conception, one that is somehow just out of reach.

One who finds this Pragmatism appealing will likely also find the Contextualist response to skepticism appealing. For, from the Pragmatist's point of view, debate about whether some cognitive grip is knowledge is debate about how we might decide to use our concept of knowledge to mark distinctions, and not about which distinctions that concept itself marks. Though she may be inclined to view such a debate as purely linguistic, I have already argued that this is a mistake. But, according to the Pragmatist, the idea that the debate with the skeptic could be properly understood, not to mention resolved, from a presupposition-free context is a mistake. We cannot occupy such a context, and even if we could, our concepts, including our concept of knowledge, would be of no use in it.

This Pragmatism echoes a theme of Quine's, that no neat distinction can be drawn between the conventional and the factual components in any truth.¹⁹ Quine characterized the idea that our concepts, by themselves, determine what is possible and what is not, as the myth of the museum. I suspect that the lure of that myth feeds the hope that we could, or perhaps even do, occupy a presupposition-free context. But that myth, according to Quine, distorts the nature of our conceptual relations to the world. One of the virtues of Contextualism, at least as I have tried to describe it here, is that it can help us to see just what this alleged distortion is supposed to be.²⁰

Notes

1. See, e.g., (DeRose, 1992) and (Cohen, 1987) for examples the first kind, and (Lewis, 1996) for one of the second kind. A different brand of Contextualism, associated with (Williams, 1991) and (Travis, 1989), is more difficult to categorize. A very helpful introductory discussion can be found in (DeRose, 1999).
2. (Stanley, 2004) develops a form of Contextualism of the first kind, though he does not endorse it. A version of the second kind is defended in (Lewis, 1996). (Schiffer, 1996) describes several accounts of the specific linguistic mechanisms involved, though he does not endorse Contextualism.
3. (Lewis, 1996).
4. This response is suggested by David Lewis's discussion at the end of (Lewis, 1996), and also by (DeRose, 2004).
5. A common criticism of the Contextualist account of what happens when the skeptic raises her doubts is that it makes defeating the skeptic too easy. Can it really be, the criticism goes, that the debate with the skeptic is just over how to use the word "knows"? This criticism is expressed in (Dretske, 1991), (Schiffer, 1996) and (Feldman, 2001).
6. This picture is clear in (Stalnaker, 1984) and the papers collected in (Stalnaker, 1999). It is also developed in (Lewis, 1983). The idea that assertion is best viewed as rational action traces back to the work of Paul Grice, see the papers collected in (Grice, 1989).
7. In Section 6, I consider whether attaining an objective conception of ourselves would require occupying a context where the set of possibilities under discussion includes all possibilities.
8. See (Lewis, 1983) and (Grice, 1989).
9. A speaker need not believe what she presupposes during a communal inquiry. She may assume something she in fact doubts, or treat as open a question she believes to be settled. In

- an ideal inquiry, the participants will be aware of the divergence between what is presupposed by all (the common ground), and what each participant in fact believes. For more on this, see (Stalnaker, 2002).
10. Thanks to a reviewer for helping me to put this point properly.
 11. There are examples in (Travis, 1989).
 12. This way of putting it may suggest that the words, all by themselves, mark distinctions among the possibilities. What really matters, though, is that the participants agree, even if only for the sake of the discussion, about which distinction an assertion using those words would mark. Successful inquiry does not require that the participants are right about which distinctions their words would mark. It does not even require that there be a fact about this. All it requires is that the participants (largely) agree on how their words will be used to mark distinctions. Their words are just conventional tools for drawing distinctions. (Even solitary inquiry requires only that the inquirer have beliefs about how her words draw distinctions: it does not require that those beliefs be true, or even that there be anything to make her attitudes true.) Of course, it can and does happen that disagreements about this go unrecognized. I return to this in the final section.
 13. She might think that in some of the worlds S could have known whether P, even though S does not know whether P. Or, she might be an extreme standards skeptic who claims, not just that S does not know whether P, but also that S could not have known whether P. This would depend on whether the capacity she claims is needed for knowing that P is one S could have exercised.
 14. Since this paper was completed, the number and variety of examples has exploded. One especially influential recent discussion is in (Stanley, 2005).
 15. It is not clear to me, though, why it is more plausible to hold that “is a blackbird” is context-dependent than to hold that “knows” is context-dependent. In any event, this is a debate about the specific mechanisms involved, and not about whether sentences of the form “S knows that P” can be used to ascribe different cognitive capacities in different contexts.
 16. A common response to Contextualism is that the context-dependence affects only what *speakers* say with sentences of the form “S knows that P” and not what these sentences themselves *literally* say. So far as I can see, this response assumes that a presupposition-free context is indeed possible. This response is at the heart of (Rysiew, 2001), who relies heavily on a “rough-and-ready” distinction between (the tentidiously labeled) “loose” and “literal” uses of such sentences. I doubt this distinction can be sharpened to do the kinds of work Rysiew needs it to do. But even if it could be made good, the fact would remain that there are different things to say in using the words “S knows that P”, which is what Contextualism claims.
 17. Some uses of “here” and “now” generate this kind of situation. For when it is an open question where (or when) the speaker is located, it is an open question what she would be saying in using these words. But these are the exception.
 18. I discuss these issues, though not in the case of knowledge, in (Hunter, 2007).
 19. (Quine, 1953).
 20. I am indebted to Kent Bach, David Braun, Ram Neta, Cara Spencer and Robert Stainton for helpful comments on earlier drafts of this paper.

References

- Cohen, S. (1987). Knowledge, context, and social standards. *Synthese*, 73, 3–26.
- Cohen, S. (1999). Contextualism, skepticism, and the structure of reasons. *Philosophical Perspectives*, 13, 57–89.
- DeRose, K. (1992). Contextualism and knowledge attributions. *Philosophy and Phenomenological Research*, 52, 913–29. Reprinted in L. Alcock (Ed.), *Epistemology: The big questions* (pp. 109–123). Malden, MA: Blackwell Publishers. References are to this reprinting.

- DeRose, K. (1999). Contextualism: An explanation and defense. In J. Greco & E. Sosa (Eds.), *The blackwell guide to epistemology* (pp. 187–205). Malden, MA: Blackwell Publishers.
- DeRose, K. (2004). Single scoreboard semantics. *Philosophical Studies*, 119, 1–21.
- Dretske, F. (1991). Knowledge: Sanford and Cohen. In B. McLaughlin (Ed.), *Dretske and his critics* (pp. 185–196). Cambridge, MA: Blackwell Publishers.
- Feldman, R. (2001). Skeptical problems, contextualist solutions. *Philosophical Studies*, 103, 61–85.
- Grice, P. (1989). *Studies in the way of words*. Cambridge, MA: Harvard University Press.
- Hunter, D. (2007). Common ground and modal disagreement. In H.V. Hansen, et al. (Eds.), *Dissensus and the Search for Common Ground*, CD-ROM (pp. 1–7). Windsor, ON: OSSA. *Proceedings of the 2007 Ontario Society for the Study of Argumentation*.
- Lewis, D. (1983). Scorekeeping in a language game. In D. Lewis (Ed.), *Philosophical papers volume 1*. New York: Oxford University Press.
- Lewis, D. (1996). Elusive knowledge. *Australasian Journal of Philosophy*, 74, 339–359. Reprinted in D. Lewis (Ed.), *Papers in metaphysics and epistemology*, 2006 (pp. 418–445). Cambridge: Cambridge University Press.
- Quine, W.V.O. (1953). Two dogmas of empiricism. In W.V.O. Quine (Ed.), *From a logical point of view*. Cambridge MA: Harvard University Press.
- Rysiew, P. (2001). The context-sensitivity of knowledge attributions. *Nous*, 35, 477–514.
- Schiffer, S. (1996). Contextualist solutions to scepticism. *Proceedings of the Aristotelian Society*, 96, 317–333.
- Stalnaker, R. (1984). *Inquiry*. Cambridge, MA: MIT Press.
- Stalnaker, R. (1999). *Context and content*. New York: Oxford University Press.
- Stalnaker, R. (2002). Common ground. *Linguistics and Philosophy*, 25, 701–721.
- Stanley, J. (2004). The linguistic basis of contextualism. *Philosophical Studies*, 119, 119–146.
- Stanley, J. (2005). *Knowledge and practical interests*. Oxford: Oxford University Press.
- Travis, C. (1989). *The uses of sense*. Oxford: Oxford University Press.
- Williams, M. (1991). *Unnatural doubts: Epistemological realism and the basis of scepticism*. Cambridge, MA: Blackwell Publishers.

On Failing to Capture Some (or Even All) of What is Communicated

Kent Johnson

Abstract This paper examines a methodological argument launched against Cappelen and Lepore’s “minimalist” semantics. The charge is that this semantic theory – and by implication a great many other ones – cannot be correct, because they fail to capture some of the “intuitive truth conditions” of the relevant sentences. In response, I argue that this charge rests on the claim that an acceptable scientific theory must (at least sometimes) capture all of the overt phenomena under study. But this claim, I contend, is false. In actual practice, scientific models will often never capture all of the behavior of the relevant phenomenon, and this feature does not undermine them as such. I maintain that semantic theorizing is just an instance of this more general aspect of scientific methodology.

Keywords Linguistics · semantics · model · error · methodology

1 Introduction

Recently, Herman Cappelen and Ernie Lepore (hereafter CL) have developed and defended a particular view, semantic minimalism, regarding the general structure and scope of semantic theories for natural languages (CL, 2005). According to semantic minimalism, semantic context-sensitivity in our expressions is limited to a handful of expressions (e.g., *I*, *you*, *she*, *this*, *that*, *tomorrow*, etc.). In particular, there do not exist unobserved (i.e., phonologically and orthographically unrealized) context-sensitive elements in either the syntactic or semantic structure of an expression. As CL show in great detail, semantic minimalism contrasts sharply with virtually every other existing view about semantic theory. In particular, it is extremely common for semanticists and philosophers of language to assume that a correct semantics for (1a) and (2a) assigns the additional structure in (1b) and (2b) as part of the correct meaning of these sentences:

K. Johnson
Department of Logic and Philosophy of Science, University of California, Irvine, CA 92697-5100,
USA

- (1) a. Mary is ready;
b. Mary is ready *to X*.
- (2) a. It is raining;
b. It is raining *in location X*.

According to the standard view, *Mary is ready* means something like Mary is ready to do some particular salient activity, or is ready for something (in particular) to happen. Similarly, *it's raining* means that it's raining in some contextually specified place. Semantic minimalism denies these claims. Instead, the semantic minimalist claims that *Mary is ready* simply means that Mary is ready, and *It's raining* simply means that it's raining. There is no specification of a semantic complement of *ready* or of the location of the raining.

Unsurprisingly, these aspects of semantic minimalism have been met with numerous objections, which CL canvass and respond to (cf. Chapters. 11–12 of CL, 2005). According to one (cluster) of these objections, which I will call the Problem of Irrelevance, semantic minimalism results in a theory that is utterly irrelevant to psychology. Human linguistic practices are essentially driven by much more complex representations of meaning than what semantic minimalism offers; we *never* intend to communicate just the Semantic Minimalist's meaning of an expression. In short, there is no real cognitive role for a minimalist theory to play. Thus, a semantically minimalist theory is otiose, and so should be rejected as a style of semantic theorizing. In response to the Problem of Irrelevance, CL argue that there is a cognitive role for Semantic Minimalist meanings. They hold that we in fact do sometimes rely on such minimal meanings in our communicative practices.

In this paper, I wish to revisit the Problem of Irrelevance. I think that it relies on a common but false assumption about the nature of linguistic theorizing. The falsity of this assumption can be brought out with some considerations about the general nature of scientific theorizing. When we treat semantic as scientific theories about human language, we can see that there is no need to make the kind of head-on response to the Problem that CL attempt (which is just as well, since I'll argue that their attempt is unconvincing). Instead we can assert something much stronger: like any other scientific theory, a semantically minimalist theory may be on the right track even if speakers *never* intend to communicate the semantically minimal meaning of an expression. Indeed, we'll see that semantic minimalism may be on the right track even if speakers *cannot* (intend to) communicate such minimal meanings. If these claims are correct, then the Problem of Irrelevance has no force whatsoever.

This paper is structured as follows. In §2, I present the Problem of Irrelevance, and I critically examine CL's reply to it. In §3, I explore some properties of scientific models, and argue that semantic theories have these properties too. In particular, I argue that, in precise parallel to other scientific models, the particular outputs of a good semantic theory may be quite different from the communicative facts that form the raw data of a semantic theory. I then show how these facts invalidate the Problem of Irrelevance. I conclude in §4.

Before beginning, a couple caveats are in order. First, in this paper, I remain officially neutral about semantic minimalism. Although I will defend CL's view against

an objection, this defense should not be construed as an endorsement of it. Second, there are a number of issues concerning how to understand the position(s) of CL's critics, such as whether they are really offering different positions, and whether they aim to be working on (the foundations of) a scientific psychology (e.g., Stainton, 2006). The present paper has no stake in any of these issues. My interest concerns CL's methodology and a criticism of their theory. Thus, it doesn't really matter where the criticism comes from; e.g., even scientists with ultimately similar views to CL's can err in criticizing CL's view.

2 The Problem of Irrelevance

A number of authors have presented the Problem of Irrelevance for minimalist theories, and CL devote an entire chapter to discussing it (cf. CL, 2005, chap. 12 and citations therein). For instance, Recanati writes:

That minimal notion of what is said is an abstraction with no psychological reality, because of the holistic nature of speaker's meaning. From a psychological point of view, we cannot separate those aspects of speaker's meaning which fill gaps in the representation associated with the sentence as a result of purely semantic interpretation, and those aspects of speaker's meaning which are optional and enrich or otherwise modify the representation in question. They are indissociable, mutually dependent aspects of a single process of pragmatic interpretation. (Recanati, 2001, 88)

Recanati's pessimism about minimalist semantic theories is driven in large part by his view that a semantic theory must capture the entire "content of the statement as the participants in the conversation themselves would gloss it" (Recanati, 2001, 79–80). Recanati expresses this in his

"Availability Principle", according to which "what is said" must be analyzed in conformity to the intuitions shared by those who fully understand the utterance – typically the speaker and the hearer, in a normal conversational setting. This in turn supports the claim that the optional elements... (e.g., the reference to a particular time in "I've had breakfast") are indeed constitutive of what is said, despite their optional character. For if we subtract those elements, the resulting proposition no longer corresponds to the intuitive truth conditions of the utterance. (Recanati, 2001, 80)

The Availability Principle does not logically imply that minimalism is wrong, but it does tightly constrain minimalist theories. Minimalist theories, almost by definition, do not capture all of what Recanati calls the "intuitive truth conditions" of utterances of sentences. Thus, given the Availability Principle, a minimalist theory would need to defend itself by showing that it does in fact exactly capture the intuitive truth conditions of our utterances, at least sometimes. A bit more carefully, we can characterize the Availability Principle as follows:

(AP) A semantic theory is acceptable only if it correctly characterizes the intuitive truth conditions often enough within some psychologically interesting range of cases.

For present purposes, I will assume that (AP) is an appropriate formulation of Recanati's Availability Principle; any divergences between the two will not matter in this paper. One might strengthen (AP) further by specifying the particular range of cases in which a semantic theory must get things right, and by specifying how often the theory must get things right. I won't worry about such strengthenings, though; since what I have to say will apply equally to all such versions of (AP).

CL do not challenge the Availability Principle. Instead, they seem to accept it, and argue that there is a psychologically interesting range of cases where minimalist theories capture the intuitive truth conditions of utterances. This range of cases involve contexts where participants in the conversation (speakers, audience, others) must appeal to the "shared fallback content" of an utterance. In such cases, the minimalist's semantic content is what all participants can be assumed and expected to know even when some parties have mistaken or incomplete communication-relevant information. Moreover, this content is what can be conveyed by indirect quotation or reproduction to other parties in different contexts from the context of utterance (CL, 2005, 184–185). For example, suppose Jolene utters *Mary is tall*, and the audience disagrees about whether Jolene means that Mary is tall for a girl, or for a Samoan, or for a basketball player, or for a college student, or for something else. In such a state of ignorance, CL argue, the audience can still give some sort of characterization of the meaning of Jolene's utterance. For instance, they can report that Jolene said that Mary is tall, and that she didn't say that Mary is short, or that Peter will be tall. This agreement about the meaning of Jolene's utterance is the shared fallback content of Jolene's utterance, and, CL claim, it is precisely what minimalist theories characterize as the semantic content of the utterance.

Despite the cleverness of CL's argument, it's hard to see why an opponent of semantic minimalism would find it compelling. Opponents such as Recanati might concede that CL had identified a psychologically interesting range of communicative contexts, but deny that the minimalist's theory expresses the intuitive truth conditions in these cases. In particular, it would be natural for them to hold that in the cases CL imagine where the context is unknown or incomplete, our communicative expectations revert to default interpretations that nonetheless contain more content than the minimalist theory offers. For instance, in a context like the one just imagined, it would be reasonable to argue that we would expect and assume the various parties to interpret Jolene's utterance of *Mary is tall* as meaning that Mary is tall for an X, where the particular value of X is assumed and expected to normally be given by the context in some appropriate way. In the present case, the context failed to supply a value for X, so the shared fallback content includes something like a representation of a variable that has not been assigned a value. Thus, an opponent of minimalism could conclude, even in the cases CL imagine, there is still much more semantic structure present in the content of utterances than the minimalist predicts. Thus, the minimalist still needs to satisfy (AP) if her theory is to be acceptable.

Where does this leave us? At best, we seem to have reached an impasse of intuitions. CL claim to have found a way for minimalist theories to satisfy (AP), and their opponents have reasonable grounds for denying this. Fortunately, I think some

headway can be made on this issue. Unsurprisingly, I think that part of the problem lies with (AP) itself. In the next section, I will argue that we have good reasons for doubting both the truth of (AP) and the soundness of the Problem of Irrelevance.

3 Linguistic Theories as Scientific Theories

Before getting started, let's look at the general structure of the Problem of Irrelevance. The argument appears to take the following form:

- (3i) In all relevant ranges of cases, the intuitive truth conditions of our utterances contain much more content than what is characterized by minimalist theories.
- (3ii) If (3i) is right, then from a psychological point of view, we cannot separate the minimalist aspects of meaning from those aspects supplied by a more enriched view of meaning (often enough, in any relevant range of cases).
- (3iii) Hence, minimalist aspects of meaning cannot be separated from those aspects supplied by a more enriched view of meaning (often enough, in any relevant range of cases).
- (3iv) But if we can't separate minimalist from non-minimalist elements of meaning (often enough, in any relevant range of cases), then minimalist theories are unacceptable.
- (3v) Hence, minimalist views are unacceptable.

Premise (3i) is an empirical claim; CL attempt to refute it by appealing to cases where we employ shared fallback content. However, we've seen that their attempt at best leads to an unresolved clash of intuitions between them and their opponents. That leaves only premises (3ii) and (3iv). Premise (3ii) comes from the quote of Recanati above (2001, 88), and premise (3iv) comes from (AP). (To see this, notice that if we can never separate out the minimalist aspects of meaning, then there must always be some non-minimalist aspects present, so the minimalist aspects of meaning never characterize the intuitive truth conditions in the utterance. Hence, by (AP), minimalist theories are unacceptable.)

In this section, I argue that neither (AP), (3ii), nor (3iv) is credible, largely for the same reasons. In order to do this, I will assume that linguistic theories, and semantic theories in particular, are scientific theories (or "models", as I occasionally call them). I consider this assumption, and my use of it in this paper, to be utterly trivial. Nonetheless, I want to register this assumption up front, since one still occasionally encounters philosophers who distinguish "doing philosophy" from "doing science". I'll say a bit more about this assumption later on, but for now let's turn to the central discussion.

In order to see what is wrong with (AP), (3ii), and (3iv), it will be useful to step back from linguistic theorizing and examine some aspects of the methodology of the (other) sciences. Ultimately, I maintain that when we look at the details of how a typical quantitative empirical theory is related to the raw data, we can precisely identify counterparts in the linguistic theory. The upshot of all this will be that just as other scientific theories of particular phenomena need not – and very often do

not – directly account for the raw data, so too linguistic theories will often fail to fully explain the raw linguistic data that drives our theories. And this divergence between the linguistic theory and the data is a perfectly normal aspect of perfectly normal science. In short, no non-linguistic theories would ever be constrained to observe appropriate counterparts of (AP), (3ii), or (3iv). These features of scientific theories that free them from such constraints are found in linguistic theories too. Hence, (AP), (3ii) and (3iv) are unacceptable for linguistics for just the same reasons that their counterparts are unacceptable in the sciences.

To set this discussion in context, it's worth observing that Chomsky has long defended parallels between linguistics and other sciences (e.g., Chomsky, 2000). My discussion is certainly of a piece with Chomsky's, but it differs in one crucial respect. Rather than attending to purely verbal characterizations of historical examples, I focus on some quantitative aspects of current scientific methodology. Thus, we will be able to address certain issues with considerably more precision and detail than Chomsky offers. It is perhaps also worth mentioning that the present approach also makes possible a critical examination and principled rejection of some claims of Chomsky's that in terms of its methodology, linguistics is just like the other sciences. I discuss this issue in detail elsewhere (Johnson, 2007b).

To get things started, let's take a simple example. Suppose we are studying the relationship between different quantities of a given additive X used in some manufacturing process and the amount of some type of atmospheric pollution Y generated by the process. The industry standard is to use n units of X per ton of product, but for a period of time, certain companies used more or less than n units. The relation between the varying amounts of X used and Y emitted are given as black diamonds in the plot below. (Zero on the x -axis represents the use of n units of X ; other values represent the respective deviations from this standard; ignore the two curves and white diamonds for the moment.)

Given this data, there are many – infinitely many, in principle – possible relations that could hold between X and Y . One extreme option would be to insist that every aspect of the data is crucial to understanding how X and Y are related. In such a case, a researcher might look for a function that captured the data precisely, as in the very complex one depicted with a solid line. The resulting theory will then perfectly predict the behavior of Y on the basis of the behavior of X . The raw data, in the form of a set of pairs of measurements $\{ \langle X_i, Y_i \rangle : i \in I \}$, is fully accounted for.

Despite its success at capturing the data, the first approach is almost never adopted. A vastly more common strategy is to assume that the real relation between X and Y is much simpler, and that Y is influenced by other factors that are unrelated to X . One might, e.g., hypothesize that relationship is given by the simple function:

$$(4) \text{ Predicted value of } Y_i = f_1(X_i) = \beta_0 + \beta_1 X_i + \beta_2 X_i^2$$

for some fixed numbers $\beta_0, \beta_1, \beta_2$. Once these numbers are determined from the data, we get the much simpler curve given by the dashed line. In the present example, the values of β_0, β_1 , and β_2 were determined by seeking those values for which $\sum_{i \in I} [(Y_i - f_1(X_i))^2]$ is as small as possible. If we simply wanted a polynomial

function that captured the data perfectly, with probability 1, f_2 , a polynomial of degree 29, will do so, for the given raw data set of size 30:

$$(5) \text{ Predicted value of } Y_i = Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \dots + \beta_{29} X_i^{29}$$

Although f_1 doesn't predict the behavior of the original data as well as its rival f_2 , many other theoretical considerations speak in its favor. For example, suppose we got hold of another sample of data, given by the white diamonds above. Then we might ask how well the two functions captured this new data. One way to do this would be to compare the sizes of the discrepancies between what f_1 and f_2 predict about the value of Y for given values of X in the new data set. E.g., we might examine the ratio:

$$(6) \frac{\sum_{i \in I'} [(Y_i - f_1(X_i))^2]}{\sum_{i \in I'} [(Y_i - f_2(X_i))^2]}$$

Here I' indexes the second set of measurements, and f_1 and f_2 are assumed to have had the particular numerical values of their parameters $\{-\beta_0, \beta_1, \beta_2\}$ in the case of f_1 , and $\{\beta_k : 0 \leq k \leq 29\}$ in the case of f_2 – fixed by the first data set. In the particular case presented in Fig. 1, (6) yields a value around 6×10^{-31} , indicating that there is vastly less discrepancy between the new data and what f_1 predicts than there is between this data and what f_2 predicts. (In a typical case, this number would be even smaller, since there are $n - 3$ and $n - 30$ degrees of freedom associated with f_1 and f_2 respectively; (6) does not factor in this discrepancy.) From a God's-eye view, this is unsurprising, because f_1 is the form that actually generated the data (although the data themselves determined the values of the β_i s).¹ The extra structure in the curve given by f_2 errs by capturing much of the variance in the data that is unrelated to the true relation between X and Y. (There's much more to be said about the general issues of model construction and model evaluation; cf. e.g. Forster and Sober, 1994; Burnham and Anderson, 2002; Bamber and van Santen, 1985 for further relevant discussion.) In short, a bizarre model like f_2 that captures all the

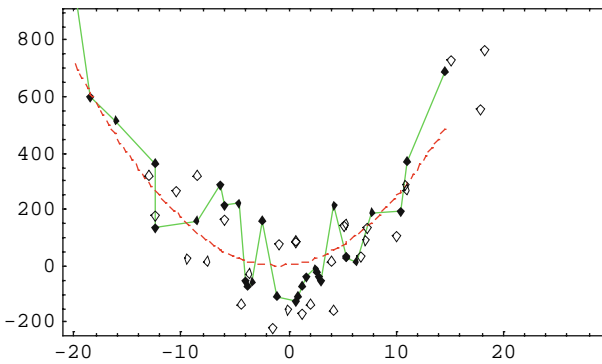


Fig. 1 Deviations from the normal amount of additive X versus amount of pollutant Y

(original) data is vastly inferior to the far more standard model like f_1 that doesn't. In particular, the simpler model does a massively better job at predicting the general trends of new data as it arrives.

How then is f_1 related to the actual raw data? Typically, this connection is formed by adding a "residual" or "error" term to our equation:

$$(7) Y_i = f_1(X_i) + \varepsilon_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \varepsilon_i,$$

The term ε_i , whose value varies as i varies, expresses whatever deviation is present between the model and the raw data (i.e., $\varepsilon_i = Y_i - f_1(X_i)$). In practice, scientific models of complex phenomena *never* perfectly fit the data, and there is *always* a residual element (ε_i) present. This is so even when the system under study is completely deterministic, with no random influences present in the data whatsoever. E.g., the true model might be something like

$$(8) Y_i = f_1(X_i) + f_3(Z_{li}, \dots, Z_{ki})$$

In such a case, Y is always an exact function of X and Z_1, \dots, Z_k . However, the influence of the Z_j s may be very small, very complicated, unknown, poorly understood, etc. Thus, for any number of reasons, it may be natural to model the phenomena with f_1 , all the while realizing that the existence of residuals in the raw data show that there is more to the full story than f_1 presents.

With this brief overview of scientific theorizing in hand, let's return to linguistics. In the relevant respects, linguistic theorizing is similar to the statistical case. Like f_1 , semantic theories are models of a complex phenomenon (the interpretation of language). The raw data of a sample of this phenomenon aren't represented as numerical measurements, as in the statistical example. Instead, they are given as assessments about certain types of idealized² linguistic behavior: what sorts of things would typical speakers communicate by uttering a given sentence, and under what conditions? In other words, the raw data of semantic theorizing are the intuitive truth conditions of our utterances, as we do or would make them in various contexts. Proceeding in a manner similar to the statistical researcher, the minimalist theorist begins by hypothesizing that there is some relatively simple structure – i.e. simple in comparison to the complexity of the raw data – that accounts for much of the collective behavior of the raw data. In order to obtain this relatively simple underlying general structure, some aspects of the raw data (i.e., the intuitive truth conditions) will have to be ignored, just as we ignore some aspects of variance in the statistical case. Here the analogy between the minimalist and statistical strategies is quite tight. In the statistical case, the raw data was decomposed into two parts: the pollution level Y is the additive combination of the effects of X (given by the population model $\beta_0 + \beta_1 X + \beta_2 X^2$) and some other effects (given by ε) that are not part of the present study. The general structure of this statistical model is thus:

$$(9) \text{Raw Data} = \text{(i) Effects of processes under study (ii) Interacting in some way with (iii) Residual Effects}$$

But, and this is a crucial point, (9) is also the general structure of the minimalist's semantic theory. The minimalist theory supplies some aspects of meaning that are

hypothesized to capture much of the general behavior of the *totality* of the data set. By assumption, the outputs of this theory needn't capture all of the raw data (i.e., the intuitive truth conditions of the utterances in the data set). In fact, the semantic theory needn't *ever* capture all of the intuitive truth conditions. Such an outcome is absolutely standard science. For instance, our pollution researcher would not assume that there will be some raw datum Y_i such that $Y_i = f_1(X_i)$, with no contribution from the residual effects. Indeed, it is quite typical to expect that ε_i will never equal 0, particularly when the phenomenon under study is extremely complex. (Indeed, when the phenomena are quite complex, a model may be considered significant even if it captures as little as 16% of the variation in the raw data (e.g. Putnam, 2000, 487).) In the case of semantics, the intuitive truth conditions (on analogy with the Y_i) are perhaps always partly determined by the minimalist theory of meaning (on analogy with the $f_1(X_i)$), and partly determined by other aspects of communication that interact with the minimalist theory in some way (on analogy with the ε_i). These other aspects of communication are just the familiar processes that are hypothesized to interact with the hypothesized minimalist linguistic capacities: background beliefs, indexical-fixing elements, demonstrations, "performance" capacities of speaker/hearers, etc.

As an aside, it's worth noting (again) just how "Chomskian" my interpretation of semantic minimalism is. In many places, Chomsky cautions that linguistic theories are not obliged to capture all the facts about various grammaticality judgments, or all of various details present in collections of data. By seeking out more general patterns, we may be able to learn about a speakers' linguistic "competence", which can be masked by additional "performance" factors that are also realized in the empirical data. For instance, Chomsky writes:

Linguistic theory is concerned primarily with an ideal speaker-listener, in a completely homogeneous speech community, who knows its language perfectly and is unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors (random or characteristic) in applying his knowledge of the language in actual performance. (Chomsky, 1965, 3)

By idealizing away from various extraneous factors, we can "smooth" out the raw data and thereby focus on some of the more significant forces and structures underlying human linguistic abilities.

To return to our main theme, a minimalist semantic theory is a *theory* about the nature of the raw data. Like any other scientific theory, one of its essential rights and obligations is to characterize those parts of the raw data it considers to be truly part of the phenomenon under study, and what other parts are due to extraneous processes; cf. (i) and (iii) in (9). From the standpoint of ordinary scientific theorizing, this point is too obvious to merit mention. But in the present case, for some reason, the point appears to some to be completely optional. Others have noticed this odd practice, too. For instance, Paul Pietroski writes that

one shouldn't try to define the scope of semantic theorizing in advance of investigation. . . . One cannot determine a priori which facts a semantic theory should explain. In any domain, what a theory *should* explain depends on what gets discovered. And we can discover

semantic facts, just as we can discover biological facts. In my view, insensitivity to this point still pervades philosophy of language, despite Chomsky . . . and the remarkably successful research program he initiated. (Pietroski, 2005, 26–27)

(Pietroski provides ten citations spanning six decades where Chomsky has urged this point; many more citations could be provided.) The fact that a semantic theory gets to characterize its own scope also means that, like any other scientific theory, a semantic theory should be judged by the ordinary, complicated but familiar, criteria of successful theories: simplicity, elegance, predictive fecundity, integration with other successful theories which collectively account for the raw data (or, more typically, hopefully someday will account for the raw data), etc. Methodologically speaking, demanding that a semantic theory sometimes exactly characterize the intuitive truth conditions of utterances appears to be just like demanding that statistical models should (at least for some interesting range of values) be like the complex f_2 , instead of the like the more common f_1 . Such a demand would be bizarre and deeply incorrect in the statistical case; I submit it is no better motivated in the case of semantic theorizing.

The points just made are absolutely crucial to the understanding and evaluation of a semantic theory. Moreover, they form the core of why (AP), (3ii), and (3iv) should be rejected. The problem with (AP) and (3iv) is that they place an unwarranted constraint on theory construction. In any other study of complex phenomena, the demand that theories perfectly capture the raw data across some interesting range of cases would grind theorizing to a halt. The only work that could be done would be purely descriptive and utterly unpredictable and unexplanatory characterizations of the available data, along the lines of f_2 . (3ii) should be rejected because it is one of the rights (and obligations) of a theory to provide a theoretically useful characterization of the phenomena it addresses. Typically this is done by constructing a good theory, and then declaring that the relevant phenomena are whatever the theory addresses. In the case of scientific models, this is the essence of why they are constructed and confirmed using statistical methods. There is no reason to deny such rationale to semantic theories, contrary to the announcement in (3ii). Thus, the Problem of Irrelevance rests on two assumptions which we have little reason to believe. So we have little reason to think that the Problem of Irrelevance poses a real problem at all.

There is another way to view the flaws with (AP), (3ii) and (3iv) that helps to bring out why the Problem of Irrelevance might appear sound. The Problem of Irrelevance fundamentally turns on an equivocal interpretation of “separability”. As Recanati and many others have made clear, the raw data of linguistic theory – the intuitive truth conditions of our utterances – are probably nearly always substantially influenced by pragmatic factors. In this sense, Recanati is probably right that pure semantic content is “inseparable” from pragmatic factors: in actual language use, you rarely if ever find the former alone, without the latter. Indeed, the contribution of pragmatics is probably thorough enough to make this last claim hold very generally. This notion of inseparability makes (3ii) plausible (recall that (3ii) says that if intuitive truth conditions typically contain more content than what the minimalist

theory supplies, then we can't separate minimalist from non-minimalist aspects of meaning). However, such an interpretation of inseparability makes it unclear why (AP) and (3iv) should hold (recall that (3iv) says that minimalist theories are acceptable only if we can (sometimes) separate minimalist and non-minimalist aspects of meaning; (AP) says that a semantic theory is acceptable only if it (sometimes) correctly characterizes the intuitive truth conditions). After all, we've seen that it's no criticism of a theory that it treats the raw data as being a product of multiple sources. If this is what separability amounts to, then the claim that minimalist and non-minimalist aspects of meaning are inseparable simply begs the question against minimalist theories. Minimalists will certainly hold that their theories are reasonable, even though these theories purport to explain the two types of aspects of meaning using different mechanisms.

How then might we make (AP) and (3iv) plausible? To do this, we need a notion of inseparability that applies to linguistic *theories*. That is, (AP) and (3iv) are plausible if inseparability here means that no reasonable total theory of language will treat minimalist and non-minimalist aspects of meaning as effects of (relevantly) distinct processes. That is, in order for (AP) and (3iv) to be plausible, the relevant notion of inseparability must require that all aspects of the intuitive truth conditions be explained by the same mechanisms. Then (AP) and (3iv) are virtually tautologies, hence trivially true. But now there is no support for (3ii). Why should the fact that the intuitive truth conditions of our utterances *do* contain both minimalist and non-minimalist aspects of content be sufficient to license the restriction that any theory of semantic content *must* capture all of these aspects? Such a view clearly begs the question against minimalist semantic theories, which propose to capture only part of the intuitive truth conditions of our utterances. Moreover, when (3ii) is interpreted this way, it is independently implausible for reasons we have already discussed; (3ii) would require semantic theories to account for all of the raw data, in dramatic contrast to the way all other theorizing proceeds.

In short, (3ii) is a kind of converse of (AP) and (3iv). (3ii) becomes plausible when we interpret inseparability as implying that our intuitive truth conditions are almost always the product of minimalist and non-minimalist aspects of meaning; however, this makes (AP) and (3iv) question-begging and implausible. On the other hand, (AP) and (3iv) are plausible when we interpret inseparability as implying that no reasonable theory will separate minimalist and non-minimalist aspects of meaning; however, this makes (3ii) question-begging and implausible. The problem with either interpretation of inseparability is the same: there's no apparent way to motivate the inference from the empirical fact that our utterances typically contain minimalist and non-minimalist aspects of meaning to the normative claim that all these aspects must be captured by a semantic theory. Thus, we can locate the structural flaw in the Problem of Irrelevance in an equivocation regarding inseparability.

In the final analysis, a lot of the present discussion hangs on the current epistemological situation with respect to linguistics. I've stressed that part of a scientific theory involves carving up the data in one way or another. This aspect of theorizing is especially important in linguistics because of how little we know about how language works. If we had a better idea of how it works, this additional knowledge

would likely constrain what structures, processes, etc. could be plausibly employed by a theory. As it is, though, our lack of knowledge about language leaves us with very few such constraints on theories (at least at the level of detail that is currently relevant here – i.e., a level where we may ask whether minimal propositions play some psychological role concerning language).

The difficulty with constraining which entities, etc. a theory may reasonably posit is further compounded by the fact that there is tremendously strong evidence that many entities may not be introspectively accessible by anyone. To see this, notice that many aspects of cognitive processes are not consciously perceivable by persons undergoing them. The many decades of literature on both human vision and judgment and decision-making have presented countless studies showing that many aspects of these types of cognitive processes are not consciously accessible to us. Similarly, people lack conscious awareness of many hypothesized aspects of language and its processing (Johnson, 2007a; cf. Townsend and Bever, 2001 for an overview). E.g., people aren't aware of such items as traces, PRO, movement, Merge, etc. that are posited by *some* linguistic theories simply because of the theoretical work they do. But opposing theorists cannot simply announce that the former theories are false, because they employ some entity which they *feel* is not truly part of language or its processing. The entity in question may well be part of the end product or part of the processing, or both, even if people are not generally consciously aware of it. (Indeed, even in the case of semantics, there is evidence that people have substantial difficulties becoming aware of certain aspects of the meanings of relatively ordinary sentences (Johnson, 2007a).)

In the present case, then, it's not enough for Recanati to simply declare that CL's minimal propositions won't find any place in a scientific theory of language. He also needs to supply enough *uncontroversial* details of a theory of language to show that no such minimal propositions will play any role in a completed correct theory. Lacking such details, CL's theory is not objectionable on these grounds. (I take it that this is just a general point about the nature of scientific theories in general. Much of the philosophy of science concerns how scientific theories will often posit unobserved entities, where the justification for treating these entities as real comes only later, as the theory is confirmed by the usual holistic criteria – simplicity, accuracy, integration with larger theories, predictive fruitfulness, etc. These are the criteria by which a minimalist theory should be judged, not by apriori speculation about what entities will appear in a completed theory of language.)

4 Rescuing the Attack on Minimalism?

In the previous section, I levied a charge against the Problem of Irrelevance. To those familiar with Chomsky's work, this charge may seem familiar. To use Chomsky's words, Recanati and the other proponents of the Problem of Irrelevance are guilty of "methodological dualism" (e.g., Chomsky, 2000, chap. 4). Methodological dualism is characterized by the insistence on a set of standards for some area of research (semantic theories in this case) that would be utterly inappropriate and trivially

unacceptable in any other area of inquiry. Chomsky has long argued that the study of language and mind is held to just such dualistic standards. Furthermore, he has argued, such standards would cripple any other developed discipline such as chemistry or physics, were they applied there.

Now of course, there is nothing *a priori* wrong with this sort of methodological dualism. Indeed, maybe language and mind are “special” in some important sense. Maybe language and mind are so radically unlike every other complex phenomena that humans have theorized about that such dualistic standards are appropriate. Perhaps assumptions like (AP) are appropriate constraints on semantic theories, even though their corresponding forms would be a travesty anywhere else. Recanati appears to believe something like this, because he provides an argument in favor of (AP). If this argument is compelling, perhaps it could justify the unique standards he imposes on semantic theories. Thus, it will behoove us to briefly examine this argument.

The essence of Recanati’s argument is given in the passage below:

Suppose I am right and most sentences, perhaps all, are semantically indeterminate. What follows? That there is no such thing as ‘what the sentence says’ (in the standard sense in which that phrase is generally used). . . . If that is right, then we cannot sever the link between what is said and the speaker’s publicly recognizable intentions. We cannot consider that something has been said, if the speech participants themselves, though they understand the utterance, are not aware that that has been said. This means that we must accept the Availability Principle and its maximalist consequences. (Recanati, 2001, 87–88)

This argument is puzzling. Given the context of this argument, Recanati’s claim that most or all sentences are “semantically indeterminate” amounts to the claim that minimalist semantic theories don’t capture the intuitive truth conditions of most or all sentences. It is unclear, though, why such a claim should be taken to imply that “there is no such thing as ‘what the sentence says’ (in the standard sense in which that phrase is generally used)”. I take it that “what the sentence says” here refers to the content that a (minimalist or other standard) semantic theory ascribes to a sentence. If that is correct, then claiming that there “is no such thing” is question-begging. We’ve seen that it’s part of the job of a theory of some complex phenomenon (semantic or otherwise) to carve out the exact sub-portion of the phenomenon that it directly deals with, leaving the remaining parts for further theorizing. Recanati’s claim that there is no such thing as “what is said” in this context is like saying there is no such thing as the true population model f_1 in the statistical case. I conclude that this inference is invalid, for largely the same reasons that (AP), (3ii), and (3iv) are unjustified. Hence, I do not believe that Recanati has successfully resurrected his case against minimalism.

How might one defend a claim like (AP)? (I focus on (AP), although similar considerations also hold for similar claims like (3ii) and (3iv).) The most obvious strategy, of course, would be to come up with a rival semantic theory that has this form, and show that it is to be preferred overall, using the standard criteria for the selection and confirmation of scientific theories. This strategy reverses Recanati’s: instead of defending a type of theory as preferable *because* it respects (AP), one

would need to defend (AP) because a type of theory that is independently determined to be preferable respects it.

Barring such a straightforward approach, there appear to be two general types of defenses of (AP). Both types attempt to show that semantics is unlike all other empirical disciplines insofar as residual effects contribute nothing to the raw data in some (or enough) instances. On the one hand, one might try to provide a *metaphysical* defense of such a principle. For instance, one might argue that no matter how integrated the phenomenon of language is within a larger interacting system (e.g., human cognition, social interactions, etc.), the phenomenon must occasionally deliver “pure” manifestations of itself. Unfortunately, this strategy looks hopeless, at least insofar as it is an empirical fact about language that the raw data we use to study it (i.e., the intuitive truth conditions of utterances) are the expression of multiple different systems. (Moreover, it also appears that we have strong independent reasons for individuating and distinguishing these systems, so that the previous claim is not arbitrary.) If, as the empirical research very strongly suggests, this is the case, then it’s an open empirical question whether the raw data is ever a pure expression of language only, or whether it always also contains other, extra-semantic, aspects (e.g. Gleitman and Liberman, 1995) Thus, it looks like any such metaphysical argument may run afoul of the empirical facts.

On this note, it is also an open, and quite interesting, question whether one could construct a context, perhaps in the form of a carefully designed experimental setting, in which subjects would reliably offer all and only what the minimalist characterizes as the meaning of our sentences. Indeed, one way to study a minimalist theory would be to look for a single context in which subjects reliably offered the minimalist’s truth conditions as the intuitive truth conditions for a broad range of expressions. Finding “interesting” contexts of this sort (e.g., *not* contexts like “Give me the minimalist truth conditions of the following utterance. . .”) could prove quite interesting. Finding such contexts could also falsify (3i) in the Problem of Irrelevance.

The second type of strategy for defending (AP) is *epistemological*. In this vein, one might argue that any semantic theory must be sophisticated enough to occasionally completely explain exactly how all the component effects work to produce the raw data. As I noted in the previous paragraph, it’s an interesting question whether there are relevant contexts in which the minimalist’s truth conditions are identical to the intuitive truth conditions. But let’s leave that issue aside. In general, we cannot demand that a theory of some complex phenomenon have this level of sophistication, for the simple reason that the theory would have to be so complex that we wouldn’t know how to go about constructing it. With a highly complex phenomenon like language, it’s vastly more reasonable to identify some (hopefully) “natural part” of the phenomenon, and work with that, letting the resulting theory determine the nature of the natural part as inquiry progresses. To demand that our theories account for all of the raw data would be to ask for a theory that we don’t have any idea how to construct (and couldn’t use if we ever did build it). In the present case, it would have to be at least a near-total theory of human psychology. And to place this kind of requirement on a theory would be to grind research to a halt. Hence such a requirement should be rejected.

Incidentally, it's worth observing that Recanati himself appears to reject any such requirement on semantic theories. After all, the raw data of linguistics are *individual persons'* assessments of the intuitive truth conditions of utterances in contexts. Recanati's discussion suggests that he is following standard practice by abstracting away from the "residual" disagreements between persons (or between one person at different times). One wonders why such a maneuver would be legitimate for Recanati here, if it is not also legitimate in the case at hand. Of course, abstracting away from differences in judgments is necessary if one hopes to carve out a project that can be addressed and worked on. Precisely similar considerations are what motivate many semanticists, minimalists included, to further abstract away from the (idealized) intuitive truth conditions of utterances in the hopes of uncovering some general structure.

5 Conclusion

We've seen that the Problem of Irrelevance fails primarily because it ignores a basic fact about research into complex phenomena. This fact is that each theory gets to determine what part of the phenomena it addresses, and typically this is only a very proper subpart of the total phenomena. The requirement that a theory accommodate all of the intuitive truth conditions often enough in some relevant range of cases is a restriction on semantic theories that has no precedent in any of the developed sciences. Indeed, it is far more typical to assume that a given theory will not account for the data. Theories typically don't account for all the raw data for both the metaphysical reason that the data are typically the expression of multiple interacting processes, and for the epistemological reason that any such theory would be too hard for us to even begin constructing, and probably too hard for us to confirm or even to understand, if we were able to construct it.

On a final note, I have occasionally encountered the view that my assumption that semantic theories are scientific theories is somehow optional, and that one could avoid all these problems by stipulating that Recanati's position is a "philosophical" view. I confess I simply don't understand this position (not least because I do not understand what a distinctively "scientific" or "philosophical" theory would be). Nonetheless, there are a few things that can be said in response. First, I mean very little by calling a semantic theory "scientific". Semantic theories deserve this appellation, I suggest, primarily because their construction and confirmation centrally involve employing some of our best known methods for obtaining knowledge about a particular empirical phenomenon (the interpretation of language). From this perspective, it is unclear how one could reasonably defend the importance of a "non-scientific" theory of language. Moreover, my present use of the idea that semantic theories are a type of scientific theory is, I believe, especially uncontentious. So if one wants to resist my conclusions by taking this route, one needs to show why the particular features of semantic theorizing that I've appealed to are not part of some other form of theorizing *that can be independently motivated as an interesting and worthwhile project*. I see no way of doing this.

Acknowledgment I am grateful to Rob Stainton for much useful feedback.

Notes

1. The data were generated using the formula $Y_i = 3 + 4X_i + 2X_i^2 + \varepsilon_i$, where ε was normally distributed with a mean of zero and a standard deviation of 100. The values for X were randomly drawn from a normal distribution with a mean of zero and a standard deviation of 10.
2. The notion of idealization in linguistics and the other sciences has been discussed at great length in many places (e.g., Liu, 2004; Chomsky, 1986, and citations therein). Since the primary data of interest in the present paper concerns “intuitive truth conditions”, the idealizations at play here are substantially less (although by no means absent!) than in other areas of linguistics.

References

- Bamber, D., & van Santen, J. P. H. (1985). How many parameters can a model have and still be testable? *Journal of Mathematical Psychology*, 29, 443–473.
- Burnham, K., & Anderson, D. R. (2002). *Model selection and multimodel inference* (2nd ed.). New York: Springer.
- Cappelen, H., & Lepore, E. (2005). *Insensitive semantics*. Oxford: Blackwell's.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Chomsky, N. (1986). *Knowledge of language*, Westport, CT: Praeger.
- Chomsky, N. (2000). *New directions in the study of language and mind*. Cambridge: CUP.
- Forster, M., & Sober, E. (1994). How to tell when simpler, more unified, or less ad hoc theories will provide more accurate predictions. *British Journal for the Philosophy of Science*, 45, 1–35.
- Gleitman, L., & Liberman, M. (Eds.). (1995). *An introduction to cognitive science, volume I: Language*. Cambridge, MA: MIT Press.
- Johnson, K. (2007a). Tacit and accessible understanding of language. *Synthese*, 156, 253–279.
- Johnson, K. (2007b). The legacy of methodological dualism. *Mind and Language*, 22, 366–401.
- Liu, C. (2004). Laws and models in a theory of idealization. *Synthese*, 138, 363–385.
- Pietroski, P. (2005). *Events and semantic architecture*. Oxford: OUP.
- Putnam, R. (2000). *Bowling alone*. New York: Touchstone.
- Recanati, F. (2001). What is said. *Synthese*, 128, 75–91.
- Stainton, R. J. (2006). Terminological reflections of an enlightened contextualist. *Philosophy and Phenomenological Research*, 73, 460–468.
- Townsend, D., & Bever, T. (2001). *Sentence comprehension*. Cambridge, MA: MIT Press.

Part III

Semantic Values

The Disunity of Truth

Josh Dever

Abstract §§3–4 of the *Begriffsschrift* present Frege’s objections to a dominant if murky nineteenth century semantic picture. I sketch a minimalist variant of the pre-Fregean picture which escapes Frege’s criticisms by positing a thin notion of semantic content which then interacts with a multiplicity of kinds of truth to account for phenomena such as modality. After exploring several ways in which we can understand the existence of multiple truth properties, I discuss the roles of pointwise and setwise truth properties in modal logic. I argue that thinking of supertruth and determinate truth as setwise truth properties allows an understanding of supervaluationist approaches to vagueness which escapes both Williamson’s objections to and a needless metalinguistic orientation of traditional supervaluationism.

1 In Media Res: Supervaluations

A natural first move in attempting to solve the problem of constructing a semantic theory for a language containing (or discussing) vagueness is to admit a multi-valued logic, and assign to sentences ascribing a property to a borderline case of that property a ‘truth value gap’ in the guise of a third truth value. Three compelling arguments to (at least naïve formulations of) the multi-valued approach to vagueness are then:

1. Multi-valued logics are revisionary about the system of validities generated by the language. On many multi-valued systems, for example, neither instances of the law of the excluded middle nor instances of the law of non-contradiction will be valid. One elegant way of bringing out the logical difficulties is to observe that, in an n -valued logic, any sentence of the form:

$$(p_1 \leftrightarrow p_2) \vee (p_1 \leftrightarrow p_3) \vee \dots \vee (p_n \leftrightarrow p_{n+1})$$

J. Dever

Department of Philosophy, University of Texas at Austin, 1, University Station C3500, Austin, TX 78712, USA

will be valid.¹ But there seems to be no reason why such a claim should be valid, especially in light of higher-order vagueness, which will allow (e.g.) borderline cases and borderline cases of borderline cases, whose status with respect to the vague predicate should not be the same.

2. Continuing with this line of thought, multi-valued approaches to vagueness have difficulty accounting for higher-order vagueness. If the proper account of a borderline case involves the ascription of a third truth value, then the proper account of a borderline case of a borderline case should involve a fourth truth value, and hence any finite-valued system will be committed to a finite limitation on the degrees of higher-order vagueness.
3. Because of their underlying commitment to truth-functionality, multi-valued accounts cannot capture what Fine (1975) calls *penumbral truths*. If *a* and *b* are both borderline cases of tallness, then both ‘If *a* is tall then *b* is tall’ and ‘If *b* is tall then *a* is tall’ must receive the same truth value, even though if *a* is taller than *b*, the second, but not the first, should come out always to be true.

This cluster of worries motivates the introduction of supervaluational accounts of vagueness, paradigmatically in Fine (1975). Supervaluational theories center the semantics around the property of supertruth, understood as a quantification over truth at complete precisifications of the language. The resulting Kripke structure allows the recapture of all classical validities, a substantive logic of higher-order vagueness, and a formal mechanism for producing penumbral truths. There are, of course, well-known objections to the claims of supervaluationism to these formal success. Some of these objections will be discussed in Sections 10 and 11 below, but I want to begin by considering a different style of objection to supervaluationism, raised by Fodor and Lepore (1996). Fodor and Lepore suggest that supervaluationism comes in conflict with what they propose as a platitude:

(P) Conceptual truths must be respected by all classical models, including classical valuations.

Since, as they argue, it is a conceptual truth that (for example) people with certain configurations of hair on their heads are borderline cases of baldness, all classical valuations, including the classical valuations quantification over which defines supertruth, must respect this conceptual truth. But supervaluationism does not respect it, because it appeals to fully precise classical valuations, allowing borderline cases, understood as gaps in supertruth, to emerge via *disagreement* among the precise classical valuations. Roughly speaking, supervaluationism tries to buy a vague semantic account with a non-vague semantic coin, and thus cannot pay the needed price.

There are many subtle objections that might be raised to this argument, and Fodor and Lepore carefully present and examine several of them. I will return to one of these objections in Section 11 below, but let me begin with an unsubtle objection,

¹ See Williamson (1994), 112, for this objection. The objection is an application to the domain of vagueness of the Gödel-McKenzie-Dugundji objection discussed in Section 6 below.

one too naïve to have been considered by Fodor and Lepore but one which I think is helpful in uncovering some of the issues at stake here. Consider the following response to Fodor and Lepore:

Of course there are no borderline cases in the classical valuations, because the classical valuations aren't the tool in the semantic theory used to explain borderline cases. Borderline cases are explained using supertruth, and supertruth *does* respect the desired conceptual truths. Consider an analogy: determinacy of the underlying microphysics does not prevent the vagueness of the macrophysical phenomenon, even in the face of supervenience of the macro on the micro. That a determination relation holds between a non-vague set of facts and something does not prevent that something from being vague. Similarly, supertruth can be determined by the precise classical valuations without preventing supertruth from exhibiting vagueness. Worrying about the precision of the classical valuations is like worrying about the precision of the microphysical.

This response has something of the flavour of Fodor and Lepore's Reply 1 ('flouting conceptual truths is something that is done all the time in classical model theory') and also something of that of Reply 6 ('*t-in-a-classical-valuation* is my word, so I get to decide what it entails'), but tries to cut the objection off earlier than either of these.

The attempt at the early cutoff comes via a tacit attempt to set aside what I take to be the deep concern underlying the Fodor and Lepore argument: the status of the truth-conditional approach to semantics. By, in essence, claiming that supervaluationism should meet intuitive constraints only at the level of 'supertruth', not at the level of classical truth, the objector abandons the project of giving semantic explanations in terms of *truth* conditions. This move is disguised by the choice of the label 'supertruth', combined with the slogan 'truth is supertruth', but so long as supertruth is defined out of truth at a classical valuation, it is hard to see how it can simply *be* truth. To try to make supertruth the semantic coin of choice is then to devalue truth, and to set aside one important strand of semantic theorizing going back through Davidson (1984c) at least to Frege (1967).

Seen in this light, supervaluationism is one of a host of challenges, often deriving from the application of high-powered technical tools to semantic problems, to the centrality of truth in a theory of language. My goal in this essay is to explore one method of making a place for these tools at the semantic table without unseating truth at the head of the table. To achieve this goal, I want temporarily to rewind to what I take to be the birth of semantic truth-conditionalism in Frege. By examining his reasons for giving truth conditions pride of place against the background of nineteenth century logic, I will uncover a path not taken, which I will argue preserves the virtues of Frege's route but also allows an increased flexibility for giving truth-conditional accounts of, for example, supervaluationism. After further developing the details of this *disunity of truth* approach, I will at the end return to a more detailed consideration of supervaluationism and objections to it, considered in light of the lessons gained along the way. I will then close with a second application

of the disunity of truth approach, showing that it allows a reconceptualization of much work in dynamic semantics which preserves the formal achievements while avoiding the thought that dynamic semantics mandates:

the emergence of a *dynamic perspective* on natural language semantics, where the center of the stage, occupied so long by the concept of truth with respect to appropriate models, has been replaced by context change conditions, with truth conditions defined in terms of those. (Van Eijck and Kamp (1997), 181)

2 Flashback: Frege's Terminality Argument

The *Begriffsschrift* is a work of many pleasures. There are, of course, the positive pleasures, those of the accomplishments and advances Frege achieves (the stunning introduction of a full polyadic quantified logic, the complete axiomatization of the first-order fragment of that logic, the definition of the ancestral). More subtle, but no less enjoyable, are the negative pleasures – those afforded by watching Frege avoid the logical pitfalls into which so much other nineteenth century work on logic falls. At times these negative pleasures are implicit, granted by Frege's failure to do wrong things which he easily could have. Use and mention are kept quietly but efficiently distinct, names and quantified noun phrases are not folded into a single logico-syntactic category, and so on. At other times, the negative pleasures become explicit, as Frege directly addresses confusions of the existing logical tradition that he wishes to avoid. The most extended such passage occurs in § 4:

We distinguish between *universal* and *particular* judgments; this is really not a distinction between judgments but between contents. *We ought to say "a judgment with a universal content", "a judgment with a particular content"*. For these properties hold of the content even when it is *not* advanced as a judgment but as a mere proposition.

The same holds of negation. In an indirect proof we say, for example, "Suppose that the line segments AB and CD are not equal." Here the content, that the line segments AB and CD are not equal, contains a negation, but this content, though it can become a judgment, is nevertheless not advanced as a judgment. Hence the negation attaches to the content, whether this content becomes a judgment or not. I therefore regard it as more appropriate to consider negation as an adjunct of a *content that can become a judgment*.

The distinction between categoric, hypothetical, and disjunctive judgments seems to me to have only grammatical significance.

The apodictic judgment differs from the assertory in that it suggests the existence of universal grounds from which the proposition can be inferred, while in the case of the assertory one such a suggestion is lacking. By saying that a proposition is necessary I give a hint about the grounds for my judgment. *But, since this does not affect the conceptual content of the judgment, the form of the apodictic judgment has no significance for us.* (Frege (1967), § 4).

I want to begin by extracting a picture of the aspect of the pre-Fregean semantic tradition that Frege is criticizing in §4, together with a characterization of Frege's present objection to that tradition. The story will follow a well-trodden path at first, but then will strike out into less familiar territory to arrive at a non-standard characterization of pre-Fregean semantics. The first step in Frege's criticism is the drawing of the distinction between force and content. This distinction is made at best

murkily in much pre-Fregean logical work. For example, in a simmering dispute about whether O-type claims ('Some S is not P') should be understood as a species of I-type claims ('Some S is P') with a negated predicate, those opposed to such reduction tend to swing freely between two formulations:

1. Negation is a feature of an entire sentence, not just of a predicate.
2. Denial is a way of making a judgment, on equal footing with assertion.

Thus Bain (1874) says:

Some logicians have proposed to do away with the distinction between affirmative and negative by transferring the sign of negation from the copula to the predicate; 'A is not B', 'A is not-B'; 'penury is not agreeable'; 'penury is disagreeable'. There is then the appearance, but only the appearance, of making all propositions affirmative. The attempt is illusory. Affirmation and Denial belong to the very nature of things; and the distinction, instead of being concealed or disguised to make an imaginary unity, should receive the utmost prominence that the forms of language can bestow. Thus, besides being either universal or particular in quantity, a proposition is either affirmative or negative. (83–84)

There is no serious attempt here to distinguish the account as one addressed toward the contribution of negation to content and one addressed toward the contribution of negation to force.

Without the content/force distinction clearly made at all, the proper placement of that distinction certainly cannot be expected. Frege's second critical point in §4 is that the semantic contribution of logical particles and their associated operations – he is explicit about negation, universality and particularity (existentiality), and suggests the same view for the conditional and the disjunction – belongs to content, rather than to force. In §4 the argument is an *overdetermination* argument:

Argument From Overdetermination: No utterance can possess more than one force (so, for example, no utterance can be both an assertion and a command). But logical particles can be used in sentences containing explicit force specifications, or in sentences that can optionally be uttered under various forces. So logical particles cannot carry force specifications as their semantic content, else force overdetermination would result.

Negation cannot rely on a force of denial, because negations can be used (and used effectively) in propositions which are not denied but are instead supposed. Were negations to introduce a force of denial, then:

Suppose that the line segments AB and CD were not equal.

would become a denial, rather than (as it actually is) a supposition. Elsewhere the point is put in terms of *embeddings*:

Argument From Embeddings: A sentence ϕ can be embedded in a context such that no speech act is performed with ϕ . However, if logical particles in ϕ were to contribute force to the semantic interpretation of ϕ , the embedded

occurrence of ϕ would result in the performance of a speech act. So logical particles cannot carry force specifications as their semantic content.

Thus in Frege (1997), Frege observes:

The thought contained in the sentence:

'If the accused was in Rome at the time of the deed, he did not commit the murder.'

may be acknowledged to be true by someone who does not know if the accused was in Rome at the time of the deed nor if he committed the murder. Of the two compound thoughts contained in the whole, neither the antecedent nor the consequent is being uttered assertively when the whole is presented as true. (145)

Call a semantic feature *terminal* if that feature appears mandatorily only as a feature of entire utterances, and is prohibited (either *de jure* or by its intrinsic nature) from appearing as a semantic feature of an expression embedded in a larger expression. The full generalization of Frege's critique of force-based accounts of logical particles is then the following terminality principle:

Terminality: An account of the semantic contribution of an expression must permit of integration into a compositional theory for the entire language. Thus if ϵ is an expression which can occur in contexts $\Sigma(\phi[\epsilon])$ for some ϵ -containing sentence ϕ and some embedding context Σ , the semantic contribution of ϵ must be entirely non-terminal.

Coupling this principle with the following semantic hypothesis:

Terminality of Force: Force is essentially a terminal semantic feature we cannot coherently assign force to a sentence which occurs as an embedded proper part of a larger utterance.

leads to Frege's thesis that force-based accounts of the semantics of logical particles must be rejected. It is in the terminality argument, I suggest, that we see the birth of truth-conditional semantics.

3 Pre-Fregean Semantic Maximalism and Minimalism

Frege's terminality argument is a forceful one, and I think it's safe to say that it has carried the day in the philosophical tradition. The argument has life in it still, continuing to raise problems for semantic theories, such as emotivism or various brands of fictionalism, whose most natural formulations require certain expressions (moral terms, modal operators, mathematical expressions, etc.) to make essentially terminal semantic contributions. However, there are some linguistic phenomena that seem to run contrary to it, and some open routes of resistance to it. Two escape routes can be distinguished. First, some examples suggest that, *despite* the terminality of force, force is indeed carried as a semantic contribution of embedded expressions. Thus consider conditional imperatives:

If someone asks you for the salad fork, give it to them!

and assertion-loaded questions:

I only wanted to do something nice for you, so why did you have to be so rude?

Exactly how best to describe, let alone theoretically account for, the phenomena here is not straightforward – perhaps two speech acts of two different forces (assertion, interrogation) occur using one sentence; perhaps a single compound force (*hypothetical command*, etc.). But some description is tempting on which such cases amount to counterexamples to the terminality principle. Alternatively, the terminality of force claim can either be denied or, more commonly, evaded. Evasive action is available through ‘value-loading’ – by assigning to lower-level constituents semantic values consisting of instructions for the production of certain semantic effects at higher levels (as in the treatment of comparison-class sensitive scalar adjectives as maps from matrix \bar{N} to modified \bar{N}), embedded particles can carry force-relevant information that fully emerges only at the terminal constructive stage. Alternatively, a kind of ‘internal accusative’ to the desired force can be introduced as a composable semantic value – a value which attempts to provide a value, capable of compositional interaction with other semantic values, which mirrors certain crucial features of the target force. Truth can perhaps be thought of as an internal accusative of assertion,² and one of the innovations emerging from the interaction of moral non-cognitivism and minimalism about truth has been the thought that truth can, in fact, function as a single internal accusative to a wider range of forces. Much work in formal semantics treating of non-assertoric constructions can be thought of in this way, as when, for example, questions are treated as carrying sets of answers (sets of possible worlds) as semantic values.

Frege’s reaction to the terminality argument is to center his semantics around a non-terminal value, and his choice of that value is truth. The choice of truth is articulated and motivated in §3 of the *Begriffsschrift*, where Frege announces that in his concept language:

Everything necessary for a correct inference is fully expressed but what is not necessary usually is not indicated. (Frege (1967), §3)

Thus Frege’s clear separation of content from judgment allows him to see that only content, construed truth-theoretically, provides a suitable domain for the semantic contribution of logical particles. So far, this is the standard reading of Frege’s reasoning in §4. From this point, I want to venture into an attempt to extract a picture of the pre-Fregean semantic tradition to which Frege is reacting with this line of reasoning. The most obvious way to take that tradition is as supposing a kind of *semantic maximalism*:

² I don’t particularly *recommend* this way of thinking about truth. If assertion is taken to be subject to an epistemically-specified regulative norm (such as ‘assert only that which you *know*’), truth construed as the internal accusative for assertion threatens to push one toward a substructural logic. If, of course, the regulative norm for assertion is truth-based (‘assert only that which is true’), then truth cannot be analyzed or illuminated by being characterized as assertion’s internal accusative.

PFS ↑: The *content* of an expression contains specification of both truth-conditional and speech-act force information. Thus ‘and’ and ‘but’ differ in content, as do ‘The cat is on the mat’ and ‘Is the cat on the mat?’

PFS ↑ can then be combined with force-based analyses of a variety of logical particles. Typical in this respect is the assumption that quantificational specification, negation polarity, molecular status, and modal status can all be understood in terms of force³ – that an utterance can be made with the force of universality, or denial, or hypotheticality, or apodeicticity. When Devey (1854) sums up the standard pre-Fregean taxonomy of the subject-predicate ‘S–P’ proposition by saying:

Regarding the two terms of a proposition as expressing its matter, and the act by which they are confirmed or separated as relating to its form, all that the logician has to advance upon this portion of the science may be treated under these two heads. (Devey (1854), 76)

he captures the **PFS** ↑ thought that it is force, determining the speech act performed, which differentiates kinds of propositions.

In addition to the maximalist strand of the pre-Fregean tradition, it also seems to me that the tradition has its *minimalist* moments, embodying a thought of the form:

PFS ↓: The basic structure of a proposition is the ‘S–P’ combination of subject and predicate, and the various modifications of quantificational specification, negation polarity, molecular status, and modal status are not alterations to the content of the proposition

The logical specifications then are understood as the result of decisions to perform, with the matrix ‘S–P’ content, speech acts of varying force. Much of what pre-Fregean logicians say appears to run directly contrary to this diagnosis. A long sequence of logic texts marches across the English landscape of the period, all following the same basic pattern. The text begins with a discussion of terms, proceeds to a discussion of propositions, and concludes with a categorization of the forms of the syllogism, together with techniques of transformation. The section on the proposition then proceeds to enumerate various *types of proposition*, such as universal and particular, affirmative and negative. Thus:

The same subject and predicate may be combined in four several manners, and so be formed into four different kinds of Propositions. viz.

Universal affirmative, as, All Men are wise.

Universal negative, as, No Men are wise.

Particular affirmative, as, Some Men are wise.

Particular negative, as, Some Men are not wise. (Bentham (1773), 40)

Especially in the Boole-influenced tradition, much of the language mirrors that used by Frege in his downplaying of the categorical/hypothetical/disjunctive distinction:

³ Quantity, quality, relation, and mood, in the classical terminology. See, for example, Lotze (1884) for a clear articulation of these categories, although Lotze is, not surprisingly, much closer to Frege than the nineteenth century norm in his inclination to identify a force-free content underlying these logical differentia.

There is no need to dwell on the usual matters given as to the distinction of universal and particular, or of positive and negative. But, I think it cannot be denied, that the distinctions may, for logical purposes, be considered as accidents of language. (Morgan (1966), 3)⁴

However, taxonomic passages such as Bentham's must be read in light of the generic pre-Fregean conflation of force and content. The pre-Fregean semantic tradition is sufficiently murky that it is hazardous to attempt straightforward interpretation of its terminology.

In the absence of a clearly-drawn use-mention line, the pre-Fregean use of the word 'proposition' becomes a prime offender, tending to swing between the syntactic expression, the speech act, and the semantic content. Thus we find assertions such as:

1. A proposition is an *act of judgment expressed in language*, and consists of three parts, a *subject*, a *predicate* and a *copula*. (Coppee (1857), 63)
2. A Proposition is a sentence which either affirms or denies, as, All men are mortal. (Boole (1952), 64)
3. Judgment is the comparing together in the mind two of the notions (or ideas) which are the objects of Apprehension, whether complex or incomplex, and pronouncing that they *agree* or *disagree* with each other: (or that one of them *belongs* or does not belong to the other.) Judgment, therefore, is either *affirmative* or *negative*. . . . An act of *apprehension* expressed in language, is called a *term*; an act of *judgment*, a *proposition*. . . . [A proposition] must not be *ambiguous* (for that which has more than one *meaning*, is in reality *several* propositions. (Whately (1825), 55, 76)

There is, I think, no disentangling the elements of these views. Frege is, in a sense, *inventing* the notion of a proposition in §§3–4 of the *Begriffsschrift*, via the determination that truth will be the recursively-combined carrier of a distinguished region of linguistic potentiality identified as content.

That said, there are moments at which the lurking thread of minimalism pre-Frege rises to the surface. The understanding of modality plays a key role here. The mode of an utterance is generally taken not to contribute to the content of an utterance, but rather to make a secondary declaration about that content. Thus in a typical such passage Bain (1874) says:

Propositions qualified as *Necessary*, first give an affirmation, and secondarily declare that such affirmation belongs to the class of necessary truths, whatever these may be. (Bain (1874), 99)

⁴ Thus, one suspects, Venn's failure in his review of the *Begriffsschrift* to comprehend the radicality of Frege's logical moves:

The merits he claims as novel for his own method are common to every symbolic method. . . . He calls attention to the fact that, on his scheme, the distinction, so important in grammar and on the predication-view of ordinary logic, between subject and predicate loses all its significance, that hypothetical and disjunctive propositions become equivalent to categorical, and so on; all these being points which must have forced themselves upon the attention of those who have studied this development of Logic. (Venn (1880), 297)

Such passages follow a Kantian line:

The *modality* of judgments is a quite peculiar function. Its distinguishing characteristic is that it contributes nothing to the content of the judgment (for, besides quantity, quality, and relation, there is nothing that constitutes the content of a judgment), but concerns only the value of the copula in relation to thought in general. (Kant (1929), 106)

This minimalism regarding modality is then combined with a frequent attempt to treat modality as a version of quantificational specification:

A *Modal* proposition may be stated as a *pure* one, by *attaching the mode to one of the Terms*: and the Proposition will in all respects fall under the foregoing rules. . . . E.g., “man is *necessarily* mortal:” is the same as “*all* men are mortal.” . . . Indeed every sign (or universality or particularity) may be considered as a *Mode*. (Whatley (1825), 108)

The result is a pressure for quantificational specification as well to drift out of content. Since the desire to unify logic under the syllogistic system produces pressure to reduce all logical modification to the case of quantification, the tendency toward **PFS**↓ results.

4 The Disunity of Truth

We have arrived at the following point. Frege’s innovation in §§3–4 of the *Begriffsschrift* is the isolation of the notion of *propositional content* from a murky pre-Fregean stew of broadly semantic phenomena. On one reading (**PFS**↑) of the historical record, the Fregean innovation occurs by way of *whittling down* – the pre-Fregean stew was too rich, and Frege had to isolate the specifically semantic ingredients in it. The isolation then proceeds via the distinction between the content stroke and the judgment stroke – a move which, in turn, sees the elevation of truth to a central location in the semantic theory. Truth, unlike assertoric force, is internally localized, and hence can enter into a calculus of compositional determination. But I’ve tried to suggest that the very muddying of the waters that Frege was attempting to clarify makes available another interpretation of the state of play. On this alternative interpretation (**PFS**↓), the pre-Fregean conception was not too rich, but (if anything) too impoverished. There is a sense in which not only force, but all logical operations are stripped from ‘the proposition’, leaving a bare matrix of the form ‘S–P’ which can be asserted in various ways.

Of course, modality, quantificational specification, negation-polarity, and other features must go *somewhere*, so if not in content, where? Again, the murkiness of the pre-Fregean tradition prohibits a clear answer, but one perspective that emerges from that tradition is that all of these features are assimilated to force, and hence become terminal contributions to a speech act. Call this version of minimalism **PFS**↓ +**F**. **PFS**↓ +**F** leads to a different interpretation of Frege’s moves in §§3–4. Now, instead of whittling down, Frege is building up – incorporating enough into content to allow for a compositional semantic calculus. A truth-theoretic picture of logical particles is thus developed. It is hence no surprise that Frege stands largely alone in the nineteenth century in providing a clear truth-conditional interpretation of the propositional connectives and quantifiers. Frege’s decision to set

aside modality – to which we'll return shortly – thus becomes predictable, since the requisite truth-conditional contribution sufficing to explain combination with modal operators is hard to come by.

I want to grant the efficacy of the Fregean terminality argument against both **PFS**↑ and **PFS**↓ +**F**, but point to another way of taking the minimalist conception that opens conceptual space for a non-Fregean understanding of the structure of a semantic theory, an understanding which I'll then argue sheds light on some recent issues in semantics. The alternative construal of the minimalist conception meets Frege half-way by granting the significance of truth as an internal carrier of semantic value, but rejects Frege's *truth monism* – the thought that there is only a *single* property going under the name of, and playing the theoretical role of, truth. The alternative picture – call it **PFS**↓ +**T** – holds that there are multiple properties going under the general heading of 'truth'. To avoid overworking the word 'truth', I will refer to a *veracity feature*, or **VF**. Thus **PFS**↓ +**T** holds, via the minimalism, that there is a content common to logical modifications of the base 'S-P', but (unlike **PFS**↓ +**F**) holds that this sole content can bear multiple veracity features, and takes the role of the logical particles to be in identifying which **VF** is currently relevant. **PFS**↓ +**T** is a specific instance of a more general hypothesis:

Disunity of Truth Hypothesis (DOT): There are multiple veracity features, and a single content can be assessed for more than one such feature.

One way to gain entrance to **DOT** is to begin with the idea of assertion as a norm-governed speech act, and then to reify the telos of that act and label it 'truth', without concern for the nature of the telos. We thus treat truth as a sort of 'internal accusative' to assertion – truths are asserted, just as appearances are seen and sounds are heard, as a terminological convenience which provides an object via which the act type can be discussed without overly much concern about the actual causal and normative structure of the act. We can then produce corresponding internal accusatives for the range of speech acts that **PFS**↓ +**F** appends to the minimal 'S-P' content, and identify those accusatives as 'modes of truth'.⁵ I'll argue below that this isn't the best way to think about the disunity of truth, but it may provide a useful initial way of understanding what it might be for more than one property to be a **VF**.

Consider an utterance of 'All men are mortal'. **PFS**↑ has it that the content of this utterance includes a force of universality; **PFS**↓ +**F** that the content has no such universal force and thus is shared with 'some men are mortal', but that force must then be added to achieve the particular universality. The Fregean picture agrees with the maximalist picture in locating the differentiation in the content, but disagrees in taking the differentiation to be truth-conditional, rather than force-based. **PFS**↓ +**T** agrees with Frege that the phenomenon to be explained (the universality) is to be accounted for truth-conditionally, but agrees with the minimalist that the content is common between the universal and the particular utterance. Room is made for this

⁵ So, perhaps, answers are 'truth values' for interrogatives and satisfactions for imperatives.

position by allowing that the minimalist content (what we might call ‘men being mortal’) admits of (at least) two kinds of truth: *universal truth* and *existential truth*.

Thus ends the historical rewind. **PFS**↓ +**T** and the **DOT** represent the path not taken, the alternative to the Fregean enthronement of (monistic) truth in pride of place in semantic theories. I now turn to the exploration of that path. I begin by developing a specific case of the **DOT** proposal in more detail. A bolder version of the present paper would perhaps argue that for modes of truth corresponding to universality, or particularity, or conditionality, but I’ll be more cautious, and limit myself to exploring the treatment of various styles of modality within a disunity of truth framework. My chosen instance of the **DOT** proposal is thus less anti-Fregean than it might have been, since Frege at the end of §4 of the *Begriffsschrift* appears happy to join the **PFS**↓ position in leaving the impact of modals outside of content. Bradley (1883), on the other hand, inveighs against the application to modality of the approach I intend to explore:

We must begin by stating an erroneous view. Modality may be supposed to affect the assertion in its formal character, and without regard to that which is asserted. We may take for instance a content S–P, not yet asserted, and may claim for modality the power of affirming this content S–P, unaltered and unqualified, in several ways. S–P, it is supposed, may be asserted, for instance, either simply or problematically, or apodeiktically, and may yet remain throughout S–P: and thus, though the content is unmodified, the assertion is modal.

There are, of course, worries about whether **PFS**↓ +**T** is just a terminological variant on the Fregean position. Such worries are inevitable when engaged in carving up the broad phenomenon of linguistic communication into fields of responsibility, and final adjudication of them is typically a delicate matter of weighing integration of various taxa into broader theoretical concerns. I will try to discharge the burden of the worries in three ways. First, the development of the details of the **DOT** treatment of modality will, I hope, highlight features of the model theory of modality which do not emerge so naturally on the traditional Fregean model. Making this point is the burden of Section 8 below. Second, I turn in Section 9 to the question of what makes a property a genuine *truth* property; the discussion in this section should help demonstrate that more than terminology is at stake. Third, I return in Sections 10 and following to our opening issues regarding supervaluationism, and show that a **DOT** perspective on supervaluation theory sheds valuable light on the contemporary discussion.

5 Main Plot: Necessity is a Mode of Truth

We begin with the thought that sentences can have two different *kinds* of truth value. One of these is the familiar truth – what I’ll call *truth*_@. The other is the property of necessary truth – what I’ll call *truth*_□. Truth_□ shouldn’t be read as an adjectively-modified version of truth_@ – to be true_□ is not to be such that it

is necessary that one be true_@.⁶ How it *should* be read, of course, is a substantial question that we'll have to work out as we proceed. Truth_□ also is not just another view about what truth is – it is not, for example, another competitor in the field of correspondence, coherence, and minimalist theories of truth. Truth_□ is not even extensionally equivalent to truth_@, so some sentences will differ in their truth_□ and truth_@ values. The present view is thus not an instance of what Lynch (2004) has called *alethic pluralism*.⁷

Question: Hasn't this already been tried, perhaps first by MacColl (1897) and then more carefully and elaborately by Łukasiewicz (1967a) and Łukasiewicz (1967b)? And wasn't the approach shown to fail via the line of objection developed by Godel (1986), J.C.C. McKenzie, and Dugundji (1940):

GMD Objection: Let \mathbb{L} be an n -valued logic, $n \in \omega$. Then, for distinct atoms ϕ_1, \dots, ϕ_n :

$$\Theta = \bigvee_{i,j \leq n+1, i \neq j} \Box(\phi_i \leftrightarrow \phi_j)$$

is a theorem of \mathbb{L} . However, Θ is not a theorem of any normal modal system **S5** or weaker, so \mathbb{L} cannot give the logic of any such system.

More expansively: Suppose a modal logic is given a multi-valued truth-functional semantics, with some privileged truth value τ representing necessity, so that $\Box\phi$ is true (or one of a set of designated truth values) just in case ϕ has truth value τ . Suppose also that a biconditional between two sentences of the same truth value has itself the truth value τ . (This is necessary in order to account for the truth of $\Box(\phi \Rightarrow \phi)$.) Then, if there are n truth values, and we form biconditionals between $n + 1$ distinct sentences, some two of those sentences must have the same truth value, and hence their biconditional must have truth value τ and its necessitation must be true. But in (for example) **S5**, a disjunction of these necessitated biconditionals can be made false simply by providing 2^{n+1} worlds, which then provide all possible assignments of truth and falsity to the $n + 1$ sentences. Thus the multi-valued system produces validities not produced by **S5** (and *a fortiori* not produced by any weaker system), and does not suffice as a semantics of modal logic.

Answer: No. The present proposal is importantly different from that of MacColl and Łukasiewicz, although the GMD objection will turn out to play an important role here. MacColl and Łukasiewicz both attempt to create a multi-valued system in which the standard features of truth and falsity are *supplemented* by additional modal veracity features – MacColl's values of ϵ (certainty), η (impossibility) and θ

⁶ Of course, eventually an adequate account must *yield* that a sentence ϕ is true_□ just in case it is necessary that ϕ is true_@. The present thought is, however, that this result will be a downstream consequence of an independent understanding of true_□, rather than the very definition of that would-be **VF**.

⁷ See the end of Section 9 below for more on this point.

(variability), or Łukasiewicz's value of $1/2$. They are thus committed to a single semantic calculus, on which the standard truth values **T** and **F** can interact with the novel modal truth values. **DOT**, on the other hand, requires genuine disunity. It's a multi-tracked system – we can evaluate a sentence correctly as being $\text{true}_@$, and we can also, and independently, evaluate the very same sentence correctly as being true_\square . And the two truth values don't interact. Thus there is no commitment to (for example) determining the truth value of $\phi\psi$ when ϕ is $\text{true}_@$ and ψ is false_\square . As a result, the view is not immediately subject to the GMD objection – the massive disjunction Θ does not become a validity by way of exhausting the supply of truth values, since the various modes of truth don't interact via the biconditional. However, the GMD objection will return shortly to haunt the disunity proposal in another way.

I now give a toy example. Let **L** be a simple propositional language with a standard collection of truth-functional connectives. A model assigns to each atom a pair of truth values – one $\text{true}_@$ value, and one true_\square value. Assignments are not required to be total – sentences can lack either a $\text{true}_@$ value or a true_\square value (although only the lack of true_\square value will feature in the discussion). **L** carries with it two consequence relations $\vDash_@$ and \vDash_\square , understood in the usual way via preservation of $\text{true}_@$ or true_\square respectively. The consequence relations can be conceived indifferently as constraining or flowing out of the notion of truth in a model. In general, call a **VF** T_X *regular* if it meets the following conditions:

1. \vDash_X is superclassical.
2. \vDash_X is compact.
3. In no model is any sentence assigned both true_X and false_X .⁸
4. In any model, if a sentence ϕ is neither true_X nor false_X , $\neg\phi$ is also neither true_X nor false_X .

True_\square is regular, and I will assume throughout that all **VF**'s under discussion are regular. $\text{True}_@$ -in-a-model can be defined in the usual way, but true_\square -in-a-model presents a challenge. When ϕ and ψ are both false_\square , we cannot predict the true_\square value of $\phi \vee \psi$. If ψ is of the form $\neg\phi$, then $\phi \vee \psi$ is true_\square , but if ψ is a (modally) unrelated atom, then $\phi \vee \psi$ will be false_\square .

The toy language lacks a compositional theory of true_\square values, because the sentential connectives are not true_\square -functional, and the language so far is purely extensional (albeit both $\text{true}_@$ -extensional and true_\square -extensional). The problem is strictly analogous to the GMD objection to finite multi-valued analyses of modality – in both cases, the language seeks to make more $\text{true}(\square)$ distinctions among certain molecular expressions than are made available by the given range of true_\square values. For the general form of the GMD objection, this requires showing that the modal language allows an infinite number of such distinctions (hence, the

⁸This constraint follows from the superclassicality of \vDash_X if we take 'truth preservation' to block the move from true_X to falsehood $_X$. In the case of true_\square , the constraint can be thought of as a realization of the **D** axiom.

underderivability of Θ in normal systems **S5** and weaker); for the specific case here, since it is set in a bivalent \square language, it is necessary only to show that three-fold disjunction is not valid.⁹

Despite the appearance of an analogous problem, the moral is not the same. The MacColl-Łukasiewicz project was to treat modal logic as a variant of propositional logic, and hence to introduce object-level modal operators semantically explained on a par with the usual truth-functional operators. The unavailability of a finite range of truth values allowing appropriate truth tables is thus fatal. But my goal is more humble – merely to recognize a novel range of truth(\square) values, and thus allow two modes of veracity-evaluation of sentences which themselves are free of overt modal language. I can thus respond to the GMD objection by conceding that the toy language, at least on its truth \square side, is not compositional. What to make of that concession will then depend on what, in general, to make of non-compositional semantic phenomena.

What do we learn when we learn that a language has a non-compositional meaning theory? Or, put another way, why have people taken compositionality to be a desirable feature of semantic theories? What danger is compositionality skirting or averting? There are a number of standard answers to that question, none of which I find very persuasive.¹⁰ I'll thus briefly sketch my own preferred answer, and then consider its consequences for the current non-compositionality with respect to truth \square . Compositionality is the principle that the meaning of a complex expression be a function of the meanings of its parts and their mode of syntactic arrangements.¹¹ Compositionality enforces a supervenience of sentential semantics on lexical semantics; its significance thus depends on the significance of lexical semantics. Why do we think that individual words should have meanings? Lexical meanings are not *necessary* to explain sentential meanings – learnability arguments will (perhaps) establish that there be some finite basis and recursive/computable pathway from that basis to sentential meanings, but such arguments will not themselves impose the further constraints that the information in the basis be semantic (rather than syntactic, phonological, contextual, meteorological, etc.) in nature, or that the articulation of the basis follow the joints of syntax or conform to any convenient conceptual breakdown. My suggestion is that the role of lexical semantics is to extract various commitments (conceptual and ontological) which lie implicit in the global linguistic practice. Our patterns of assent and dissent determine (for example) our ontological commitments, but such commitments are not overt in sentential semantics. We identify, and care about, a category of nominal term semantics because of the role of such terms and their meanings in exposing our previously inchoate ontological commitments. But for

⁹ Hence the result extends slightly further, blocking a truth-functional analysis for logics **S5-Alt**₂ and weaker.

¹⁰ See both Szabo (2000) and Dever (2006) for some of the shortcomings of the standard answers, and Dever (2003) for some problems with the novel answer that Szabo defends.

¹¹ More or less; for refinements, see the discussion in Dever (2006).

lexical semantics to play this role, features of the global linguistic practice must be properly reflected into the lexical semantics. Compositionality enforces this reflection – absent compositionality, semantic features of the linguistic practice can ‘float free’ (for example, the epistemological commitment of Fregean senses can emerge (noncompositionally) only at the sentential level, and hence fail of manifestation in our ontological commitments), not appearing anywhere in the lexicon.

Exempli Gratia: At least since the *Gorgias*, deontic debates have turned on whether there is a single notion of *good*, or whether *good* must be understood as relativized to ends (or objects or agents). Semantic considerations are unlikely to settle this dispute, of course, but certain linguistic data at least appear to pose a problem for the monistic account of the good. Thus consider Geach (1956)’s linguistic considerations. If ‘good’ is unrelativized, then ‘Susanne is a good dancer’ should entail ‘Susanne is good’, so ‘Susanne is a good dancer and Susanne is a pianist’ should entail ‘Susanne is a good pianist’.¹² The entailment, of course, is not in fact valid. Assuming compositionality, the semantic behaviour which underwrites the invalidity of the argument needs to be traced back to the semantics of the lexical items involved, which then at least strongly suggests a view on which ‘good’ contributes a function from argument noun phrase to a characterization of goodness *for that object type*. Without compositionality, however, there is no need for the meanings of ‘good dancer’ and ‘good pianist’ to be built up out of the meanings of ‘good’, ‘dancer’, and ‘pianist’, and hence no pressure to introduce a relativized semantics for ‘good’ (the relativizing behaviour can appear as an emergent property of the phrases). The result seems to be a refusal of the semantic theory to grapple properly with legitimate questions about the property or properties picked out by adjectives.

If this picture is correct, then what we learn from the non-compositionality of truth \square is that the portion of the semantic theory which describes the distribution of truth \square values does not exhaust the commitments – especially conceptual and ontological commitments – deriving from the linguistic practice associated with the use of truth \square values of sentences. That is, however, no fatal objection to the use of truth \square values in semantic theorizing.¹³ Of course, the distribution of truth \square values must receive some explanation, so there is pressure to *supplement* the picture thus far given with further information (and lexico-semantic information, if we wish to be responsive in particular to the pressure for compositionality) which allows computation of truth \square values.

¹² This version of Geach’s point is taken from Szabo (2000).

¹³ We also know that, once we leave the extensional realm, there is no compositional treatment of truth $\textcircled{\text{a}}$ values either, and hence that these values cannot by themselves reveal our ontological and conceptual commitments.

6 Modality and ‘Possible Worlds’

Here’s a sketch of the required supplementation, following a familiar road. By the Lindenbaum lemma, we know that any consistent \Box set of sentences can be extended to a maximal consistent \Box set.¹⁴ Given a particular model (that is, a particular distribution of truth \Box values), we select a subset of the collection of maximal consistent sets by taking all sentences true \Box in the model and then, for each sentence neither true \Box nor false \Box , adding it to the base set and taking all maximal consistent \Box expansions of that collection. We then, given a particular maximal consistent \Box set w , say that ϕ is true $_w$ if and only if ϕ is a member of w . The resulting feature true $_w$ then allows explanation of the distribution of truth \Box : we trivially obtain the results:

- ϕ is true \Box if and only if ϕ is true $_w$ for all w .
 ϕ is false \Box if and only if ϕ is false $_w$ for all w .¹⁵

So, not surprisingly, a standard Kripke semantics for the modal dimension of the language **L** emerges. But two points should be emphasized, one formal and one philosophical:

1. One might worry about the apparent ability to construct a Kripke semantics, given that we started with extremely minimal assumptions about the distribution of truth \Box values in models of **L**. Why think that we’ve done enough to ensure the minimal floor of the **K** axiom and **NEC** rule, let alone the full **S5** system we should expect given the absence of an accessibility relation?

The key fact here is that **L** contains no overt modal operators, and hence can express neither iterated modalities nor molecular combinations of unary modal claims. This limitation makes (what is available of) both **K** and **NEC** quite innocent. **NEC**, in this context, imposes only the requirement that the consequence \Box relation be superclassical, which I’ve taken as an assumption. All we can require is that those claims which are always true $@$ also be always true \Box . Since their truth \Box does not show up in a syntactically distinct form (i.e., in the form of $\Box\phi$), there is no *further* question about the universal truth ($@$ or \Box) of the necessary truth of the classical theorems. All we could observe is the triviality that, if a claim is always true \Box , it is always true \Box . **K**, of course, cannot even be formulated in the current context. The best we can do is an inferential version of it, requiring that $p, p \rightarrow q \vdash_{\Box} q$. Again, this will follow from the superclassicality of the consequence \Box relation. But the claim falls considerably short of the full strength of **K**, given the limitation to non-modal substitution instances.

¹⁴ The applicability of Lindenbaum’s lemma follows from the regularity of truth \Box . Superclassicality suffices to guarantee that at most one of $\phi, \neg\phi$ can consistently \Box be added to a consistent \Box set.

¹⁵ The right-to-left direction of both conditionals requires the assumption that if ϕ is neither true \Box nor false \Box , then $\neg\phi$ also is neither true \Box nor false \Box . This assumption is guaranteed by the regularity of truth \Box .

2. I have used truth_w values to derive the distribution of truth_\square values, but this should not be taken as imputing a conceptual priority to truth_w values. At this point in the development, the ‘possible worlds’ introduced, and the associated features truth_w , are purely theory-internal devices, with no independent route to comprehension. Any understanding we have of what it means for something to be true_w comes by way of a prior understanding of what it means for something to be true_\square , rather than the other way around. For comparison, consider the status of Kripke models for intuitionistic logic, from the perspective of the constructivist. The constructivist will reject the interpretation of evaluation points as states of knowledge in a process of knowledge growth, viewing such an interpretation as unduly classical (the constructivist takes the entire Kripke model to deliver a single knowledge state). He will, instead, take the evaluation points to be conceptually parasitic on an independently understood notion of constructivist truth.

Truth_w values were introduced to achieve compositionality, and compositionality was sought to achieve ontological and conceptual honesty in our semantic theory. So what have we learned from this honesty? Lewis, famously, takes the appeal to truth_w to entail a commitment to possible worlds. But at least in the current context, such a move is too quick for two reasons. First, the ‘worlds’ appealed to are mere formal devices, and we have no reason yet to associate them with anything that, intuitively, we might be inclined to call a world. Second, the current thought is that the lexical semantic values, and in particular the semantic values of referential expressions, are the bearers of ontological commitment (in a properly compositional semantic theory) – were we to expand the current propositional language to a quantified one, we would find no referential expressions picking out worlds, or even mere possibilities. Here we perhaps see an indirect argument for direct reference: the distinction between directly referential and quantificational denotational terms such a theory provides allows us to maintain ontological caution by restricting the assignment of referents to actuals, while allowing denotational identification of possibilities.¹⁶

7 The Emergence of Modal Language

I have posited a language whose sentences can systematically be evaluated in two different ways – either for $\text{truth}_\text{@}$ or for truth_\square . Such a language allows a certain amount of modal discourse, but falls short of the full range of such discourse permitted when modality is represented by object-language modal operators creating sentences of modal *content* evaluated univocally for $\text{truth}_\text{@}$. I now want to tell a brief just-so story about the transition from the **DOT** picture I’ve sketched to the standard unity picture.

¹⁶ I set aside here worries about, e.g., pronouns under modal subordination, and other facets of the generally Meinongian tendencies of natural language reference.

Consider a claim which admits of multiple forces, such as:

Open that box and you'll get a surprise.

This claim can be read (at least) as carrying the force of (i) a request, together with a conditional promise, (ii) a threat, or (iii) an assertion. It's tempting, although not mandatory, to take the claim to involve a single content which, when subject to different forces, gives rise to the various readings (rather than positing multiple *contents* corresponding to the various readings). Such a view would give a contemporary realization of the **PFS**↓ +**F** position. Notice, however, that the claim can be disambiguated by the addition of various force-specifying expressions:

1. Open that box, please, and you'll get a surprise. [Enforced 'request' reading]
2. Open that box, I'm warning you, and you'll get a surprise. [Enforced threat]
3. Open that box, I'm told, and you'll get a surprise. [Enforced assertion.]

Suppose we posit similar expressions for the **DOT** analysis of modality – expressions which mandate, or at least suggest, that a claim be received with an eye to one of its many truth-values. Thus the addition of 'necessarily' could serve to focus attention on the truth_□ value of an assertion, rather than the truth_@ value. Then compare the following two pictures:

1. On the **DOT** picture, interpret 'Necessarily, ϕ ' as involving the content *that* ϕ , and focusing on the truth_□ conditions of that content.
2. On the classical picture, interpret 'Necessarily, ϕ ' as involving the content *that* $\Box\phi$, and focusing on the (only available) truth_@ conditions of that content.

Assuming that truth_□ values are distributed to reflect our modal judgements, these two pictures will coincide in their final evaluation of 'Necessarily, ϕ '.

Now suppose there arises a desire to assign truth_@ conditions uniformly, despite the presence of the truth_□-mandating expression 'necessarily'. The obvious move is to take 'Necessarily, ϕ ' to be true_@ whenever the semantic theory assigns the value of true_□ to ϕ . The coercion of truth_□ conditions to truth_@ conditions thus gives rise to an object-language modal operator. This approach requires that the dictates of the semantic theory receive truth_@ conditions. A more general question then arises of which **VF**'s we seek when we do semantic theorizing. There are deep issues here, but I'll skirt the depths by supposing, for the modes of truth which will interest us here, that semantic claims are uniform in their evaluation across all modes. Thus, the claim ' ϕ is true_@' will be both true_@ and true_□. When we step into the depths, there emerges the possibility of wholly ineffable semantic claims – semantic theories about truth_X, for some *X*, which are themselves only true_X, or only interestingly patterned along the dimension of truth_X.

The availability of object-language modal operators, and hence distinctively modal content, can then develop in three stages. At the first stage, there is mere communicative act duplication, as acts previously available via the truth_□ evaluation of ϕ become reduplicatively available via the truth_@ evaluation of 'Necessarily, ϕ '. At the second stage, molecular combinations of modal claims become available. In

the disunity picture, modal combinations are unavailable – we cannot, for example, express the negation of a necessity claim, because $\neg\phi$ is either evaluated for $\text{truth}_@$, in which case it is wholly non-modal, or it is evaluated for truth_\square , in which case the negation effectively takes narrow scope with respect to the modal. But $\neg\square\phi$ can be evaluated for $\text{truth}_@$ once \square is added and forced into the possession of $\text{truth}_@$ -conditional content, since negation can then operate in a straightforward way on the $\text{truth}_@$ conditions of $\square\phi$.¹⁷ In the third phase, iterated modalities become available. This third phase, however, is both less natural in genetic origin and more difficult in implementation. There is no obvious reason why iterations of the modality-forcing particle should be licensed, even once that particle is coerced into the $\text{truth}_@$ scheme. And even if such iterations are licensed, the story given here provides no obvious route to the attribution of content. The attribution of $\text{truth}_@$ conditions allows molecular combinations, because molecular combinations involve logical operators sensitive to $\text{truth}_@$ conditions, provided by the coercion to \square . But a further application of \square seems to refer us back to the truth_\square arena, and there's no good reason why that arena will provide what is needed. The most obvious approach is to treat $\square\square\phi$ as asserting the truth_\square of $\square\phi$, which will in turn assert the truth_\square of the claim that ϕ is true_\square . But, given the earlier commitment to the $\text{truth}_@$ and truth_\square of all semantic claim, the result will be the triviality of iterated modalities (hence, in the current context, an **S5** system).

Before turning to the question of semantic implementation of the iterated modalities, let me extract from this just-so story a tentative prediction:

The Iteration Hypothesis: A semantic phenomenon which is best understood using a **DOT** picture will often be characterized by a difficulty in interpretation of iteration of the operators characteristic of that phenomenon.

The direction of the current investigation is to suggest that certain aspects of natural language may be best understood via the **DOT** – that we have been misled by thinking that these aspects involve a distinctive kind of content subject to the usual $\text{truth}_@$ evaluation, rather than a familiar content subject to a novel form of truth-evaluation. The just-so story I've given suggests, in turn, that such cases will tend to grow into full-fledged object-language representations, but also that difficulties in making sense of iteration will tend to emerge. Note, for example, that modal logic certainly shares this feature. Iterated modalities aren't a normal feature of natural language.¹⁸ And the non-trivial implementation of iterated modalities requires a novel idea (in the typical formulation, the introduction of an accessibility relation;

¹⁷ The same emergence of molecular claims occurs elsewhere when speech act specifiers are added to a language. The negation-containing command 'Don't do that!' can be read only as a command not to perform some action. However, once a specifier 'I command' is added, we can distinguish between 'I command you not to do that' and 'I don't command you to do that.'

¹⁸ A Google search for 'necessarily possibly', for example, turns up only two uses that are not part of philosophical discussions. One:

The compiled list is not necessarily possibly incomplete.
looks like a misprint to me; the other:

we will return shortly to another way of thinking about the requisite novelty); absent that novelty, a ‘flat’ **S5** modality is the result of a flat-footed iteration.

The history of the subject is suggestive here. In Lewis and Langford (1932) there are no iterations of the modal operator \Box and \Diamond in the main text – such iterations occur only in the second appendix, and there only as a result of definitional reinterpretations of claims involving iterations of the strict conditional \prec . The real impetus toward iterative modalities in Lewis and Langford is the iteration of the strict conditional – an iteration presumably encouraged by what has traditionally been taken to be a looseness about use and mention which encouraged conflation of metalinguistic implication with object language conditionality, but which in the current context might better be taken as another manifestation of the multiple truth-mode-evaluability of conditional claims.¹⁹

8 Truth Pointwise and Setwise

On the **DOT** model, how do iterated modalities arise? Answering this question requires setting out a distinction between *pointwise* and *setwise* truth properties. Suppose we have identified a certain regular **VF** truth_X . Following the above reasoning, we can generate a set of *evaluation points*, and define truth_X as truth_w at each evaluation point w . It will sometimes happen that $\text{truth}_@$ can be identified with truth at some one of the evaluation points – call that evaluation point the *actual world*, or $@$. This needn’t happen – since truth_X is permitted a superclassical consequence relation, it could happen that the set of $\text{truth}_@$ forms an inconsistent $_X$ set, and hence fails to appear among the evaluation points. Call a **VF** *actualizable* if $\text{truth}_@$ can be identified with truth at one of its generated evaluation points. More generally, given any two **VF**’s truth_X and truth_Y , it can happen that the evaluation points induced by truth_Y are a subset of those induced by truth_X . In this case, we call truth_Y *commensurate* with truth_X . My discussion will focus primarily on commensurate **VF**’s, although interesting questions are raised by the possibility of incommensurate **VF**’s in semantic analysis.

When truth_X is an actualizable **VF**, we can compare it and $\text{truth}_@$ by describing truth_X as a setwise **VF** and $\text{truth}_@$ as a pointwise **VF**. Given the underlying evaluation point space, assignment of $\text{truth}_@$ values supervenes entirely on the features of a single evaluation point, while assignment of truth_X values supervenes on the

Well, doesn’t one necessarily possibly mean the other? If you’re going to disengage, you may have to pull people out of their houses?

is suggestive in its apparent use of ‘necessarily’ as an indicator of discourse structure, rather than sentential content.

¹⁹ Also suggestive, returning briefly to the thought that quantificational specification might be seen as a mode of truth, is the historical preference for a logic of quantification which avoids quantificational iteration, although there are important distinctions between this case and the modal case.

features of a set (the maximal set) of evaluation points. From here, we can generalize to relativized notions of truth, and thus distinguish between:

1. Calling a sentence true relative to a particular evaluation point – this is the relativization of $\text{truth}_@$, and in the modal case, amounts to counterfactual truth conditions.
2. Calling a sentence true relative to a set of evaluation points – this is the relativization of truth_X , and what to make of it is as yet an open question.

We can think of the relativization of $\text{truth}_@$ in the following way. The context of utterance of a claim provides with it a world with respect to which truth is to be evaluated. Supposing that we have a coherent way of thinking of each evaluation point as providing an instance of truth with respect to a world, we can then think of the evaluation of a claim for $\text{truth}_@$ as proceeding in two stages: first, the selection of a proper evaluation point, and second, the determination of $\text{truth}_@$ with respect to that evaluation point. Following the same line of thought, we can also think of the context of utterance as providing a set of worlds, and then think of the evaluation of a claim for truth_X as proceeding in two stages: first, the selection of a proper *set* of evaluation points, and second, the determination of truth_X with respect to that set.

Given this picture, to ask whether a claim holds $_X$ at a particular evaluation point is an ill-formed question, since holding $_X$ is something which happens at a set. Similarly, asking whether a claim holds $_@$ relative to a set of evaluation points is an ill-formed question, since holding $_@$ is something which happens relative to a point. Pointwise and setwise modes of truth look for different inputs to their evaluation. This then provides a hurdle to be overcome when we want to induce the formation of an object-language modal (or, more generally, truth_X) operator, since that operator is intended to form claims evaluable pointwise, rather than setwise. How are we to do pointwise evaluation of a setwise-determined property? More generally, there's a question about whether setwise and truthwise properties can be brought into semantic interaction. There are two obvious ways to go here:

1. Pointwise properties can be lifted up to setwise ones, via the stipulation that pointwise property P holds at a set S of evaluation points if and only if P holds at every point in S . If we think of a pointwise property as a function from points to $\{0, 1\}$ and a setwise property as a function from sets of points to $\{0, 1\}$, it follows that the lifts of pointwise properties form a subset of the setwise properties. This lifting technique frequently features in dynamic semantic systems, in which (for example) claims are evaluated at, or are maps among, epistemic states consisting of sets of worlds, and atoms are used as filters on sets, reducing those sets to the points within them that pass a pointwise test.
2. Setwise properties can be pushed down to pointwise properties, by mandating a correlation between points and associated sets. This is, in effect, the role of the accessibility relation in a Kripke semantics – it allows the intrinsically setwise modals to be evaluated pointwise by providing each point with an associated set – the set of accessible worlds. The pushing down then allows meaningful iteration of modal operators, since the setwise modal evaluation can now be done

pointwise, and the possibility of variation in associated set from world to world in turn allows a nontrivial (i.e., not merely **S5**) logic to the iteration. Note that it's an immediate consequence of Cantor's Theorem that not all setwise properties can be represented pointwise in this way, so there's an essential loss of expressive power in opting for pointwise, rather than setwise, properties as basic.

One way of taking the central thesis of this paper is as a plea for setwise conceptions of truth. However, the 'setwise' terminology here can suggest a less radical reading of that thesis than I intend. An insistence that certain semantic phenomena are best accounted for via setwise truth properties can – when coupled with a background picture on which the setwise property is actualizable, and the points of evaluation are taken to be understood via the pointwise notion of actual truth – be seen as mere terminological shuffling. If the space of evaluation points has its origin in standard pointwise truth, then what difference whether we speak pointwise (via object-language operators quantifying over points via an accessibility relation) or setwise, since actual pointwise truth lies at the conceptual foundations either way? Part of my suggestion is that certain setwise semantic properties may be primitively setwise – be such that the space of evaluation points can only be understood in terms of the previously grasped setwise mode of truth, rather than the other way around. The 'setwise' terminology is damaging here, since it suggests a characterization of the **VF** in terms of the set of evaluation points.

9 What Makes a Mode of Truth a Mode of *Truth*?

I have suggested that necessity can be understood as the possession of a distinctive property truth_\square by claims, and that that property can in turn be understood as one of a potential collection of setwise **VF**'s. But there's a natural line of resistance to these suggestions which needs to be addressed. Why should truth_\square be thought of as a kind of *truth*, rather than just as necessity under an idiosyncratic label? In this section I sketch three tests for determining whether a putative **VF** should, indeed, be thought of as a veracity.

Let T_X be a property proposed as a **VF**. The first question one should ask about T_X is whether it supports the Tarskian disquotational scheme. This question can be asked in two forms:

Biconditional Disquotation: Given mode T_X , do we have, for every sentence ϕ , the truth of $\phi \leftrightarrow T_X \ulcorner \phi \urcorner$?

Inferential Disquotation: Given mode T_X , do we have, for every sentence ϕ , that $\phi \models T_X \ulcorner \phi \urcorner$ and $T_X \ulcorner \phi \urcorner \models \phi$?

One might think that the answer to this question (in either form) for T_\square is obviously 'no'. After all, the biconditional $\phi \leftrightarrow \square\phi$ and the dual inferences $\phi \models \square\phi$ and $\square\phi \models \phi$ are characteristic of the trivial modal logic **Triv**, and of course in **Triv** necessity is equivalent to truth, meaning that in that modality, T_\square just is T_\oplus . And there's a natural generalization to this thought. The disquotational scheme fixes the

extension of the truth predicate, so if we insist on the disquotational scheme as a mark of a genuine truth property, have we not immediately given up **DOT** and insisted on a single truth property?

However, this line of argument is too quick. Inferential disquotation appeals because implication requires coordination in truth value, and the truth predicate reports on that very coordinated feature. Similarly, biconditional disquotation appeals because the biconditional requires coordination in truth value, and the truth predicate reports on the coordinated feature. Disquotational principles are forms of *internal alignment* tests. The test, then, should be properly calibrated. If we are serious about treating truth_X as a mode of truth, then the desired disquotational scheme ought to be constructed using truth_X rather than $\text{truth}_@$. Truth_X should be evaluated on its own terms, rather than on the terms of an alien truth property. Only this native evaluation will reveal whether truth_X is internally aligned in the way we expect of a truth predicate.²⁰ In the case of truth_\square , we should thus be asking whether $\phi \leftrightarrow T_\square \ulcorner \phi \urcorner$ is true_\square , and whether ϕ implies $\square T_\square \ulcorner \phi \urcorner$ and vice versa.

Consider the second question first. Implication \square is preservation of truth_\square , so ψ implies $\square \theta$ if any model making $\text{true}_\square \psi$ also makes $\text{true}_\square \theta$. Recast in more familiar terms, this is the requirement that $\square \psi \vDash @ \square \theta$, where $\vDash @$ is defined using $\text{truth}_@$. Thus the inferential disquotational principle for truth_\square requires:

$$\square \phi \vDash @ \square \square \phi$$

$$\square \square \phi \vDash @ \square \phi$$

Both of these principles hold in **S4**.²¹ So, if the modality underlying truth_\square is an **S4** modality or stronger, then the inferential disquotational principles hold, when properly construed as involving implication \square .

The biconditional disquotational principle, however, is not so easily secured. The truth_\square of the disquotational biconditional $\phi \leftrightarrow T_\square \ulcorner \phi \urcorner$ amounts to the $\text{truth}_@$ of $\square(\phi \leftrightarrow \square \phi)$. That principle holds only in modal logics at least as strong as **T**, **c**, which are not logics useful for an interesting notion of modality. The failure of biconditional disquotation for truth_\square should come as no surprise, though, and presents no real threat to truth_\square 's standing as a **VF**. The fundamental difficulty here is that the connective \leftrightarrow is a test on coordination of $\text{truth}_@$ values, so there's no reason to expect truth_\square to produce interesting correlations when coupled with \leftrightarrow . However, we could easily define a conditional which checks for coordination in truth_\square value:

$$\phi \leftrightarrow \square \psi \text{ is true at } w \text{ iff } \phi \text{ is true at every point } w' \text{ accessible from } w \text{ iff } \psi \text{ is true at every point } w' \text{ accessible from } w.$$

Again, in an **S4** logic, the resulting biconditional disquotational principle, formulated using \leftrightarrow_\square , will come out valid.

²⁰ For a more detailed argument that disquotational principles should hold with respect to the truth property under evaluation, rather than uniformly with respect to $\text{truth}_@$, see Asher, Dever, Pappas (2005)

²¹ In fact, both hold in a weaker logic requiring transitivity plus the reflexivity of all worlds which are accessed by some world.

Once a mode of truth is judged on its own terms, then, we see that it will exhibit appropriate disquotationality just in case its relation to the underlying evaluation point space is (approximately) **S4**.²² Metaphysical modality, for example, will pass this test, and hence truth_\perp has some reason to be treated as a genuine **VF**. Not everything passes, though. For example, falsity fails to be disquotational, even when evaluated on its own terms. The falsity biconditional:

$$\phi \leftrightarrow T_\perp \ulcorner \phi \urcorner$$

is indeed, as desired, true_\perp (that is, false). However, the inferential disquotational principle fails. ϕ does not imply $\perp T_\perp \ulcorner \phi \urcorner$ – implication $_\perp$ requires that the truth $_\perp$ of the premise force the truth $_\perp$ of the conclusion, but the falsity of ϕ does not force the falsity of $\neg\phi$. Nevertheless, the disquotational test is too weak by itself. Consider, for example, a ‘mode of truth’ T_∞ , where T_∞ is a property possessed by every sentence in the language. Then the appropriately localized disquotational principles, both inferential and biconditional, will hold trivially. Everything implies $_\infty$ everything else (T_∞ is *always* preserved), so the inferential principle must hold, and the disquotational biconditional, like everything else, is T_∞ .

The universality of implication $_\infty$ suggests another test for a genuine **VF** – that it produce a reasonable logic. I am deliberately vague here about what counts as ‘reasonable’. So strong a requirement that the logic be superclassical, or even precisely classical, although compatible with modes of truth other than truth_\oplus , would have what I take to be the undesirable consequence of ruling intuitionistic truth not a **VF**. But there should be *some* resemblance between the logic generated by a putative veracity feature T_X and inferential principles we are pretheoretically inclined to accept. T_∞ , in validating every inference, fails such a test as badly as is possible. Truth_\square , on the other hand, is close to being superclassical in its inferential structure. The non-modal fragment is straightforwardly superclassical. However, some classical inferential patterns are lost when truth_\square is extended to a language enriched with a truth_\square predicate. Superclassicality gives us $p \vee \neg p$ as a theorem \square , and the inferential disquotational principle gives us $p \vDash_\square T_\square \ulcorner p \urcorner$ and $\neg p \vDash_\square T_\square \ulcorner \neg p \urcorner$, but we should not expect as a theorem $T_\square \ulcorner p \urcorner \vee T_\square \ulcorner \neg p \urcorner$, since such a theorem would commit us to the necessitarian thesis that everything that happens, happens of necessity. Implication $_\square$, on the full language, thus fails to support proof by cases. The formal difficulties here are similar to those besetting the consequence relation of supervaluation theory, for reasons that we’ll return to below. Despite the failure of proof by cases, the inference scheme of truth_\square is reasonably similar to that of truth_\oplus , and it seems not unreasonable to say that truth_\square is here, as with disquotation, a fair candidate for a **VF**.

The third test for a genuine **VF** is even more nebulous. Truth, as has been stressed by a number of philosophers, is tightly integrated into a host of other important concepts. Dummett (1959) has emphasized the analogy between truth and the

²² Earlier truth_\square was related to the evaluation point space universally, so that truth_\square was truth_w for all evaluation points w . However, setwise modes needn’t receive this universal interpretation. If truth_Y is commensurate with truth_X , then it can be given setwise truth conditions relative to the larger evaluation point space generated by truth_X .

winning of a game, holding that it is essential to truth that it serve as the end of various practices. Davidson, especially in Davidson (1984a, b), has stressed the role of truth in integrating the practices of meaning ascription and belief content ascription. Lepore, especially in Lepore (1982, 1983), and Lepore and Loewer (1983), has explored the use of Davidson-style constraints on the concept of truth as a tool for ruling out certain notions (such as model-theoretic *truth in a structure*, infinite list-like definitions of truth, and truth defined via substitutional quantification) as genuine notions of truth. To assert a claim is to be committed to the truth of that claim, and the norm governing both assertion and belief is arguably truth-based.²³ True beliefs, as Horwich (1990) puts it, ‘tend to facilitate the achievement of practical goals’ (40). Even roughly sketching the outlines of the role of truth in our larger cognitive architecture would vastly exceed the scope of this paper, but whatever that infrastructure is, any would-be **VF** had better more or less mirror it. Again, T_{\perp} falls short here – falsity isn’t something we aim for in assertion and belief, so it can’t with any plausibility be called a truth value.

How does truth $_{\square}$ fare? That a claim is true $_{\square}$ does indeed entail that one ought to/is justified in asserting or believing it, or that acceptance of it will, under normal conditions, promote practical goals. However, so much is a predictable consequence of the **T** axiom – since truth $_{\square}$ entails truth $_{@}$, a claim which is true $_{\square}$ will inherit the integration into broader practice enjoyed by one which is true $_{@}$. But since $\neg_{\square}\phi \not\equiv_{@} \neg\phi$, the shortcomings of truth $_{\square}$ reveal themselves when we consider the cognitive role of claims which are not true $_{\square}$. Here we do not find what we would expect of a **VF**. That a claim fails of truth $_{\square}$ is in itself no good reason to refrain from asserting or believing it; claims which are not true $_{\square}$ can be reliably useful in practical reasoning.²⁴ For this reason, I am inclined – after coming all this way! – to think that truth $_{\square}$ is not, after all, best understood as a mode of truth. However, I’ll argue below that it has close cousins which are best so understood.

But is that even possible? There’s a worry here similar to that posed by the use of the disquotational principles in identifying truth. Those principles threatened to force a unique concept of truth on us, as the sole solution to the system of equations the disquotational instances represent. Similarly, why shouldn’t the constraints imposed by integration in our broader cognitive practices impose a unique solution? If, for example, that which is true is to be believed, and that which is false is not to be

²³ Williamson (2000), among others, argues that the regulative norm of assertion is based on knowledge, rather than truth, and Sutton (2005) extends the claim to belief as well. The point will play no significant role in the current discussion, but I am inclined to think that a norm of truth suffices, and that the apparent role of knowledge in the norms is a result of the fact that the norm is a rule to be followed, and to *follow* a rule is to adopt a course of action which one knows to be in keeping with the rule.

²⁴ Can we retrench by pointing out the role of false $_{\square}$? I don’t think so; such a retrenchment still fails to account for the appropriateness, and perhaps even the obligatoriness, of believing and otherwise cognitively integrating claims which are not true $_{\square}$. But there are messy questions about the role and conceptual coherence of truth value gaps here; the considerations raised by Dummett (1959) will be relevant.

believed, then the normative constraints on belief will fix the extension of truth. The response to the earlier challenge was to suggest that the disquotational principles had to be properly situated within the context of the putative **VF** at hand. An analogous move here, however, would be unfortunate. Were we to suggest that truth_X need only be integrated with belief_X , via the normative requirement to believe X that which is true_X , we would simply vitiate the cognitive integration requirement, and lose our grasp of what is distinctive about truth. The disquotational scheme should be relativized and localized, because it imposes an internal coherence constraint on a **VF**, and that internal coherence can then be locally realized. But there is a genuine, independent, and unique cognitive practice into which any notion of truth needs to be integrated.

However, the uniqueness of the practice is insufficient to fix the uniqueness of the integrated **VF**. Belief and assertion, for example, may be subject to a number of (perhaps conceptually closely related) norms. Thus there may be a sense of 'should' according to which one should believe what is $\text{true}_@$ and not believe what is not $\text{true}_@$, but another sense of 'should' according to which one should believe what is true_X , for some X , and not believe what is not true_X . The line of thought here has certain parallels to a line which has attracted much attention in the development of 'minimalist' theories of truth. This second line holds, more or less, that various parts of our discourse and practice will be, as Wright (1992) puts it, 'truth-apt', in that the structure of those regions of practice will support the introduction of a disquotational truth predicate. However, these different regions of discourse can nevertheless differ substantially in the norms of assertion and evidence governing them, so that there is, in some important sense, no single substantive notion of truth common to all. This is a disunity of truth picture, but it is not precisely the picture I am offering here. The 'no substantive common truth' picture puts together the following two thoughts:

1. Discourses of different vocabulary and subject matter are subject to different norms of assertion and evidence.
2. These normative variations give rise to a multiplicity of local truth conceptions which are metaphysically diverse (some, for example, 'realist' and others 'anti-realist') but logically unified, via their common participation in the disquotational scheme (hence, for example, the ready availability of arbitrary truth-functional combinations of statements drawn from diverse regions).

The proposal I am currently making, on the other hand, puts together these two thoughts:

1. There are (or can be) norms of assertion and evidence which hold across discourses of all vocabulary and subject matter, but there can be more than one such norm, such that one can be acting appropriately in one sense in believing or asserting that ϕ , but acting inappropriately in another sense. Which norms are effective in guiding behaviour may then vary from context to context.
2. These varying norms each gives rise to a logically unified conception of truth adhering to a (properly localized) disquotational scheme, but there is no single scheme in which the multiplicities are brought together.

The move is rather abstract at the moment, and intended only to create conceptual space for disunity in the face of a unique cognitive architecture.

10 Conclusio Res: Vagueness is a Modality

The discussion of the disunity hypothesis thus far has been rather abstract, and the one specific example I've worked through – truth $_{\square}$ – I've proceeded to argue in fact should not be thought of as a mode of truth. I now want to turn to a genuine example, attempting to show that supervaluationist theories of vagueness are best understood as introducing one or more setwise conceptions of truth.

I begin with a brief sketch of the standard semantics for a supervaluationist account of vagueness, derived (but somewhat simplified) from Fine (1975). Call a *Fine space* an ordered triple $\langle S, \rho, r \rangle$, where S is a set of evaluation points, ρ is a reflexive accessibility relation about the evaluation points, and r is a root point among the members of S . Each evaluation point is then a *partial first-order model*; hence, evaluation point w is itself an ordered triple $\langle D, \llbracket \cdot \rrbracket_w^+, \llbracket \cdot \rrbracket_w^- \rangle$, where D is the domain (common to all evaluation points in the space) and $\llbracket \cdot \rrbracket_w^+$ and $\llbracket \cdot \rrbracket_w^-$ are, respectively, extension and anti-extension functions on $D^{<\infty}$. Given an assignment function g , we then define truth (\vdash) and falsity (\dashv) at an evaluation point for a first-order language amplified with a determinacy operator \mathcal{D} in the obvious way:

$$\begin{aligned} w, g \vdash \Pi^n x_{i_1} \dots x_{i_n} &\text{ iff } \langle g(x_{i_1}), \dots, g(x_{i_n}) \rangle \in \llbracket \Pi^n \rrbracket_w^+ \\ w, g \dashv \Pi^n x_{i_1} \dots x_{i_n} &\text{ iff } \langle g(x_{i_1}), \dots, g(x_{i_n}) \rangle \in \llbracket \Pi^n \rrbracket_w^- \\ w, g \vdash \neg \phi &\text{ iff } w, g \dashv \phi \\ w, g \dashv \neg \phi &\text{ iff } w, g \vdash \phi \\ w, g \vdash (\phi \wedge \psi) &\text{ iff } w, g \vdash \phi \text{ and } w, g \vdash \psi \\ w, g \dashv (\phi \wedge \psi) &\text{ iff } w, g \dashv \phi \text{ or } w, g \dashv \psi \\ w, g \vdash \exists x_i \phi &\text{ iff for some } o \in D, w, g[o/x_i] \vdash \phi \\ w, g \dashv \exists x_i \phi &\text{ iff for every } o \in D, w, g[o/x_i] \dashv \phi. \\ w, g \vdash \mathcal{D}\phi &\text{ iff for every } v \text{ such that } \rho wv, v, g \vdash \phi. \\ w, g \dashv \mathcal{D}\phi &\text{ iff for every } v \text{ such that } \rho wv, v, g \dashv \phi. \end{aligned}$$

Evaluation point w' *extends* point w (or $w' \geq w$, if $\llbracket \cdot \rrbracket_w^+ \subseteq \llbracket \cdot \rrbracket_{w'}^+$ and $\llbracket \cdot \rrbracket_w^- \subseteq \llbracket \cdot \rrbracket_{w'}^-$). Point w is *complete* if, for any predicate Π^n , $\llbracket \Pi^n \rrbracket_w^+ \cup \llbracket \Pi^n \rrbracket_w^- = D^n$. We then say that ϕ is *supertrue* in a Fine space if ϕ is true at all complete extensions of the root r of the space.

The standard conception of supervaluationism carries, on top of this formal apparatus, certain philosophical commitments to the interpretation of that formalism. Supertruth is treated as the unique **VF** produced by supervaluationism, and the points of evaluation, on which the behaviour of supertruth rests, are conceived of metalinguistically, as describing the semantic behaviour of a permissible completion of a semantically underdetermined linguistic practice. These commitments contain

three elements that, from the perspective of the **DOT** hypothesis, appear optional. First, it insists on a *single* conception of truth – ‘truth as supertruth’ – underlying the vague portion of our linguistic practice. I will suggest, in contrast, that we may profitably distinguish determinate truth, supertruth, truth at an evaluation point, and perhaps a version of ‘actual truth’ corresponding to truth at a distinguished evaluation point or set of points. Second, the standard conception treats supertruth as conceptually posterior to truth at an evaluation point. I’ve emphasized above that such a perspective is optional, and that there’s no reason why we can’t recognize setwise **VF**’s as giving conceptual rise to their associated spaces of evaluation points, rather than the other way around. Third, the standard conception is committed to a particular view of the nature of the semantic feature ‘true at evaluation point w ’, opting for a metalinguistic gloss on it, rather than some other reading. Even if one opts to treat evaluation point truth as prior to setwise supertruth, one could give a non-metalinguistic read on that pointwise notion, perhaps following the object language operator reduction strategy sketched above.

So far this merely points to some room in logical space that the standard conception of supervaluationism has left unexplored and uninhabited. To promote the new real estate a bit, I want to point out some (well-known) problems with supervaluation theory, and suggest that greater attention to the **DOT** hypothesis can resolve these problems. These problems will fall into two categories. First, I’ll rehearse some non-standard features of the logic generated by supertruth. Next, I’ll point out that supervaluation has difficulty accommodating a rich notion of higher-order vagueness.

Fine’s original presentation of a supervaluationist account of vagueness already observes that the deduction theorem fails for the language containing a determinacy operator (or, alternatively, a supertruth operator). Define *superentailment* as preservation of supertruth:

Γ **superentails** ϕ ($\Gamma \models_S \phi$) if, for every Fine space \mathcal{F} , if for every assignment function g and every complete extension w of r in \mathcal{F} , $g \vdash \Gamma$, then for every assignment function g and every complete extension w of r in \mathcal{F} , $w, g \vdash \phi$.

It then follows immediately that $\phi \models_S T_S \ulcorner \phi \urcorner$ and $T_S \ulcorner \phi \urcorner \models_S \phi$, where T_S is an object language supertruth predicate. Complete evaluation points can be extended only by themselves (or other isomorphic evaluation points), so supertruth is a modal operator with the logic of **Triv** on complete evaluation points. Since superimplication considers the supertruth of premise and conclusion, it considers the behaviour of premise and conclusion only at complete evaluation points, so the overall (super)implicational structure of supertruth is that of a trivial modal logic. This result is, of course, to be desired if we are thinking of supertruth as a **VF**, since we then want it to satisfy the disquotational principle, the inferential version of which we’ve just secured. Similarly, if we (following the intention of Fine’s presentation) require the accessibility relation ρ to be a subset of the extension relation, then we again secure $\phi \models_S \mathcal{D}\phi$ and $\mathcal{D}\phi \models_S \phi$, so that determinacy also acts truth-like in obeying the inferential disquotational principle.

However, both supertruth and determinate truth fail the biconditional version of the disquotational principle. Let ϕ be a claim that is true at some but not all complete extensions of the root r . Then $T_S \ulcorner \phi \urcorner$ will be false (or gappy, depending on how it is implemented in the object language) at all complete evaluation points, so the biconditional will fail at those complete points at which ϕ is true, and thus will not be supertrue. A similar argument shows that $\phi \leftrightarrow \mathcal{D}\phi$ needn't be supertrue. I've already argued that the failure of these biconditionals should be expected, and isn't a decisive objection to the treatment of supertruth or determinate truth as **VF**'s – we merely have to replace the pointwise \leftrightarrow with a setwise modal biconditional, and an appropriately reworded disquotational biconditional then will hold for supertruth and determinate truth (and will fail for truth at a point or 'actual truth'). But the failure of the disquotational biconditional, for the material biconditional, does have implications for the logic of supervaluation. We have instances in which $\phi \models_S \psi$ but $\not\models_S \phi \rightarrow \psi$, so the deduction theorem fails for superimplication. From the failure of the deduction theorem, other troubling consequences flow, such as the failures of conditional proof, proof by cases, and indirect proof. Thus, the first objection to supervaluationism: despite its stated goal of retaining a classical logic for vagueness-infested portions of the language, it ends up casting aside basic forms of inference. Williamson (1994) has made much of this objection.

Precisely because supervaluationism *does* deliver an inferential disquotational principle for the determinacy operator, it is poorly positioned for yielding a rich account of higher-order vagueness. With a logic of **Triv**, prefix blocks built from the operator \mathcal{D} will always reduce to a length of one, and hence always reduce to first-order vagueness. The issue is complicated somewhat by the absence of a deduction theorem – in light of this failure, molecular claims with various irreducible nestings of determinacy operators are compatible with the universal atomic prefix block reduction. Nevertheless, the claim that there's simply nothing to distinguish truth, determinate truth, determinately determinate truth, and so on is hard to put together with a real possibility of higher-order vagueness. (In fact, it seems to me that there's a genuine tension in our inferential intuitions regarding determinacy, pushing us both toward and away from the inferential disquotational principles. More on this below.)

Supervaluationist accounts of the logic of \mathcal{D} , and hence of higher-order vagueness, are also subject to a line of objection originating in Wright (1987), in which arguments are produced from *prima facie* plausible principles governing \mathcal{D} to show that a contradiction results. This line of arguments is not specific to (the standard understanding of) supervaluationism, although other accounts of vagueness will have different conceptual motivations of \mathcal{D} and hence will have to be examined case-by-case to determine whether they motivate the necessary inferential principles. Let R be a paradigm vague predicate, and let $\{x_i\}$ be a sorites sequence for R , beginning with an unproblematically R object and proceeding in small (perhaps epistemically indistinguishable) differential increments to an object x_n which is unproblematically not R . Let x' be the successor of x , for any x in this sequence. The following *gap principles* then seem intuitively appealing:

$$(\mathbf{Gn}) \forall x(\mathcal{D}^n Rx \rightarrow \neg \mathcal{D} \neg \mathcal{D}^{n-1} Rx')$$

$$(\mathbf{Gn}') \forall x(\mathcal{D} \neg \mathcal{D}^{n-1} \neg Rx \rightarrow \neg \mathcal{D}^n \neg Rx').$$

The gap principles capture the thought that at no level of determinateness do we find that determinate (to that degree) possession of R transitions in a single step to determinate lack of determinate (to that degree but one) possession of R . A sequence of arguments, beginning with Wright and running through Fara (2003) and Asher et al. (2005), seeks to extract contradictions from the gap principles. Wright's argument requires the principle:

(DET) If $\Sigma \models \phi$, and every atomic sentence of Σ is in the scope of a \mathcal{D} operator, then $\Sigma \models \mathcal{D}\phi$.²⁵

Fara's proof instead requires a rule of determinacy introduction:

(\mathcal{D} -intro) If $\Sigma \models \phi$, then $\Sigma \models \mathcal{D}\phi$.

and avoids use of the rule in conditional proofs. Finally, the argument can be run, as in Asher et al., with very minimal assumptions about the logic of \mathcal{D} . Permit arbitrary determinateness of the sorites sequence endpoints (positive at one end; negative at the other) and appeal to arbitrary determinatizations of the gap principles, and contradiction again follows.²⁶

11 A Disunified Treatment of Vagueness

I suggested earlier that a difficulty with iterations of an object-language operator for a modal notion should be taken as an indication that that modality is best understood directly as a mode of truth on the **DOT** model. Given the problems with higher-order vagueness just rehearsed, I thus want to think about a disunity treatment of vagueness. The central thought in this treatment will be that supertruth and determinate truth are both setwise notions of truth, and that sensitivity to this fact will shed light on otherwise confusing features of supervaluationism.

If supertruth and determinate truth are setwise notions of truth, then (following the path set out above) underlying them (compositionally, if not conceptually) is a notion of pointwise truth at an evaluation point. As a result, talk of truth, and of truth-related notions such as implication, will be systematically ambiguous in vagueness contexts. In particular, we can distinguish between multiple notions of implication:

1. **Superentailment:** Γ superentails ϕ ($\Gamma \models_S \phi$) if, for every Fine space \mathcal{F} , if for every assignment function g and every complete extension w of r in \mathcal{F} , $w, g \vdash \Gamma$, then for every assignment function g and every complete extension w of r in \mathcal{F} , $w, g \vdash \phi$. (As above.)

²⁵ **DET** requires that the logic of \mathcal{D} obey both **4** and **E**.

²⁶ See Asher et al. (2005) for full presentation of all three arguments.

2. **Determinate entailment:** Γ determinately entails ϕ ($\Gamma \models_S \phi$) if, for every Fine space \mathcal{F} , if for every assignment function g and every evaluation point w in \mathcal{F} , for every evaluation point v such that ρwv , $v, g \vdash \Gamma$, then for every assignment function g and every evaluation point w in \mathcal{F} , for every evaluation point v such that ρwv , $v, g \vdash \phi$.
3. **Pointwise entailment:** Γ pointwise entails ϕ if, for every Fine space \mathcal{F} , for every evaluation point w in \mathcal{F} , for every assignment function g , if $w, g \vdash \Gamma$, then $w, g \vdash \phi$.

Depending on the details of the model theory and the implementation of an object-language supertruth predicate, superentailment and determinate entailment can have some interesting logical differences. However, the fundamental difference here is between a setwise (global) consequence relation and a pointwise (local) consequence relation. There are two important points to make about the choice between the global and local consequence relations:

1. The local consequence relation is, in many ways, much better behaved. Like the global relation, its logic on the \mathcal{D} -free fragment of the language is superclassical. However, unlike the global relation, it allows a deduction theorem and hence remains superclassical even when \mathcal{D} and a supertruth predicate are added to the language.²⁷ The inferences from ϕ to $T_S \ulcorner \phi \urcorner$ and $\mathcal{D}\phi$ fail of local validity (given plausible assumptions about the shape of the Fine spaces), so room is made for genuine higher-order vagueness. Wright's principle **DET** and Fara's **\mathcal{D} -intro** both fail locally, so neither the Wright nor the Fara anti-gap-principle arguments succeed when reconstrued locally. The final version persists, but although we cannot have arbitrary determinatization of the gap principles, we can have determinatization to any degree n we desire.
2. There is good reason to think that the local consequence relation is what is wanted when we are dealing with the logic of determinacy or object-language expressed supertruth. The global consequence relations, by checking the truth of premise and conclusion throughout the evaluation point space, are poorly designed to make logical distinctions which depend on the changing shape of the accessibility relation over the space. Put another way, the setwise consequence notions, by quantifying freely over sets, fail to make distinctions which depend on looking at particular sets in particular contexts (especially contexts of imbeddings). Perhaps a more sensitively designed global consequence relation could do the job, but the local consequence relation straightforwardly implemented does what is wanted.

Here, then, is the **DOT** picture of supervaluation theory. Supervaluation proposes *modes of truth* – supertruth, determinate truth. These modes are regular and

²⁷ Depending on how truth-value gaps are treated, it may be necessary to restrict the evaluation point quantifier of the local consequence relation to complete points. See Asher et al. (2005) for more on the logical issues here.

commensurable,²⁸ and hence can be thought of as setwise truth properties. Thus we have available both local and global consequence relations. We get into a muddle because both the local and the global have their attractions here. Recognizing supertruth and determinate truth as genuine **VF**'s, we expect disquotational principles. The inferential version of disquotation requires the global consequence relation, so we have reason to attend to the global. On the other hand, we want, when talking about the semantics of vagueness – especially in the guise of determinacy talk – to recognize logical complexities that emerge only when we attend to the local consequence relation. There are deep tensions in our attitude toward truth in matters of vagueness. The impulse to superclassicality – an overt goal of supervaluation – leads us to favour instances of the excluded middle. Even in cases of vagueness, object *o* either is or is not red.²⁹ But recognition of vagueness pushes us away from semantic recognition of the excluded middle – we don't want to endorse the claim that either it's true that *o* is red, or it's false that *o* is red. Williamson (1994) takes this split to show that supervaluation cannot hold onto disquotation, but the **DOT** picture suggests instead that we are simultaneously endorsing the local validity of the excluded middle and the global validity of the disquotational principles. Locally, the metasemantic excluded middle follows, but is unproblematic, and fails to reflect on supertruth, since the disquotational principles fail locally. Globally, the metasemantic excluded middle fails to follow, due to the failure of proof by cases, and the disquotational principles instead persist. Again, consider the naive attitude toward the questions 'Is that red?', 'Is that truly red?', 'Is that truly truly red', and so on. One feels simultaneously the attraction of insisting that a positive answer to the first question trivially entails a positive answer to the subsequent questions in the sequence and the legitimacy of the thought that the stakes are being raised as the 'truly' operators are being nested. A distinction between global and local consequence, resting on a distinction between setwise and pointwise truth, explains the conflict.

12 Two Objections to Supervaluationism

12.1 First Objection

We return at last to Fodor and Lepore's objection to supervaluationism. Recall that the objection holds that supervaluationism flouts a basic principle (**P**) that 'conceptual truths must be respected by all classical models, including classical valuations' (512), because it seeks to explain vague, and hence gappy predicates using a model theory on which the predicates are non-gappy at all relevant points. Earlier,

²⁸ Whether they are, in addition, actualizable will depend more or less on whether epistemicism about vagueness is true.

²⁹ Maybe – not everyone feels the draw of this. A disunity approach accounts also for the split in intuitions here.

I argued that this objection was, at its core, an attempt to protect the status of truth-conditional semantics by refusing to license supertruth as the fundamental semantic coin. Seen in this light, the objection rests on a refusal to take supertruth seriously as a **VF**. Imagine a similar objection raised against a Lewisian account of modality in terms of quantification over concrete worlds – the worlds themselves represent only non-modal facts, and we can't ground an account of modal truth on a wholly non-modal basis (one can perhaps see Kripke (1980)'s Humphrey argument as the first instance of such an objection). One way of responding to this argument, in keeping with the current project, is to explain that the possible worlds are a theory-internal device, and that the notions of necessity and possibility are conceptually prior to that of a possible world. Another way is to hold on to the constitutive role of the worlds in explaining the modal concepts, but then hold that the notion of truth at a world must, in turn, be understood in a modally loaded way (again as above, perhaps as truth under certain counterfactual suppositions). Both of these responses, it seems to me, remove entirely the objection. Only if one is wholly reductionist (as Lewis is), and wants the modal notions to depend on a non-modally conceived underlying notion, does the objection have teeth.

As in modality, so in vagueness. Supervaluationism is a threat to truth-conditional semantics only if supertruth is not a genuine **VF**. But why should it not be? If we take supertruth to be fundamentally a setwise truth notion, with the pointwise notions derivative on it, then the bivalence of those points poses no threat to the genuine vagueness and gappiness of supertruth. If we take supertruth to be conceptually dependent on pointwise truth, but explain our grasp of pointwise truth via our grasp of vagueness-controlling object language operators, then there is no threat to the genuine vagueness and gappiness of supertruth. Only if we are whole-heartedly reductionist about supertruth, analogously to Lewis, is there a worry. The line I am taking here has some similarity to Fodor and Lepore's Response 6 ('truth-in-a-classical-valuation is *my* word, so *I* get to decide what it entails' (529)), but either (a) simply accepts that pointwise truth doesn't genuinely *explain* how supertruth works (except in a technical sense), or (b) offers an account of pointwise truth in terms which *do* respect the conceptual truth at hand. *Perhaps* the Fine-ian supervaluationist, with the metalinguistic account of pointwise truth as truth in a determinate specification of the language, is sufficiently Lewisian in spirit to be subject to this line of attack (then again, perhaps not – the appeal to complete evaluation points is, modulo some alteration of the logic, dispensable, in favour of a sequence of infinite sequence of increasingly specified points), but there is much room outside this perspective on supervaluationism.

12.2 Second Objection

Williamson (1994) objects to what he argues is a failure of disquotation in supertruth. Importantly, he argues that the kinds of disquotation I have offered here are insufficient:

The supervaluationist did allow the statement that definitely A to entail and be entailed by the statement that A . In the same way, the supervaluationist may allow the statement that ‘ A ’ is supertrue to entail and be entailed by the statement that ‘ A ’. Were ‘if and only if’ to be used for mutual entailment, the disquotational schema would have a reading acceptable to the supervaluationist. It is not Tarski’s reading, on which ‘if and only if’ is the material biconditional. More important, the mutual entailment reading fails to capture the disquotational idea. If the truth predicate really does have the effect of stripping off quotation marks, then the material biconditional that ‘ A ’ is true if and only if A strips down to the tautology that A is true if and only if A . The supervaluationist denies that supertruth behaves like that; the availability of the mutual entailment reading is an irrelevance. (162–163)]

A raw insistence on the necessity of a *material* biconditional in the disquotational scheme would merely beg the question against the supervaluationist, so the central criticism concerns the origin of the disquotational biconditional as a consequence about content preservation under the truth predicate. Trivially, we have $\phi \leftrightarrow \phi$. Given that application of the truth predicate preserves content, we can replace ϕ with $T\ulcorner\phi\urcorner$ to obtain the disquotational biconditional with the material biconditional. The failure of ϕ and $T_S\ulcorner\phi\urcorner$ to intersubstitute here show that supertruth does not preserve content and is not a genuine truth property.

But this line of objection is deeply rooted in truth monism. Once a **DOT** picture has been accepted, it becomes clear that the appeal to content preservation contains an ambiguity. An appeal to the preservation of truth@ conditions will result in a truth predicate allowing intersubstitutability in material biconditionals. An appeal to the preservation of supertruth conditions, on the other hand, will result in a truth predicate allowing intersubstitutability in a setwise-defined supertruth-coordinating biconditional. There is no reason to think that there is *any* interesting biconditional that will allow intersubstitutability based on preservation of *total* content across all modes of truth, and hence no reason to favour one formulation of the biconditional scheme over another.

Williamson’s truth monism manifests three more times as the argument against supertruth proceeds. First, he says:

Truth _{T} is disquotational; supertruth is not. In order of definition, truth _{T} is primary; supertruth is secondary. (163)

(Here ‘truth _{T} ’ is truth at an evaluation point in a Fine space.) The first point has already been discussed. The second supposes, without adequate justification, that the setwise truth property of supertruth is to receive a reductionist understanding. But I have been suggesting throughout that multiple modes of truth can be understood independently, and that setwise notions of truth can be conceptually prior to their induced evaluation points. Second, Williamson says:

Once the supposed advantages of supertruth are seen to be illusory, it becomes overwhelmingly plausible to equate ordinary truth with the property that meets Tarski’s disquotational condition, truth _{T} . (163)

The move from truth _{T} to ordinary truth is too quick here – such a move requires a way of identifying some one of the evaluation points induced by supertruth as *the* privileged point which characterizes the property of ‘ordinary truth’. But there

may be no such way (the resistance to epistemicism about vagueness can be taken as a skepticism that there is such a way) – supertruth may not be an actualizable truth property. Even if supertruth is actualizable, the thought that there is a single notion of ‘ordinary truth’ which needs to be identified with a single **VF** is, again, a manifestation of truth monism. Third, Williamson says:

There remains the ‘definitely’ operator, with its semantics of admissible interpretations. However, this apparatus has lost its privileged connection with the concept of truth. Of any admissible valuation, we can ask whether it assigns truth to all and only the true sentences of the language and falsity to all and only the false ones. At most one valuation has that property. (164)

The truth monism is fully overt in the final sentence. But even in the thought that we can ask whether a valuation assigns truth to the true sentences, we suppose that the genuine property of truth is a pointwise property, which again overlooks the possibility that one way of speaking of truth of vague claims is a setwise way, which does not admit of meaningful evaluation at all with respect to a particular point.

13 Dynamic Semantics and the Omniscience Argument

Consider what I will call the *Omniscience Argument*:

Suppose one accepts the following principle:

Motan Principle: If ϕ entails ψ , then the following claim is true:

- If ϕ , then it must be the case that ψ .

Here ‘must’ is to be understood as an epistemic modal, tracking truth in all worlds compatible with what the relevant agent knows/believes.³⁰ The principle is, if not examined too closely, quite compelling, as consideration of instances demonstrates (‘If that’s a poodle, it must be a dog.’).

Since ϕ entails itself, the Motan Principle tells us that if ϕ , it must be the case that ϕ . Since $\neg\phi$ entails itself, the Motan Principle tells us that if $\neg\phi$, it must be the case that $\neg\phi$. Since, by the law of the excluded middle, either $\neg\phi$ or ϕ , it follows (via proof by cases) that either it must be the case that ϕ or it must be the case that $\neg\phi$. Thus, for any agent and any claim, either the agent knows/believes that the claim is true, or he knows/believes that it is false.

Readers will, I suspect, quickly identify some serious flaw in the Omniscience Argument (although there may be some disagreement about where the flaw is). Rather than discuss these flaws immediately, however, I will present a semantic framework within which the Omniscience Argument comes very close to working, and serves as a tool for exposing an interesting and perhaps troubling semantic lesson. After

³⁰ If the epistemic modal is interpreted in terms of belief instead of knowledge, the argument is more properly an omnidogmatic, rather than omniscience, one. Nothing I say here will depend on which interpretation is taken.

discussing the interaction between that lesson and the **DOT** hypothesis, I will close with a discussion of the more obvious diagnoses of the Omniscience Argument.

The update semantics of Veltman (1996) is a specific implementation of the guiding idea of dynamic semantics: the thought that the meaning of a sentence is to be given not via a truth-condition-possession proposition, but instead via a rule for transition between one state (understood perhaps as conversational context, or perhaps as the doxastic description of an agent) and another. In the case of update semantics, sentence meanings are functions from prior to posterior information states, tracking the impact of the sentence on the belief system of an agent in the conversation. Information states are modeled as sets of possible worlds: those worlds compatible with everything that the agent knows/believes.³¹

Consider the following simple update semantics. Let \mathcal{L} be a non-modal propositional language with the connectives \neg , \wedge , \vee , \rightarrow and \Rightarrow . On top of \mathcal{L} we build a simple modal language, by defining:

$$\mathcal{L}' = \mathcal{L} \cup \{\Box\phi : \phi \in \mathcal{L}\} \cup \{\Diamond\phi : \phi \in \mathcal{L}\}^{32}$$

Finally, we allow molecular combinations of the singly-modally-modified sentences of \mathcal{L}' by defining \mathcal{L}' to be the closure of \mathcal{L}' under \neg , \wedge , \vee , \rightarrow , and \Rightarrow . Let a model \mathcal{M} assign to each atom p a set of worlds $\llbracket p \rrbracket$ (intuitively, the worlds at which p is true). Then the update functions of sentences can be recursively defined as follows:

1. For atomic p , $s[p] = s \cap \llbracket p \rrbracket$.
2. $s[\neg\phi] = s - s[\phi]$.
3. $s[\phi \wedge \psi] = s[\phi][\psi]$
4. $s[\phi \vee \psi] = s \cap (s[p] \cup (s - s[p])[q])$
5. $s[\phi \rightarrow \psi] = s - (s[\phi] - s[\phi][\psi])$
6. $s[\phi \Rightarrow \psi] = \begin{cases} s & \text{if } s[\phi][\psi] = s[\phi] \\ \emptyset & \text{if } s[\phi][\psi] \neq s[\phi] \end{cases}$
7. $s[\Box\phi] = \begin{cases} s & \text{if } s[\phi] = s \\ \emptyset & \text{if } s[\phi] \neq s \end{cases}$
8. $s[\Diamond\phi] = \begin{cases} s & \text{if } s[\phi] \neq \emptyset \\ \emptyset & \text{if } s[\phi] = \emptyset \end{cases}$

\Box , \Diamond , and \Rightarrow act as tests, either leaving the information state unchanged (if the test is passed) or reducing it to absurdity (if the test is failed). \Box tests an information state to see if every point in it supports the truth of the matrix sentence; \Diamond tests a

³¹ I use here the simpler of the update systems developed in Veltman (1996), ignoring the complications added by the introduction of expectation patterns in the less simple system.

³² Following the most common practice in the update semantics literature, I implement the epistemic modals only when they have scope over non-modal sentences, disallowing iterated modal operators. Note the interaction between this practice and the Iteration Hypothesis set out above.

state to see if some point in it supports the truth of the matrix, and \Rightarrow tests a state to see if, when updated with the antecedent, every point in it supports the truth of the consequent. A simple inductive proof shows that, when $\phi \in \mathcal{L}$, updating an information state σ with ϕ is equivalent to taking the intersection of σ with the set of worlds at which ϕ , construed classically, is true.

Since update semantics does not give truth conditions for sentences, logical consequence cannot be construed as preservation of truth. Instead, we introduce the *update-to-test* notion of consequence:

UTC Consequence: $\phi_1, \dots, \phi_n \Vdash \psi$ iff for all information states s , $s[\phi_1] \dots [\phi_n] = s[\phi_1] \dots [\phi_n][\psi]$.³³

UTC consequence takes a sentence to be a consequence of some premises if, given any information state, the result of updating with all of the premises (consecutively) is the same as the result of updating with all of the premises and then the conclusion.³⁴ Given the earlier observation that the update functions mimic classical semantics on the modal-free \mathcal{L} , it follows that UTC consequence is classical on \mathcal{L} .

The following two features of update semantics are central for considering the status of the Omniscience Argument within this framework:

1. The combination of the test conception of \Box and UTC consequence strongly suggests that $\phi \Vdash \Box\phi$. For this implication to hold is for updating an arbitrary state σ with ϕ to produce the same result as updating σ first with ϕ and then with $\Box\phi$. But updating with $\Box\phi$ merely tests to see that each point in σ supports ϕ , which it presumably will, after the update with ϕ . Somewhat more carefully, we reason as follows:

Since $\Box\phi \in \mathcal{L}'$, $\phi \in \mathcal{L}$. Thus $\sigma[\phi] = \sigma \cap \llbracket\phi\rrbracket$. By the clause for \Box , $\sigma[\phi][\Box\phi]$ is just $\sigma[\phi]$ if $\sigma[\phi] = \sigma[\phi][\phi]$, which in turns is true just in case $\sigma \cap \llbracket\phi\rrbracket = \sigma \cap \llbracket\phi\rrbracket \cap \llbracket\phi\rrbracket$. But this does hold, so $\phi \Vdash \Box\phi$.

The implication also holds in the other direction: $\Box\phi \Vdash \phi$. To see this, we need that $\sigma[\Box\phi] = \sigma[\Box\phi][\phi]$. But $\sigma[\Box\phi]$ is either empty (if σ fails the test) or just σ (if σ passes the test). If $\sigma[\Box\phi]$ is empty, then further updating with ϕ will keep it empty, so the desired identity holds. If $\sigma[\Box\phi]$ is σ , then, because the test was passed, $\sigma[\phi]$ was σ , so updating again with ϕ will still yield σ . Thus ϕ and $\Box\phi$ are always mutually entailing.

2. Both the test conditional \Rightarrow and the material conditional \rightarrow support a deduction theorem: if $\phi \Vdash \psi$, then $\Vdash \phi \Rightarrow \psi$ and $\Vdash \phi \rightarrow \psi$. For the test conditional, this claim is trivial: if $\phi \Vdash \psi$, then updating with ϕ then ψ produces the same result as updating with ϕ , but this is exactly the test that \Rightarrow imposes. For the material conditional, the claim is less trivial, but still true:

³³ This logical consequence relation is \Vdash_2 , the second of the three consequence relations Veltman (1996) considers for update semantics.

³⁴ Thus the conclusion follows from the premises just in case any information state updated with the premises is a fixed point of the update function of the conclusion.

Suppose first that $\phi \Vdash \psi$. Then $s[\phi][\psi] = s[\phi]$, so $s[\phi \rightarrow \psi] = s - (s[\phi] - s[\phi][\psi]) = s$. Hence $\Vdash \phi \rightarrow \psi$. Suppose on the other hand that $\Vdash \phi \rightarrow \psi$. Then $s[\phi \rightarrow \psi] = s$, so $s - (s[\phi] - s[\phi][\psi]) = s$. Thus $s[\phi] - s[\phi][\psi] = \emptyset$, and $s[\phi] = s[\phi][\psi]$. Thus $\phi \Vdash \psi$.

From the combination of these two features, the Motan principle follows. Suppose $\phi \Vdash \psi$. Since $\psi \Vdash \Box\psi$, $\phi \Vdash \Box\psi$.³⁵ Then, by the deduction theorem, $\Vdash \phi \Rightarrow \Box\psi$ and $\Vdash \phi \rightarrow \Box\psi$. Thus the Motan principle holds on either the test or the material conditional interpretation of the relevant ‘if’.

While update semantics does endorse the Motan principle, it nevertheless can easily be seen not to endorse the conclusion of the Omniscience Argument. Consider the following *partiscient* model. Let σ be an information state containing only worlds w_1 and w_2 , and let $\llbracket p \rrbracket = \{w_1\}$. The Omniscience Argument predicts that $\Vdash \Box p \vee \Box\neg p$, and hence that $\sigma = \sigma[\Box p \vee \Box\neg p]$. However, the identity fails to hold:

1. $\sigma[\Box p \vee \Box\neg p] = \sigma \cap (\sigma[\Box p] \cup (\sigma - \sigma[\Box p])[\Box\neg p])$ [By the clause for disjunction.]
2. $= \sigma \cap (\emptyset \cup (\sigma - \emptyset)[\Box\neg p])$ [Since σ fails the test imposed by $\Box p$.]
3. $= \sigma \cap \sigma[\Box\neg p]$
4. $= \sigma \cap \emptyset$ [Since σ fails the test imposed by $\neg\Box\neg p$.]
5. $= \emptyset \neq \sigma$

So the Omniscience Argument fails in update semantics. This is good news for update semantics, of course, but, given the endorsement of the Motan Principle, where does the argument fail? The lacuna comes in the failure of UTC consequence to support proof by cases. One can have $\phi \Vdash \theta$ and $\psi \Vdash \theta$ without having $\phi \vee \psi \Vdash \theta$. Roughly, each of ϕ and ψ may perform an update which is sufficient to yield an information state supporting θ , but the disjunction of the two, precisely because of its agnosticism, may fail to form any useful update, or indeed any nontrivial update at all.

The failure of proof by cases is not the only logical peculiarity of update semantics and UTC consequence brought out by the Omniscience Argument. The Omniscience Argument can be restructured to appeal to modus tollens. By the Motan principle, $\phi \Vdash \Box\phi$. Thus by contraposition, $\neg\Box\phi \Vdash \neg\phi$. But by a second application of the Motan principle, $\neg\phi \Vdash \Box\neg\phi$, so by transitivity, $\neg\Box\phi \Vdash \Box\neg\phi$. Then by the deduction theorem, $\Vdash \neg\Box\phi \rightarrow \Box\neg\phi$. But since the material conditional is

³⁵ UTC consequence is easily seen to be transitive. Suppose $\phi \Vdash \psi$. Then:

1. $\sigma[\phi] = \sigma[\phi][\psi]$

Suppose also that $\psi \Vdash \theta$. Then (taking one particular instance):

2. $\sigma[\phi][\psi] = \sigma[\phi][\psi][\theta]$

Chaining together the two equalities, we obtain $\sigma[\phi] = \sigma[\phi][\psi][\theta]$. But applying the first identity, we then obtain $\sigma[\phi] = \sigma[\phi][\theta]$, so $\phi \Vdash \theta$.

introduced into update semantics as a notational variant on the usual negation-disjunction combination, it follows that $\Vdash \Box\phi \vee \Box\neg\phi$, which is the omniscience conclusion. Update semantics escapes this time by denying the validity of modus tollens. Suppose $\phi \Vdash \psi$, and ψ is a test failed by some state σ , but passed by σ updated with ϕ . Then σ updated with $\neg\psi$ will remain unchanged, and need not be a fixed point of ϕ .

14 Update, Truth, and Disunity

The considerations of the previous section show that the Omniscience Argument, perhaps already easily evaded via less controversial means, fails to bring down even a system so amenable to the argument's starting assumptions as update semantics, because of the logical irregularities of update semantics. One possible moral, then, is that the Omniscience Argument just isn't a very good argument. But I think there is more to be said. How *can* a logic deny the validity of proof by cases or modus tollens? If one agrees that things are either a p way or a q way, and that if they are a p way they are an r way, and also if they are a q way they are an r way, how can one not agree that things are an r way?

The canonical answer on the part of update semantics is that it is no part of update semantics to track *the way the world is*, but rather to track *the way agents are committed to things being*. Given an information state σ , if $\sigma[\phi] = \sigma$, we say $\sigma \Vdash \phi$, and that σ is committed to ϕ . But this notion of commitment should not be expected to support proof by cases or modus tollens, for the sorts of reasons given above.

However, consider the following bridge principle, designed to bring out conceptual connections between agent doxastic commitment and truth:

(Bridge): If A is committed to ϕ and things are as A takes them to be, then ϕ is true.

If the strategy of update semantics is to rationalize the failures of proof by cases and modus tollens by gesturing to the displacement of truth by commitment in the conceptual foundations of the semantics, then a supporter of update semantics needs to reject Bridge if he is not to have truth, and hence proof by cases and modus tollens, and hence the Omniscience Argument, thrust back upon him. And indeed the dominant understanding of update semantics does reject Bridge, by way of rejecting the entire truth-conditional project of associating with each sentence a propositional content which serves to bifurcate the ways the world could be into two classes. Thus Veltman says:

Sentences of the form *might* ϕ are not persistent; they do not express a proposition; their informational content is not context independent. If you learn a sentence ϕ of $[\mathcal{L}]$, you learn that the real world is one of the worlds in which the proposition expressed by ϕ holds: the real world is a ϕ -world. But it would be nonsense to speak of the '*might* ϕ -worlds'. If ϕ might be true, this is not a property of the world but of your knowledge of the world. (Veltman (1996), 233)

Dynamic semantics (here in the particular guise of update semantics), like supervaluationism, threatens the truth-conditional conception of semantics by trying to place an alternative semantic notion at the center of the theory. As with supervaluationism, I want to suggest that the **DOT** can let us have the best of both worlds, retaining the technical innovations of the dynamics and the conceptual foundations of truth-conditional semantics.

The epistemic modal \Box can be understood as a setwise **VF**. The validity of the Motan principle under update semantics already shows that such a **VF** would meet (both the inferential and the biconditional version of) the disquotational test set out above. The preservation of classical logic on the non-modal fragment \mathcal{L} of the update system yields a plausible case that the **VF** meets the inferential test. And epistemic modals are well-suited for accounting for the integration between truth and other aspects of our practice that form the third test for a legitimate **VF**. Consider the norm of assertion: ‘Assert only that which is true_{\Box^u} ’. A claim is true_{\Box^u} if it holds in every world compatible with what the asserting agent knows/believes, so this norm amounts to the norm of assertion: ‘Assert only that which you know/believe’ (with the choice between knowledge and belief resting on the prior choice of whether the information states in an intended model track consistency with what is known or with what is believed), which is not an implausible norm.

If we see update semantics as introducing a setwise **VF** ‘ true_{\Box^u} ’, then the notion of truth becomes ambiguous: in addition to the setwise ‘ true_{\Box^u} ’, there is also a pointwise **VF**, representing the holding of a claim at a particular point. Because update semantics, unlike supervaluationism, allows its setwise **VF** to vary in the set with respect to which it is evaluated, there is a non-trivial question about how (in the manner of the ‘push down’ strategy of Section 8) to associate a set with each point for use in evaluating modal claims. Given the setwise and pointwise **VF**’s, we can then introduce two notions of logical consequence, mirroring the global and local consequence relations of supervaluationism. As with supervaluationism, the global consequence relation will have some non-classical features (here, failure of proof by cases and modus tollens), but will support the inferential version of the disquotational principle for the setwise **VF** (hence, the Motan principle). The local consequence relation, on the other hand, will be fully classical, but will (assuming sets are associated with points in some way more generous than simply mapping each point to its own singleton) fail to support the Motan principle.

Given a **DOT** perspective on update semantics, the Bridge principle becomes four ways ambiguous, depending on whether each of ‘things being as *A* takes them to be’ and ‘ ϕ being true’ is understood setwise or pointwise. But seeing truth as setwise yields a natural understanding of why a **VF** might not support proof by cases or modus tollens, so for the Bridge principle to do the work in forcing update semantics into acceptance of the Omniscience Argument, it will have to be the following mixed reading:

If *A* is committed to ϕ and things are *global* as *A* takes them to be, then ϕ is $\text{true}_{\text{local}}$.

There is, of course, good reason to reject this mixed reading. But, importantly, rejecting *it* does not carry a pressure toward rejecting the utility of truth-conditionality in understanding update semantics, or a pressure toward the rejection of propositional meaning as a central feature of semantic theorizing in a dynamic framework. Update semantics, seen through a **DOT** lens, offers no threat to truth-conditionality, but simply provides propositions which can be evaluated set-wise (with respect to an information state) for (global) truth, together with object-language operators reporting on the global truth feature.

I mentioned earlier that readers will undoubtedly have swiftly identified some flaw or other in the Omniscience Argument. Relatively few of these flaws, I suspect, will have involved situating the argument in update semantics or some other dynamic framework. I want, therefore, to close with a brief survey of four responses to the Omniscience Argument, no longer tied to its development in update semantics, in order to show that the insights that these responses bring can be seen also as manifestations of the underlying thought that the Omniscience Argument turns on the interaction between global and local notions of truth.

1. Perhaps the most obvious response to the Omniscience Argument is to hold that the appeal of the Motan principle relies on a scope equivocation: that the claim ‘if ϕ , then it must be the case that ψ ’ is plausible (given the background assumption that ϕ entails ψ) when the modal is given wide scope and the claim is read as ‘it must be that: if ϕ , then ψ ’, but implausible when the modal is given narrow scope and the claim is read as ‘if ϕ , it must be that: ψ ’. Since it is the narrow scope reading which is needed for the Omniscience Argument, the argument is thus defused.

A serious difficulty in the scope response, however, is that there is no syntactic evidence that the modal does, indeed, take wide scope. On the contrary, there is evidence that it takes narrow scope. The modal ‘must’ can, for example, be used in a separate sentence from the hypothetically introduced antecedent, as in ‘Suppose that ϕ . Then it must be that ψ .’ Absent syntactic evidence for the wide-scoping, the scope response threatens to become nothing more than the insistence that the inferential behaviour of the modal matches that of a conventionally (i.e., pointwise) understood wide-scoped modal. As a claim about the inferential potential of ‘must’ in this context, such a claim is, of course, correct, but it is a claim also yielded by the update semantics treatment of ‘must’, and by the setwise interpretation of that semantics. Update semantics, in fact, makes equivalent $\phi \Rightarrow \Box\psi$ and $\Rightarrow (\phi\Box\psi)$ – the dynamic perspective, taking meanings to be transition rules between information states, make tradition questions about the (semantic) scope location of a particular logical operator difficult to pin down.

2. A second plausible response to the Omniscience Argument is to claim that the apparent truth Motan principle in fact reflects features of *assertability*, rather than truth. If ϕ entails ψ , and this entailment is known to an agent *A*, then if the agent *A* is in a position to assert ϕ (rather than ϕ simply being true), then

he is also in a position to assert that it must be the case that ψ , since his knowledge/belief that ϕ (guaranteed by its assertability for him) will couple with his awareness of the entailment fact to put him in a position to assert the (egocentric) epistemic modalization of ψ . The Motan principle, properly understood, thus reflects a conditional connection between assertabilities, rather than a conditional connection between truths. But, of course, such a conditional connection is insufficient to run the Omniscience Argument.

But assertability plausibly meets the constraints on being a **VF**, and the egocentric epistemic ‘must’ modal is then plausibly a truth predicate for that setwise **VF**. So this response to the Omniscience Argument can, in fact, be seen as an endorsement of the truth, in a setwise sense of truth, of the Motan principle, coupled with inferential constraints on the use of that truth in the Omniscience Argument. Seen in this light, the response becomes another instance of the **DOT** strategy.

3. A third response to the Omniscience Argument rejects the appeal to the law of the excluded middle, used in providing the premise $\lceil \phi \vee \neg\phi \rceil$ used in moving via proof by cases from the two conditionals to the final omniscience conclusion. The rejection of the law of the excluded middle may be particularly tempting when the Omniscience Argument appears in its guise as the fatalist argument in free will debates – here certain pictures of the ‘open future’ may see the rejection of excluded middle for future contingents as the key to evading fatalism.

But failure of the law of the excluded middle once a truth predicate is added is an expected feature of setwise notions of truth – thus, for example, supervaluationism rejects the universal truth of sentences of the form ‘ ϕ is supertrue or $\neg\phi$ is supertrue’. If the rejection of the law of the excluded middle derives from the status of a setwise-inflected version of the law, such as this, then its failure is already predicted and explained by the **DOT** picture.

4. A fourth response to the Omniscience argument involves defanging the modal ‘must’ by giving it a reading weak enough that the conclusion of the Omniscience Argument can be unproblematically accepted. Thus, for example, Dummett (1964) allows, in discussion of fatalism, the inference from ‘you will not be killed’ to ‘if you do not take precautions, you will not be killed’, but then disallows the inference from this conditional to ‘any precautions you take will have been superfluous’. This move by Dummett can be seen as an endorsement of a version of the Motan principle, with ‘will’ serving as the relevant modal, combined with giving that modal a reading weak enough that consequences (such as those for practical reasoning) which normally follow from modals are now blocked.

Defanging the modal amounts to treating it as a modal ranging over only a very small subset of possible worlds (perhaps including only the actual world), and hence represents a resistance toward the introduction of a setwise operator, or the imposition of a constraint on that setwise operator which forces its inferential behaviour to match that of pointwise operators. Given such a resistance or such a constraint, the conclusion of the Omniscience Argument then reduces to the pointwise tautology that ϕ or $\neg\phi$, and poses no threat.

Acknowledgments I am grateful to Nicholas Asher, Anthony Gillies, and Rob Stainton for comments on an earlier draft of this paper. Portions of this material were presented to audiences at the University of Michigan and Princeton University, and benefited from comments and questions received there.

References

- Asher, N., Dever, J., & Pappas C. 2005 Supervaluations debugged.
- Bain, A. 1874 *Logic: Deductive and inductive*. D. Appleton and Company.
- Bentham, E. 1773 *An introduction to logick*. W. Jackson & J. Lister.
- Boole, G. 1952 The mathematical analysis of logic, being an essay towards a calculus of deductive reasoning. In R. Rhees (Ed.), *Collected logical works, Volume I: Studies in logic and probability*. The Open Court Publishing Company.
- Bradley, F. 1883 *The principles of logic*. William Clowes and Sons.
- Coppee, H. 1857 *Elements of logic; designed as a manual of instruction*. J.H. Butler and Company.
- Davidson, D. 1984a belief in the basis of meaning. In *Inquiries into truth and interpretation*. Oxford University Press.
- Davidson, D. 1984b Radical interpretation. In *Inquiries into truth and interpretation*. Oxford University Press.
- Davidson, D. 1984c Truth and meaning. In *Inquiries into truth and interpretation*. Oxford University Press.
- Dever, J. 2003 Review of Zoltan Gendler Szabo, problems of compositionality. *The Philosophical Review*, **112**, 254–258.
- Dever, J. 2006 Compositionality. In E. Lepore & B. Smith (Eds.), *The oxford handbook of philosophy of language*. Oxford University Press.
- Devey, J. 1854 *Logic; or the science of inference: A systematic view of the principles of evidence, and the methods of inference in the various departments of human knowledge*. Henry G. Bohn.
- Dugundji, J. 1940 Note on a property of matrices for Lewis and Langford's calculi of propositions. *The Journal of Symbolic Logic*, **5**, 150–151.
- Dummett, M. 1959 Truth. *Proceedings of the Aristotelian Society*, **59**, 141–162.
- Dummett, M. 1964 Bringing about the past. *Philosophical Review*, **73**, 338–359.
- Fara, D. G. 2003 Gap principles, penumbral consequence, and infinitely higher-order vagueness. In J. Beall (Ed.), *Liars and Heaps: New essays on the semantics of paradox*. Oxford University Press.
- Fine, K. 1975 Vagueness, truth, and logic. *Synthese*, **30**, 265–300.
- Fodor, J., & E. Lepore 1996 What cannot be valued cannot be valued and it cannot be supervalued either. *Journal of Philosophy*, **93**, 516–535.
- Frege, G. 1967 Begriffsschrift. In J. van Heijenoort (Ed.), *From Frege to Godel: A sourcebook in mathematical logic, 1879–1931*. Harvard University Press.
- Frege, G. 1997 Negation. In M. Beaney (Ed.), *The Frege reader*. Blackwell Publishers.
- Geach, P. 1956 Good and Evil. *Analysis*, **17**, 33–42.
- Godel, K. 1986 On the intuitionistic propositional calculus. In S. Feferman, J. W. Dawson, S. C. Kleene, G. H. Moore, R. Solovay, & J. van Heijenoort (Eds.), *Kurt Godel: Collected works Volume I, Publications 1929–1936*. Oxford University Press.
- Horwich, P. 1990, *Truth*. Oxford University Press.
- Kant, I. 1929, *Critique of pure reason*. Macmillan and Co., Ltd.
- Kripke, S. 1980, *Naming and necessity*. Harvard University Press.
- Lepore, E. 1982 In defense of davidson. *Linguistics and Philosophy*, **5**, 277–294.
- Lepore, E. 1983 What model theoretic semantics cannot do?. *Synthese*, **54**, 167–187.
- Lepore, E. B. Loewer 1983 Three trivial truth theories. *Canadian Journal of Philosophy*, **13**, 433–447.
- Lewis, C. I. Langford C. H. 1932, *Symbolic logic*. Dover Publications.

- Lotze, H. 1884 *Outlines of logic and of encyclopaedia of philosophy*. Ginn and Company.
- Lukasiewicz, J. 1967a On the notion of possibility/on three-valued logic. In S. McCall (Ed.), *Polish Logic: 1920–1939*. Clarendon Press.
- Lukasiewicz, J. 1967b Philosophical remarks on many-valued systems of propositional logic. In S. McCall (Ed.), *Polish Logic: 1920–1939*. Clarendon Press.
- Lynch, M. 2004 Truth and multiple realizability. *Australasian Journal of Philosophy*, **82**, 384–408.
- MacColl, H. 1897 Symbolic reasoning II. *Mind*, **6**, 493–510.
- Morgan, A. D. 1966 On the syllogism. In P. Heath (Ed.), *On the syllogism, and other logical writings*. Routledge and Kegan Paul.
- Sutton, J. 2005 Stick to what you know. *Nous*, **39**, 359–396.
- Szabo, Z. G. 2000 *Problems of compositionality*. Garland Publishing, Inc.
- van Eijck, J. Kamp H. 1997 Representing discourse in context. In J. van Benthem A. ter Meulen (Eds.), *The handbook of logic and language*. Elsevier, pp. 179–237.
- Veltman, F. 1996 Defaults in update semantics. *Journal of Philosophical Logic*, **25**, 221–261.
- Venn, J. 1880 Review of Frege's conceptual notation. *Mind*, **5**, 297.
- Whately, R. 1825 *Elements of logic*. Scholars' Facsimiles and Reprints.
- Williamson, T. 1994 *Vagueness*. Routledge.
- Williamson, T. 2000 *Knowledge and its limits*. Oxford University Press.
- Wright, C. 1987 Further reflections on the sorites paradox. *Philosophical Topics*, **15**, 227–290.
- Wright, C. 1992 *Truth & objectivity*. Harvard University Press.

Descriptions, Negation, and Focus

Michael Glanzberg

Abstract This paper argues that some familiar cases of interaction between definite descriptions and negation are not best analyzed as scope interactions. Attention to the role of focus, and a number of related semantic and pragmatic factors, shows that the cases give no evidence of scope interaction. However, these factors can generate an illusion of scope. In particular, focus can generate illusions of scope, which may lead us to think sentences display scope ambiguities they do not. These conclusions offer limited support to non-quantificational treatments of definite descriptions.

Keywords Definite Description · Negation · Focus · Quantifier · Quantifier Scope · Presupposition

One of the mainstays of the theory of definite descriptions since Russell (1905) has been their interaction with negation. In particular, Russellians, who advocate the view that definite descriptions are a kind of quantifier, point to these interactions as evidence in favor of their view. The argument runs roughly as follows:¹

- (1) a. Definite descriptions show a number of important interactions with negation (as well as with other quantifiers, with intensional contexts, etc.).
- b. These interactions are best analyzed as *scope interactions*.
- c. Such scope interactions are clearly and easily predicted if we treat definite descriptions as quantifiers. They are not predicted at all, or only by roundabout means, if we do not treat definite descriptions as quantifiers.
- d. Hence, we have evidence in favor of the quantificational treatment of definite descriptions.

M. Glanzberg

Department of philosophy, University of California, Davis, One Shields Avenue, Davis, CA 95616, USA

¹ For instance, Neale (1990, p. 49) writes, ‘Since descriptions are treated as quantifiers ... all sorts of interesting scope interactions are predicted; not just with negation and other quantified noun phrases, but also with various types of nonextensional operators (Chapter 4).’

This is a powerful argument, and potentially far-reaching. Even so, focusing entirely on the case of negation, I shall argue in this paper that it is not sound. Premise (1a) is certainly true, and an observation of great importance. It is primarily premise (1b) that I shall challenge here, for the special case of negation. I shall argue that the interactions we see in this case are not scope interactions. With this, of course, premise (1c) and the conclusion become equivocal, at best. If the interactions we see are not scope interactions, then a theory that explains them as scope interactions finds no great support.

The main goal of this paper is thus to take a fresh look at some long-standing data. This data, I shall argue, gives us no evidence of scope interaction when it comes to definite descriptions and negation. The interactions we see are the result of a number of factors I shall explore here: factors which can generate an *illusion* of scope. Understanding how this illusion arises, and why it is not really a genuine observation of scope, will help us to better sort out the data before us. The main source of the illusion is the phenomenon of *focus* (corresponding, roughly, to where stress falls in a sentence). Focus can lead us to think sentences display scope ambiguities they do not, and more generally, focus can lead to illusions of scope. Careful attention to the role of focus, and a number of related semantic and pragmatic factors, will give us a better understanding of the data before us. With it, we will see that the evidence from negation gives us no indication of scope for definite descriptions. Thus, in the case of negation, the argument in the form of (1) does not succeed.

I have been careful to put my conclusion in terms of what a certain body of evidence indicates. I have not proposed to argue that definite descriptions do not take scope. Though I shall argue that a certain rather narrow body of evidence is compatible with this strong conclusion, this evidence does not preclude the quantificational analysis, which treats definite descriptions as on par with canonical scope-taking operators. There are a number of reasons for approaching the stronger conclusion cautiously. Scope is an issue that relates to a huge range of linguistic phenomena, including some very general ones about the nature of human languages. Thus, no limited range of data should lead us to jump to conclusions. Furthermore, scope is as much a *theoretical* issue as a descriptive one. What takes scope, and when and why, interacts with a great many theoretical aspects of linguistic theory. Deciding what takes scope can require deciding some high-level issues in linguistic theory as much as deciding what a range of data shows. The points I shall make here about negation show another way in which questions about scope can be difficult. Because of the possibility of illusions of scope, we must be careful about the data itself, before we get to the wider theoretical questions.

I shall begin this paper with an overview of the data on descriptions and negation in Section 1. I shall then present some background on the interpretation of definite descriptions, and related issues about presupposition, in Section 2. In Section 3, I shall return to the data. I shall argue there that cases of proper definite descriptions do not reveal any scope interactions with negation. I shall also show there how focus can create an illusion of scope in cases like these, which might lead us to think we see scope interactions between descriptions and negation where there

are none. I shall extend this argument to cases of improper definite descriptions in Section 4, and to negative quantifiers in Section 5. I shall close with a brief concluding Section 6.

1 Initial Observations

Our starting point is the observation that sentences like the following show a scope ambiguity:

- (2) Every politician is not corrupt.

This sentence can mean either that every politician fails to have the property of corruptness, or that it is not the case that all politicians have this property. To bring out the difference, note that if in fact half the politicians are corrupt and half not, the claim is false on the first reading, and true on the second.

I am taking it for granted that we have good evidence that (2) really is ambiguous. At least, speakers can see the two readings, with the corresponding truth-value judgments (or be induced to see them with only ‘harmless’ prompting). As we will discuss certain sorts of evidence at length, it should be noted that the evidence here is indirect. We do not have direct evidence of ambiguity, so much as a hypothesis which explains a range of data. That the ambiguity hypothesis is right has been challenged; but nonetheless, I shall take it as our starting point here.²

I shall take it as given that the ambiguity in (2) is a scope ambiguity. Its two readings are:

- (3) Every politician is not corrupt.
- a. $\forall x(P(x) \longrightarrow \neg C(x))$ ($\forall > \neg$)
 - b. $\neg[\forall x(P(x) \longrightarrow C(x))]$ ($\neg > \forall$)

The two glosses in (3) capture the truth conditions of the two readings. It is, of course, a much more substantial theoretical claim that the sentence is in fact structurally ambiguous, and it corresponds to two distinct logical forms, showing scope relations between the quantifier and negation along the lines of (3). Indeed, it is even a substantial claim that this is what a scope ambiguity is. I shall generally assume scope ambiguities are structural ambiguities, and that scopally ambiguous sentences are associated with distinct logical forms. But it should be noted that we have already crossed the line between data and very substantial theory.³

² Though they do not dwell on this particular kind of sentence, representative argument against scope ambiguity hypotheses include Kempson and Cormack (1981), Reinhart (1979, 1983), Wilson (1978). A defense of standard ambiguity claims can be found in Chierchia and McConnell-Ginet (1990), while more recent criticism is offered by Pietroski and Hornstein (2002). Of course, if you do not accept that (2) is ambiguous, then you are not likely to see ambiguities with definite descriptions either, and so you may take the main claim of this paper as a given.

³ For our purposes here, we may assume that the logical forms in question will mark scope differences in something like the way they are marked in (3). For discussion of how this may be

Niceties of what counts as theory aside, quantifiers are among our primary examples of scope-taking operators, and examples like (2) show that they enter into non-trivial scope interactions with other operators. Judgments supporting ambiguity are our main source of data on scope interaction. With canonical quantifiers, like the universal quantifiers *every* or *all*, judgments supporting scope ambiguity with negation are widespread. However, when we come to definite descriptions, they are not. In particular, we do not see such judgments with:

- (4) The president of the United States is not corrupt.

This does not appear to be ambiguous. In particular, the kinds of judgments supporting ambiguity for (2) are not to be found for (4). When speakers come to see (2) as ambiguous, sometimes after being offered a range of scenarios and having truth-value judgments elicited, they still do not see (4) as ambiguous.

The lack of ambiguity for definite descriptions and negation is quite general. We see it just as much in:

- (5) John did not read the book.

We likewise see no ambiguity in most cases involving definite descriptions and negative quantifiers like *no one*. We observe none in any of:

- (6) a. The president of the United States likes no one.
b. No one likes the president of the United States.

In contrast, replacing the definite description with a quantifier can reinstate ambiguity in some cases, such as:

- (7) No one likes two great American novels.

The general pattern seems to be that with quantifiers and negation we get at least some judgments of ambiguity, while for similar cases with definite descriptions, we do not. As I mentioned, I shall take the evidence of ambiguity in these cases to be evidence of scope ambiguity. We thus have evidence of scope ambiguity for quantifiers/negation interactions, while our initial glance at the data shows us none for definite description/negation interactions.

Why then has the idea that there are such scope ambiguities with definite descriptions and negation become so widespread? Because of the king of France, of course. Since Russell's seminal work (e.g. Russell, 1905, 1919; Whitehead & Russell, 1927), it has often been argued that we see scope ambiguities when we have improper definite descriptions. Hence, notoriously, Russellians see a scope ambiguity in:

implemented in syntactic theory, see Heim and Kratzer (1998), May (1985), or my survey (2006). It is a point of debate whether or not scope ambiguities in natural language really are syntactic ambiguities; see Jacobson (2002) for a critical discussion. Among those who do take them to be mainly syntactic, it is a point of contention whether the syntax of logical form completely suffices to disambiguate scope. May (1985) argues it does not. For detailed discussion of sentences like (2), see Acquaviva (1993) or Büring (1997), among places.

(8) The king of France is not bald.

(An improper definite description is one for which there is no unique object answering to the description. The only improper definite descriptions I shall be concerned with here are those for which no object answering to the description exists at all; I shall not be concerned with failures of uniqueness.) To Russellians, as is well-known, this sentence is ambiguous between a true reading, in which the description takes narrow scope with respect to negation, and a false reading, in which the description takes wide scope with respect to negation.

It has been common to put this in terms of the argument I sketched in (1). It is argued both that the judgment of ambiguity for (8) shows us an interaction between negation and definite descriptions (premise 1a), and that it is a scope ambiguity (premise 1b).⁴ Whether or not such an argument should be attributed to Russell himself is less clear. Russell does indeed claim that a definite description can take one of two scopes (which he calls ‘primary’ and ‘secondary’ occurrences). Initially, in Russell (1905, p. 53), he simply says that the distinction allows us to ‘deal with the question of whether the present King of France is bald or not bald, and generally with the logical status of denoting phrases that denote nothing.’ He later says specifically that the sentence is ambiguous (Russell, 1919; Whitehead and Russell, 1927). Interpreting Russell on this matter is not entirely straightforward, as Russell seems more concerned with the syncategorematic structure of the proposition derived from a definite description, and the way it allows certain metaphysical and epistemological puzzles to be addressed, than he is with the analysis of natural language. Indeed, he distances himself from an analysis of natural language in Russell (1957).

Whether or not it is due to Russell, appeals to sentences like (8) to support the kind of argument I sketched in (1) have become standard. I shall attempt to show they are mistaken. We have already noted that until we encounter difficult cases like (8), we do not see any evidence of scope ambiguity with negation and definite descriptions. We do certainly see something important in the cases of improper definite descriptions. I shall argue that it is not scope, and it is cases like these which will bring out how an illusion of scope can be created.

2 Approaches to Definite Descriptions

Before pressing on to the main arguments, it will be helpful to stop and review some ideas about definite descriptions. In particular, I shall review two common theories of definite descriptions. The data under examination is often offered as favoring one over the other, so it will be useful in looking at the data to have these theories clearly in mind. I shall also review some ideas about the presuppositions of definite descriptions, which will be important to some of the data we will consider.

⁴ I believe this kind of argument is in Neale (1990, Chapter 4). Neale does phrase things cautiously, saying that it is ‘at least arguable’ that sentences like these have the readings Russellians claim, and that ‘presumably’ they will be captured by the scope behavior of the definite description.

2.1 Interpreting The

Our starting point, which really does go back to Russell, is with the truth conditions of sentences with definite descriptions. Compare:

- (9) a. The F is G .
 b. $\exists x(F(x) \wedge \forall y(F(y) \rightarrow x = y) \wedge G(x))$

Russell argued that the truth conditions of (9a) are given by (9b). When it comes to *proper* definite descriptions, for which there is a unique F , this is widely accepted, and I shall not challenge it.

The theory of descriptions as Russell presents it makes definite descriptions syncategorematic: there is no constituent in the logical form of a sentence corresponding to the definite article.⁵ Modern neo-Russellian theories of descriptions generally do not accept this conclusion. For instance, Neale (1990) proposes that though (9b) does give the truth conditions of (9a), the logical form is given by a construction involving a *restricted quantifier*:

- (10) [*the* $x: F(x)$] $G(x)$

The truth conditions of this form are still as they are given in (9).⁶

On neo-Russellian views like Neale's, there is a constituent in logical form [*the* $x: F(x)$] corresponding to *the* F . But following the core Russellian idea, this is interpreted as a *quantifier*. This means that *the* F is not interpreted as an individual, any more than \forall or \exists are. There are a number of formal devices for interpreting structures like (10). In generalized quantifier theory, for instance, the restricted quantifier [*the* $x: F(x)$] is interpreted as essentially a set of sets: $\{X \mid |F| = 1 \wedge |F \setminus X| = 0\}$. The full details of generalized quantifiers will not be of concern to us here.⁷ What will be important about this kind of analysis is that definite descriptions are interpreted as quantifiers, which take scope in much the same ways as canonical quantifiers do, and that as such, they are not interpreted as individuals. Let us refer to the family of views of definite descriptions which follow this neo-Russellian route as the *quantificational approach*.

The quantificational approach has become something of an orthodoxy in philosophy of language. Like all orthodoxies, it has not gone unchallenged, for instance,

⁵ In Russell (1905, 1919), he describes the theory as giving a way of associating the surface form of a sentence with a logical form, though that form does not contain any constituent directly corresponding to the definite article. In Whitehead and Russell (1927), descriptions are directly introduced as defined symbols of a formal language. (I am indebted to Fara (2001) for clarifying some of these points.)

⁶ Other authors who endorse similar proposals include Sainsbury (1979) and Sharvey (1969).

⁷ The idea that quantifiers are to be interpreted as sets of sets goes back to Frege (e.g. Frege, 1879, 1891). Contemporary development of the idea begins with Montague (1973), and then Barwise and Cooper (1981), Higginbotham and May (1981), and Keenan and Stavi (1986). The subsequent literature is huge, and the view of quantifiers they pioneered has become a mainstay of modern semantics. A survey targeted at philosophers is given in my (2006).

by Fara (2001). It is much less of an orthodoxy in the linguistics literature, which contains a number of alternative analyses of definite descriptions, none of which seems to have achieved the status of a received view. One important class of competitors to the quantificational approach is that of dynamic theories, which interpret definites more or less as variables (in a setting which makes important modifications to the standard treatment of variables and binding from first-order logic).⁸ Another important alternative treats definite descriptions as semantically structured expressions which pick out individuals.

This latter sort of approach will facilitate comparison with the quantificational approach, and motivate some further assumptions about the behavior of definite descriptions, so I shall sketch it in some more detail, and refer back to it as our discussion progresses. I should stress, though, that I choose it as our non-quantificational alternative only because it facilitates these comparisons. I shall not argue one way or the other whether it or any of its competitors are superior in the end.

The non-quantificational view I have in mind treats *the F* as a semantically structured phrase, whose nominal *F* is interpreted as a predicate in the usual way. But unlike the quantificational approach, this view interprets the entire definite description as picking out an individual:

$$(11) \llbracket \textit{the } F \rrbracket = \begin{cases} \text{the unique element of } F & \text{if } |F| = 1 \\ \text{undefined} & \text{otherwise} \end{cases}$$

($\llbracket \alpha \rrbracket$ is the interpretation, or semantic value, of α .) It is important to stress that for proper definite descriptions, where there is exactly one *F*, (11) gives exactly the same truth conditions to *the F is G* as we saw in (9). But it does so in a different way. On the neo-Russellian quantificational approach, (9a) has a logical form like (10), with the definite description interpreted as a quantifier, binding a variable. On the treatment of (11), *the F* simply contributes an individual to the truth conditions. It is not interpreted as a quantifier, nor do we need a separate quantifier and variable in the logical form. It is customary to trace this sort of definition back to Frege (1893), so let us call this the *Fregean* approach to definite descriptions.⁹

Though according to this treatment, definite descriptions contribute individuals to computations of truth conditions, it is important to stress that they do so in a

⁸ These theories originate with Heim (1982) and Kamp (1984), and then Groenendijk and Stokhof (1991). For more recent surveys, see van Eijck and Kamp (1997) and Kadmon (2001).

⁹ Fregean treatments of the definite article like (11) are given in Heim (1991) and Heim and Kratzer (1998). An extended defense is given in Elbourne (2005).

The label 'Fregean' is in some ways unfortunate, and might tend to mislead. Most importantly, it has nothing to do with the dispute between Fregeans and direct reference theorists in the theory of reference. In other work (Glanzberg, 2007), I have opted to call it the '*e*-type approach', to emphasize that its main feature is interpreting definite descriptions as picking out individuals, and thus interpreting them differently from quantifiers. I avoid that terminology here, as the apparatus of types is not relevant to this discussion. Regardless, the label 'Fregean' has become more or less become standard.

way very different from names or pronouns (the sorts of expressions for which direct reference theories are an option). First of all, definite descriptions on the Fregean approach are semantically structured, with the nominal *F* playing a significant role. Because of this, definite descriptions on this view are not rigid. Though they contribute an individual (unlike a quantifier), they can contribute different individuals in different worlds. Moreover, it is possible to bind variables in the nominal *F*, e.g. in:

- (12) Every man respects the woman he marries.

Thus, the description can contribute different individuals not just in different possible worlds, but relative to different assignments of values to variables.

The quantificational approach builds in the idea that definite descriptions take scope in many environments (as canonical quantifiers do). It thus predicts a range of scope ambiguities for sentences with definite descriptions, though in many cases, there may be no truth-conditional differences between the readings of these scopally ambiguous sentences. As we discussed with respect to argument (1), finding evidence of scope ambiguity thus speaks in favor of the quantificational approach. The Fregean approach, on the other hand, predicts that there will never be truth-conditionally distinguished scope ambiguities generated by definite descriptions. Not finding evidence of scope ambiguity thus supports the Fregean approach, or more carefully, it undercuts one line of argument in favor of the quantificational over the Fregean approach. It is compatible with the Fregean approach that definite descriptions never take scope. If supplemented with the right ideas about logical form, it is compatible with the Fregean view that descriptions take scope much as the quantificational approach has it, or even that descriptions always take fixed scope (e.g. narrow scope), but that these scope assignments never matter truth-conditionally. For this sort of reason, as I mentioned, deciding issues of scope can be a highly theoretical matter. But still, we see, evidence of scope ambiguity, or lack of it, can be significant to choosing between the quantificational and Fregean options.

2.2 *Presupposition*

The Fregean analysis of descriptions I gave in (11) makes the semantic value of *the F* undefined if there is no unique *F*. This assumes that definite descriptions carry a *presupposition* of existence and uniqueness. Whether definite descriptions carry presuppositions, and if so what they are, is highly controversial. So, several comments about this assumption are in order.

The Fregean analysis invites the sort of semantic presupposition reflected in (11), and provides an easy way to implement it, while the quantificational approach does not. But in fact, the issues of presupposition and of whether Fregean or quantificational approaches are preferable are largely independent of each other. It is possible to avoid the presuppositions of (11) on a Fregean approach, by fixing that *the F* is

G is *false* if the description is improper.¹⁰ Likewise, it is possible to write semantic presuppositions into the interpretations of quantifiers, including one which otherwise functions like we see in (10). Technically, for both the quantificational and Fregean approaches, we can either have presuppositional or non-presuppositional meanings for *the*.¹¹

Though no doubt the issue is still hotly controversial, I take it that in some cases, we do get fairly strong judgments of *infelicity* for sentences with improper definite descriptions. For old standbys like *The king of France is bald*, many speakers see an infelicity. Finding absolutely reliable tests for infelicity is not a simple matter. Sometimes, as Strawson (1950) suggested, it will go with refusal to give truth-value judgments. Sometimes, as von Stechow (2004) suggests, it will go with a response like *Hey wait a minute, France does not have a king*. Sometimes, as I proposed in my (2005b), it will go with an unwillingness to make certain kinds of assessments or indirect speech reports without initiating a *repair*. Regardless of which tests we choose, I think that in some simple cases, we have fairly firm evidence of presupposition.¹²

A presuppositional account of definite descriptions captures this data nicely. It also faces some well-known difficulties. For one, it is not easy to state entirely

¹⁰ There are some technical complications for doing this. One way to do it is to introduce a kind of 'default object' of which no simple predicate holds. Other options including type shifting, or departures from classical logic. Some of these possibilities have been explored in the literature on choice functions (which is generally more concerned with indefinite than definite descriptions), notably by Reinhart (1997) and Winter (1997). Though it works in the setting of dynamic semantics, the comparisons of presuppositional and non-presuppositional treatments of descriptions in van Eijck (1993) is also noteworthy.

¹¹ As Heim and Kratzer (1998) note, there are some quantifiers which fairly clearly seem to carry presuppositions, including *both* and *neither*. There remains a lively dispute about whether all quantifiers carry a presupposition of non-empty domain. The idea that they do goes back to Strawson (1952), and has essentially been defended by Diesing (1992). Alternatively, it has been argued that some sub-classes of quantifiers carry such a presupposition, e.g. by Barwise and Cooper (1981). The view that it is exactly the so-called strong determiners that do so is developed by de Jong and Verkuyl (1985). The position that quantifiers do not generally carry presuppositions is defended by Lappin and Reinhart (1988).

¹² Abbot (2004, p.127) goes so far as to say, 'Since the publication of Strawson's paper, there has been fairly unanimous support for the intuitions he expressed, but less agreement on how best to give an account of these facts.' I suspect a number of critics of Strawson, from Sellars (1954) to Neale (1990), might well have taken their points to cut deeper than that.

Definition (11) not only gives definite descriptions presuppositions, it makes these presuppositions conventional in nature, triggered by the semantics of the definite article. I think this is right, but I hasten to add that I do not think all of the presuppositions that have been discussed in the literature are this way. (For a recent survey of some ways presuppositions can be triggered, see Kadmon (2001). These issues are also discussed in my (2005b).) On the other extreme, there has been a long tradition of seeking to explain away presupposition as a combination of implicature and entailment. A survey especially sympathetic to this reductionist approach is given by Levinson (1983). Though I am not a proponent of the reductionist approach, the issue is not really one that is of importance here. An alternative semantics for definite descriptions, together with a different account of the source of their presuppositions, could serve the argument I shall give here equally well.

accurately what the presuppositions of definite descriptions are. Though the existence presupposition is relatively straightforward, the uniqueness presupposition is not. The right account of uniqueness has been the subject of intensive investigation over the years. However, our concerns with negation will relate only to existence presuppositions, so we may put the matter of uniqueness presuppositions aside.¹³

There are some commonly voiced objections to the presuppositional account of definite descriptions that I believe are really no more than reminders that a full theory of presupposition needs to include an account of presupposition projection: how presuppositions are inherited by larger sentences from their parts. (Indeed, facts about projection form the core around which modern theories of presupposition are built.) For instance, one sometimes sees cases like the following offered as objections to a presuppositional treatment of definite descriptions:

- (13) a. If Burkina Faso has a king, then the king of Burkina Faso is very likely worried about the situation in Niger.
 b. Ponce de Leon thought the fountain of youth was in Florida.

(Examples like these are found in Soames (1987). Example (13b) is from Neale (1990).) Neither of these sentences has a presuppositional reading. But this is just one of many data points that good theories of presupposition projection explain. It would take us too far afield to delve into the details of presupposition projection, and we will not be concerned with these sorts of examples in what follows. So I note these only to put them aside.¹⁴

One notoriously difficult issue that we will not be able to put aside for this discussion is a family of examples such as:

- (14) a. This pen [demonstrating a pen] is owned by the king of France.
 b. What royalty attended the gallery opening?
 The king of France attended the gallery opening.

Examples of these sorts have been observed since Strawson (1964). Unlike the clear infelicity judgments in simple cases like *The king of France is bald*, at least some speakers judge these to be false. These judgments are notoriously weak, and vary with speakers. They may simply point to more facts about presupposition projection, but unlike the cases in (13), there is no consensus the matter. We will return to this issue in Section 4.

Though some have taken examples like (14) and even (13) to be straightforward objections to any presuppositional account of definite descriptions (e.g.

¹³ The literature on the uniqueness of definites is huge. Some snapshots are to be found in Abbot (2004), Heim (1982), Kadmon (1990, 2001), and Roberts (2003), among many places.

¹⁴ For surveys of some important ideas about presupposition projection, see Beaver (2001) and Kadmon (2001). Recent work especially concerned with examples like (13a) includes van der Sandt (1992), recent work especially concerned with examples like (13b) includes Heim (1992).

Neale, 1990), this strikes me as unwarranted, given the clarity of presuppositional judgments in simple cases, and the delicacy of the judgments in (14) (as has been observed since Donnellan 1981). The more promising approach, I believe, seeks to offer a good explanation of what is happening in these cases, without giving up on the basic presuppositional analysis of definite descriptions. At least, I shall rely on ideas about presupposition as we proceed.

3 Negation and Focus

The initial observations we made in Section 1 indicated a lack of evidence of ambiguity in many cases of definite descriptions and negation. If this pattern of lack of evidence stands, as we have seen, it undercuts an argument for the quantificational approach, and lends at least some support to the Fregean approach. The main reason to doubt the pattern really does stand comes from cases of improper definite descriptions such as (8). Before turning in Section 4 to whether these in fact provide evidence of scope ambiguity, we should pause to look more closely at cases of proper definite descriptions.

Our initial observation was that though we see scope ambiguity in the interaction of a quantifier and negation in (2), we do not see it with a definite description in (4). In spite of this initial impression, it is tempting to say that there might really be a scope interaction between descriptions and negation, which informants are just not noticing. Furthermore, it might be suggested, we can bring out the scope differences with the following:

- (15) a. The president of the United States is not CORRUPT.
 b. The PRESIDENT OF THE UNITED STATES is not corrupt.
 c. The president of the United States is NOT corrupt.

Here the capital letters indicate ‘stress’ (more on this in a moment). The placement of stress certainly affects how we understand the sentence, and it might very well seem like it is indicating scope. Very roughly, we have an impression as if the stressed element is somehow taking wide scope. For instance, (15a) seems to say something like *of corruption, the president of the United States does not have it*, whereas (15b) says *of the president of the United States, he does not have the property of being corrupt*, and (15c) says that *the situation does not obtain of the president of the United States being corrupt*. (These paraphrases will be substantially modified in a moment, but they will do for a first pass.) As my paraphrases indicate, each sentence seems to have the stressed element doing something that looks like taking wide scope. Of course, there are no truth-conditional differences between these sentences, but it might appear that there is still a scope difference. We can certainly get informants to note this difference, by showing them the sentences in (15) with stresses in place. Thus, perhaps there is evidence to be had of a scope ambiguity.

If there is a scope ambiguity, what does stressing certain expressions have to do with it? The natural idea is that the placement of stress helps to disambiguate the sentence, and that disambiguating it multiple ways makes clear that the sentence is in fact ambiguous to begin with.

We might bolster this idea by looking at sentences which, unlike (4), clearly display scope ambiguity. When we do, there is a strong impression that stress does indeed *disambiguate* the sentences. For instance, when we add stress marking to sentences like (2), they cease to appear ambiguous, as has been discussed in detail by Büring (1997). Thus, we see only the reading with negation taking wide scope in:

(16) All politicians are NOT corrupt.

We see the same thing in:

(17) He does not hate MOST of the songs.

As discussed at length by Kadmon and Roberts (1986), without stress (17) appears to be ambiguous. But with the stress as marked, only the reading with *most of the songs* taking wide scope over negation seems to be available.

Stress, it may seem from examples like (16) and (17), leads to wide scope. Stressing the definite description or negation in (15b) and (15c) seems, intuitively, to have a similar effect, which looks like fixing scope relations between the description and negation. So, it is tempting to conclude that there really is a scope ambiguity in (4), and that there really are scope interactions between descriptions and negation, which can be brought out by stress. They do not make a truth-conditional difference (when the definite description is proper, at least), but, it might be argued, stress helps us to see that the scope interactions are there.

This argument, I shall argue, is wrong in a number of ways. It is wrong about what stress does, and it is wrong about what our intuition about the readings in (15) really say. It is a tempting line of thought, but it is mistaken.

3.1 *Background on Focus*

To show this we need to begin by exploring, if only in a cursory way, what role stress, as we see in sentences like (15), (16), and (17), is playing in semantics and pragmatics. In these examples, what I have called ‘stress’ marks what linguists call *focus*.¹⁵

¹⁵ Reading the capitals with ‘emphatic stress’ will mark focus. However, most thinking about the phonology of focus these days suggests that it is not the stress that marks the focus, but the intonational prominence that goes with it. In fact, many theorists hold that only a particular intonation contour (a particular pitch accent) marks focus. The right intonation is the one you hear in an appropriate answer to a question, as in (19). See Kadmon (2001) for a survey of some phonology relevant to focus, or the more extensive Ladd (1996). For our purposes here, I shall not worry about the phonological details, and continue to talk about ‘stress’. I am assuming that the phonology realizes an underlying focus feature, F, in logical form, so that the LF of a sentence like (18a) will look like *John likes [Jane]_F*.

There are a number of semantic and pragmatic aspects to focus—more than can be quickly surveyed here. To better understand its relation to scope, let me mention a few. Focus indicates a kind of *contrast*. Take a simple example, like:

- (18) a. John likes JANE.
b. JOHN likes Jane.

Though the truth conditions of (18a) and (18b) are the same, (18a) indicates that John likes Jane, as opposed to John liking Sue or Mary or Yolanda. Likewise, (18b) indicates that John likes Jane, as opposed to Bill or Steve or Ted liking Jane.

Focus also enters into question-answer congruence: the felicity of an answer to a question. We see:

- (19) Who likes Jane?
a. JOHN likes Jane.
b. # John likes JANE.

(‘#’ indicates infelicity.) With the focus in the right place, the answer is felicitous in context, with the focus in the wrong place, it is not.¹⁶

One approach to these and other focus-phenomena is through what is called the *alternative semantics* for focus. Sentences are assigned, in addition to their usual semantic values, an *alternative set*. For a sentence like *John likes JANE* (18a), this is essentially $\{[John\ likes\ x] \mid x\ \text{an individual}\}$. More generally, the alternative set for a sentence is the set of semantic values resulting from replacing the focused element with arbitrary values of the right type. It was a fundamental observation of Rooth (1985) that one can develop a compositional theory of these values, though we will not need the details here.

One of the pragmatic aspects of focus is that for a sentence to be felicitous, its alternative set must be, in the appropriate way, active in the discourse. We see this with question-answer congruence, for instance. Associated with a question is a set of propositions that are (partial) answers to it. On some views (e.g. Hamblin, 1973), this is the semantic value of a question.¹⁷ A very rough first approximation of what is happening in question-answer congruence is that the focused answer is felicitous when the alternative set of the answer is the same as the set of answers to the question (the semantic value of the question). The focused sentence requires the alternative set to be somehow available in the discourse, which asking the right question can bring about.

One of the important effects of focus, which will be especially relevant to considerations of scope, is that it induces a kind of *semantic partition*. Intuitively, we can gloss the effects of focus in (18) as something like:

¹⁶For a more thorough, philosophically friendly discussion of focus, see my (2005a). Other surveys include Kadmon (2001), Kratzer (1991), Rooth (1996), and von Stechow (1991).

¹⁷Hamblin’s idea is modified and developed further by Karttunen (1977). An important alternative is presented in Groenendijk and Stokhof (1984).

- (20) a. John likes JANE.
 = John likes someone, and that person is Jane.
 b. JOHN likes Jane.
 = Someone likes Jane, and that person is John.

There are a number of different approaches to this effect, but one makes use of the apparatus of alternative sets.

Traditionally, this effect is often described in terms of a *focal presupposition*. When a sentence like (20a) is felicitous, it is at least under discussion that John likes someone. Getting an exact characterization of this requirement has been controversial, but a very rough characterization is that it must be presupposed that at least one element of the alternative set of the sentence is true. This amounts to existentially quantifying out the focus position. So, for our simple sentence (20a), we have:

- (21) a. John likes JANE.
 b. Focal presupposition: John likes someone, i.e. $\exists x(\text{John likes } x)$.

As I said, this is a rough approximation of the effect, and there are some well-known ways it might be inadequate. But it will suffice for our purposes here.¹⁸

It should be clear that the focal presupposition is not part of the asserted content of a sentence with focus. The asserted content of (20a) does not include existential quantification. Rather, it is an effect triggered by discourse, and, according to the theory we are pursuing here, the semantics of focus represented by alternative sets. What kind of effect? I called it a presupposition. This is common terminology, and I shall follow it. It is reasonable, as the focal presupposition corresponds to a felicity condition (as we see, for instance, with question-answer congruence). But I am not really concerned here with exactly how the effect is generated. Though I think the presuppositional account is plausible, it would not affect what we are doing here if it turned out to be an implicature, or some other sort of discourse effect.¹⁹

¹⁸ I am inclined to follow a number of authors, including Jackendoff (1972) and Rooth (1999), in holding that focal presupposition is too strong. For instance, Rooth offers the following example:

- (i) Did anyone win the football pool this week?
 Probably not, because it is unlikely that MARY won it, and she is the only person who ever wins.

In this discourse, it is not presupposed that someone won the pool, but the focus is still felicitous. What seems to be required here is only that the set of alternatives for who won the pool has to be in the right way active in the discourse. As Jackendoff put it, these alternatives have to be 'under discussion' (p. 246). Of course, this view is highly controversial. Defenses of existential focal presupposition can be found in Geurts and van der Sandt (2004) and Herburger (2000). However, the issues that are at stake in this debate do not seem to be relevant to our discussion here, and it will simplify matters to talk about focal presupposition. So, even though I am inclined to the opposing view, I shall talk about focal presupposition for purposes of this discussion.

¹⁹ Just what the discourse effect of focus is, and how it is generated, are explored by Roberts (1996) and Rooth (1992). Other approaches to the kind of semantic partition induced by focus

3.2 Focus, Scope, and Negation

Now, with this background in hand, let us return to the issue of focus and scope. It has from time to time been suggested that foci are simply assigned wide scope by a syntactic mechanism (indeed, a syntactic mechanism similar to one that is often proposed to account for quantifier scope).²⁰ But the fact seems to be that focus does not genuinely fix scope. I shall review two arguments for this conclusion from the literature, before turning to the main issue of definite descriptions.

First, Kadmon and Roberts (1986) argue that (17) is in fact ambiguous, and the appearance that the focused quantifier must take wide scope is a pragmatic effect. Focus makes contexts in which we hear the reading with *most* taking narrow scope very remote, but such contexts can be found, they argue. Here is their example:

- (22) Nirit: He likes ‘Smooth Operator’, but MOST of the ‘Top 40’ things he HATES, right?
 Craige: No.
 Nirit: What do you mean ‘no’? He always has some disparaging remark to make about them.
 Craige: OK, so he hates MANY of the songs. All I said was he does not hate MOST of the songs.

This dialog sets up an unusual context, in which we still have focus on *most* and the salient reading is the one where *most of the songs* takes narrow scope with respect to negation.

Second, in examining (16), Büring (1997) notes that the appearance of fixing scope relations is highly specific to the universal quantifiers *all* or *every*. We see no effect of disambiguation in:

- (23) Two thirds of the politicians are NOT corrupt.

Again, focus does not appear sufficient to fix scope. It may, in some contexts, help to disambiguate scope ambiguities; but it does so by making one reading more salient, not by fixing scope relations in a sentence directly.

What is happening in cases like these? A rough approximation is as follows. First of all, there is a discourse effect, related to focus. Contexts in which both a focus is felicitous and the reading is available on which the focus takes narrow scope are often very remote, as we see in Kadmon and Roberts’ example. Hence, we might simply over-generalize and assume focus must fix scope.

Furthermore, there is often an *illusion of scope* created by focal presupposition. The semantic patterning we explained in terms of alternative sets and focal presupposition can look like scope. In a case like (21b), focus triggers the presupposition

include those that introduce structured propositions (e.g. Krifka, 1991; von Stechow, 1991), and those that partition a Davidsonian event decomposition (Herburger, 2000).

²⁰ An idea along these lines is considered by Chomsky (1976). Direct arguments against movement-based approaches to focus are given by Rooth (1985).

that John likes someone (or something close to that). This presupposition will be in the background to the felicitous assertion. Hence, the new information added is that the person doing so is Jane, i.e. Jane is such that she fits the presupposition. The effect, roughly, is to ‘pull’ the focused material out of the rest of the sentence. This can look very much like scope, as any effect of separating off an element from the environment in which it is embedded can. But it is an illusion. We know it must be an illusion, as we have seen how different scope assignments are compatible with the same focus (not to mention the obvious fact that focus affects a wide range of elements we do not think take scope). We can begin to explain how the illusion arises, once we see that the discourse effect of focal presupposition structures the information in an assertion. Much more may be said about how this illusion works, but I think we have enough to be armed against the illusion when we come to look at negation.²¹

I have argued that focus creates an illusion of scope, but does not fix scope. This is not to say that focus has no effect on negation. In some way or another, negation is sensitive to the kind of semantic partitioning that focus induces. We see this in (15) above, where our judgments about what a negated sentence says are influenced by the placement of focus.

How does negation interact with focus? In a typical case, we have:

(24) JOHN is not corrupt.

Alternative set: {[*x is corrupt*] | *x* an individual}

Focal presupposition: Someone is corrupt. (At least one element of the alternative set is true.)

Effect of assertion: Someone is corrupt, but among the corrupt people is not John.

Jackendoff (1972) suggested that this arrangement is written into the semantics of negation, which winds up saying that the focused element does not produce a true proposition. However, much of the more recent literature has argued that the effect here is not generated by the semantics of negation. Rather, the presence of the focal presupposition helps to determine what new information we get out of an assertion of (24), as we see in (24, effect of assertion). This is compatible with negation taking its normal semantics.

In many cases, negation does not affect alternative sets. That is how I represented the alternative set in (24). But there is good reason to think that whether or not it does is a context-dependent matter. Kadmon and Roberts (1986) argue that in their example (22), the alternative set must be negative, corresponding to the question *He hates many of the songs, but how many of the songs does he not hate?* (This is

²¹ For instance, Büring (1997) offers a very detailed analysis of (16). I should note that his analysis takes into account other aspects of the way information is marked phonologically and packaged in discourse than focus. His theory is fascinating and subtle, but we will have to make do with a much rougher-hewn explanation for the moment.

determined in part by the intonation required for *does not*, though I shall not explore that subtlety.) Another example is:

- (25) We are playing the game of confessing to having not read famous works (thanks to David Lodge). John did not read Plato's Republic. Bill did not read Hume's Treatise or Inquiry. Jane did not read most of Aristotle. What about Fred? What did he not read? Fred did not read ALL OF KANT.

On the most salient reading in this context, the alternative set for the last sentence is the set of propositions of the form $\llbracket \textit{Fred did not read } X \rrbracket$. The last sentence is ambiguous between negation wide and narrow readings, though in this context, the negation wide reading appears to be the more salient.²²

3.3 Descriptions and Focus

We have now seen how focus can create an illusion of scope where there is none. We have also seen some of how negation and focus interact, and particularly, how focus can affect the readings of negated sentences. With this in mind, let us look back at our cases of definite descriptions and negation. I suggested above that it might seem like focus reveals scope interactions between descriptions and negation. But we can now safely conclude that it is yet another case of an illusion of scope, created by focus. What is really going on in examples like (15) is:

- (26) a. The president of the United States is not CORRUPT.
 Alternative set: $\{\llbracket \textit{the president of the United States is } F \rrbracket \mid F \text{ a property}\}$
 Focal presupposition: The president of the United States has some property.
 Effect of assertion: The property is not corruption.
- b. The PRESIDENT OF THE UNITED STATES is not corrupt.
 Alternative set: $\{\llbracket x \textit{ is corrupt} \rrbracket \mid x \text{ an individual}\}$
 Focal presupposition: Someone is corrupt.
 Effect of assertion: That person is not the president of the United States.
- c. The president of the United States is NOT corrupt.
 Alternative set: $\{\llbracket \textit{the president of the United States is corrupt} \rrbracket, \llbracket \textit{the president of the United States is not corrupt} \rrbracket\}$
 Focal presupposition: The president of the United States is or is not corrupt (i.e. we are addressing the question *Is the president of the United States corrupt?*).
 Effect of assertion: The answer is that he is not.

²² See Herburger (2000) for an extensive discussion of the kinds of readings that can be generated by interactions between negation and focus. One attempt to work out an account of how focus and negation interact based on the semantics of negation is given by Kratzer (1989).

For (26a) and (26b) I have given the reading in which negation does not affect alternatives (which is the more salient reading ‘out of the blue’). But within the focus semantics of cases like these, we do have a further ambiguity, depending on how the negation and alternative set interact. We have two readings, with alternative sets (27a) and (27b):

- (27) The president of the United States is not CORRUPT.
- a. {[[*the president of the United States is F*] | *F* a property]}
 - b. {[[*the president of the United States is not F*] | *F* a property]}

Depending on where the focus falls, and further depending on how focus and negation interact, we have different readings. These are genuine differences. Even if they are not truth-conditional, they are differences in focal presupposition, which lead to different effects of assertion. But none of these is a difference in the scope of the description.

We have thus seen some real differences among the sentences in (15). They are differences that involve negation, and how negation and the definite description interact. But, I have argued, they are not scope differences. They can look like scope, as we have seen. They can, because focus triggers a combination of a semantic partition and a discourse effect, which creates an illusion of scope. In cases involving negation, which can ‘associate’ with focus, that effect can be important to interpretation. This may lead us to think that what we see in cases of descriptions and negation is a scope relation. But more careful consideration shows us that it is a separate phenomenon. No scopes are fixed by focus, but focus does help us to explain away an illusion of scope in cases of descriptions and negation.

Where does this leave the interpretation of definite descriptions? If there is no need to assign the definite description scope with respect to negation, then the data we have looked at so far gives us no reason to prefer the quantificational approach to the Fregean one. The analysis I gave in (26) and (27) is entirely compatible with the Fregean approach. Neither the semantics of focus, nor the semantics of negation, gives us any reason to treat descriptions as quantifiers. Of course, as I have been cautious to note all along, it gives us no definitive reason to reject the quantificational approach, either. But a careful look at the initial data about negation lends no support to the kind of argument for the quantificational view I sketched in (1), just as it appeared on our cursory look in Section 1.

4 Improper Descriptions and Negation

So far, I have suggested that a certain range of data gives us no evidence of scope ambiguities between definite descriptions and negation. I paused to look at that data closely, and suggested that at best, we can find in it illusions of scope created by focus effects. But so far, we have not confronted the cases which Russellians typically highlight: those of improper definite descriptions like (8). Even if there is no evidence for scope interaction with negation when we look at proper definite descriptions, do the improper ones tell a different story?

The Russellian's argument that they do proceeds along the lines of argument (1). There are, according to Russellians, two readings of a sentence like (8): one on which it is judged to be true, and one on which it is judged to be false (cf. step 1a of argument 1). The best way to explain this, Russellians claim, is to see the description as a quantifier. The reading on which the sentence is true is the one in which the description takes narrow scope with respect to negation. The reading on which it is false is the one in which the description takes wide scope (cf. steps 1b and 1c). Even if we do not see evidence of scope in other examples, the Russellian might hold, this is enough evidence that there is scope interaction with negation. As such, it is evidence in favor of the quantificational approach to definite descriptions (cf. step 1d).

Are there really such judgments as the Russellian claims? As I discussed in Section 2.2, many speakers have Strawsonian judgments about the particular case of (8), holding it to be a presupposition failure. Needless to say, this sort of judgment undercuts the Russellian argument, as it is a uniform judgment of infelicity, which does not support two (truth-conditionally distinct) readings of the sentence. I myself share Strawson's intuitions, and I find when I teach this material that a large number of my students do as well. Nonetheless, Russellian judgments are also well-documented. At best, I think we can say, the state of the judgments is unclear.

There are some ways that the judgments supporting the Russellian argument can be made more firm. In particular, judgments that (8) is *true* can be induced by care about some details of the sentence and its presentation. For instance, even speakers who normally get only the presupposition failure reading can be induced to see a reading on which the sentence is true by a combination of unusual intonation and a further gloss:

(28) The king of France is NOT bald—there is no king of France.

Such configurations have been studied in detail, especially by Horn (1989). They are instances of what Horn calls *metalinguistic negation*.

Very roughly, by Horn's lights, what we have here is a configuration which signals that something is wrong with an *utterance* of a sentence (hence, metalinguistic). He observes that any number of features of an utterance can be rejected by such a configuration. For instance, a manner implicature can be rejected:

(29) That is NOT Slick Willy—it is the president of the United States.

Likewise, aspects of the phonology can be called inappropriate:

(30) He did NOT call the pólíce—he called the pólice.

Metalinguistic negation can be done with other configurations as well, but this is a very typical one.²³

²³ In configurations like this, it might be that the negation is in focus (though when we look at the details of which pitch accents mark focus, this is not at all clear). It is more clear that material in the scope of the metalinguistic negation gets a special phonology: a kind of 'quotation intonation'. In other configurations for metalinguistic negation, pitch accent on negation does not seem to be

Horn lists a number of tests for metalinguistic negation. For instance, metalinguistic negation is marked by being unavailable with incorporated negation, as in *unhappy*. Another test involves what are called positive polarity items: words including *already* which are only acceptable in non-negated environments. Metalinguistic negation allows positive polarity items. Both diagnostics work for (28). We see:

- (31) a. # The king of France is unhappy—there is no king of France.
 b. The king of France is NOT already bald—there is no king of France.

I think it is safe to hold that (28) is a case of metalinguistic negation.²⁴

In cases of metalinguistic negation, judgments of truth cannot support any claim of scope ambiguity. It is clear that there is no scope involved in (29) or (30). When it comes to (28), the metalinguistic negation analysis puts it exactly on par with these two. We simply have a form that rejects the presupposition of a definite description, which does not involve the description taking scope with respect to negation.

So, one possible explanation of what is happening with (8) is that we have the coincidence of two distinct phenomena. On the one hand, we have the focus-related behavior of negation discussed in Section 3.2, and on the other hand, we have the availability of a metalinguistic negation reading. The latter explains the truth-value judgment to which Russellians appeal, while the former explains why we might have been inclined to attribute the judgment to scope interaction. We have already seen that the appearance of scope induced by focus is an illusion, and the truth-value judgment induced by metalinguistic negation is likewise not based on scope. We can thus explain why Russellians might have thought there was a scope ambiguity, and also why that is a mistake.

It may be that this is all we need. Insofar as the real force of the Russellian argument rests on the judgment of truth for (8), and insofar as this judgment is due to metalinguistic negation, it is. But when it comes to the very murky judgments surrounding these cases, there is one other way to elicit Russellian judgments that we need to consider.

This brings us to the difficult cases discussed in Section 2.2, which seem to undermine the presuppositional readings of definites. We can get a little bit firmer

necessary, but the quotation marking does seem to be required. This has not, to my knowledge, been discussed at much length in the literature, but see Potts (2005). For a somewhat different take on the phenomena at issue here, see Geurts (1998).

²⁴ The characterization of positive polarity items as only being able to occur in non-negated environments is well-known to be very rough, and the empirical situation is in fact somewhat more subtle. For a survey of some of the issues involved, see Ladusaw (1996). Some recent discussions paying attention to positive polarity items include Progovac (1994), Szabolcsi (2004), and van der Wouden (1997). As an anonymous referee pointed out, this can make the application of the positive polarity item test difficult, as it requires sorting out whether we are seeing a metalinguistic negation, or one of the other constructions which make the generalization about negative environments subtle. Nonetheless, it does appear that the combined force of the tests, run on the particular construction in question, strongly indicates we have metalinguistic negation.

judgments of truth value than we get for (8) ‘out of the blue’ by putting the definite description in the kinds of environments I mentioned in (14):

- (32) a. This pen [demonstrating a pen] is not owned by the KING OF FRANCE.
 b. What royalty attend the gallery opening?
 The KING OF FRANCE did not attend the gallery opening.

Though as I have mentioned, judgments about these cases are often very weak, at least some speakers find these acceptable, and judge them to be *true*. Moreover, some speakers who get presupposition failure in (8) find these better, and are at least willing to give truth-value judgments.²⁵

As I mentioned in Section 2.2, there is little consensus about what is happening in cases like these. The phenomenon is often-noted, but not well-understood. It is not even agreed just what the phenomenon is, e.g. whether these are cases of non-presuppositional readings, or whether there is still a presupposition failure, but one which somehow does not interfere with truth-value judgments. For our purposes here, it will be enough to note that the phenomenon is there. Rather than try to explain it, I shall give it a name. To give it one which I hope is somewhat neutral, let us call these cases of *presupposition obviation*.²⁶

I marked foci in (32) where they seem natural. Where the focus should fall is clear for the question-answer pair in (32b), while (32a) is very natural in response to a question like *What about the writing instruments around here? Are any of them related to royalty?* It is a common idea that focus (or related notions) is involved in cases of presupposition obviation. I do not want to take a stand on that here. But regardless of what the right analysis is, focus can help to bring out the presupposition-obviating readings.²⁷

Though I shall not try to sort out what is really behind the phenomenon of presupposition obviation, I shall show that it is not anything to do with scope with negation, and hence, examples like (32) do not give us any evidence that definite descriptions enter into scope relations with negation. The reason for this is really quite simple: presupposition obviation arises without negation. We saw this in example (14) of Section 2.2, repeated here, with foci marked:

²⁵ For some discussion of the range of judgments seen for some related cases, see Reinhart (1995).

²⁶ I did try to explain some aspects of presupposition obviation, based on the way context is set by discourse, in my (2002). Early work on the problem includes Strawson’s own paper (1964) and Fodor (1979). The idea that there is still presupposition failure in cases of presupposition obviation, but that speakers are able to reach truth-value judgments regardless, is explored by von Stechow (2004) and Lasnik (1993). Both of the latter, and my own contribution, offer ways we can make sense of presupposition obviation even for the kind of semantically coded presupposition that is built into (11).

²⁷ In (32a), the subject *this pen* will typically have a distinct intonation contour (a distinct pitch accent), which is often taken to mark a contrastive *topic*. The idea that focus, or more often the related notion of topic, is fundamental to presupposition obviation has been proposed by a number of authors, including Gundel (1974), Horn (1986), Kadmon (2001), Reinhart (1981, 1995), and the original Strawson (1964).

- (33) a. This pen [demonstrating a pen] is owned by the KING OF FRANCE.
 b. What royalty attended the gallery opening?
 The KING OF FRANCE attended the gallery opening.

We get the same sorts of judgments for these cases as we do for (32). Some speakers find them acceptable, and in these cases, *false*. As with (32), these judgments are often weak, but we still see at least some speakers finding these better than canonical cases of presupposition failure for definites like (8). We thus have more cases of presupposition obviation. Yet the examples in (33) do not involve negation. There is no negation for the description to take scope with, even if it were to be analyzed as scope-taking. Hence, we need an analysis of presupposition obviation that does not rely on scope with negation. Once we find one, we can apply it to cases like (32) directly. However this will be done, it appears entirely compatible with the Fregean approach. Once again, we find nothing in the data which supports the quantificational approach to the exclusion of the Fregean. We find no evidence of genuine scope interaction between definite descriptions and negation.

We have now seen two different sorts of ways that we might induce Russellian judgments for improper definite descriptions. They might be understood as cases of metalinguistic negation, or as cases of presupposition obviation. I have argued that neither phenomenon is based on scope interactions with negation. Hence, I suggest, the right explanation of the judgments to which Russell appealed in his discussion of (8) is not scope with negation. I have already noted that the illusion that these judgments are derived from scope may be made all the stronger by the interaction with focus. Focus creates an illusion of scope, and focus can indicate presupposition obviation, and can go together with metalinguistic negation as well. I thus come to the same conclusion about (8) as I did about (15). There is no evidence of scope interaction between definite descriptions and negation, though a range of factors, semantic and pragmatic, can lead us to think there is. The factors that create the illusion may be stronger in cases of presupposition obviation, but it is an illusion nonetheless.

5 Negative Quantifiers

Another category of cases that have been argued to show more clear scope interactions with descriptions are those of negative environments other than the one generated by *not*. Neale (1990) notes the effects of negative (i.e. *monotone decreasing*) quantifiers:

- (34) a. Nobody has kissed the king of France.
 b. Few Frenchmen have seen the king of France.

Neale is careful to say that it is ‘at least arguable’ (p. 120) that these have readings which come out true. According to the Russellian, they come out true in virtue of the description taking narrow scope with respect to a negative element; in this case, a quantifier rather than negation itself.

Now, the first thing to say is that, as with some of the cases we just discussed in Section 4, these judgments are rather weak. Like the cases we just reviewed, when and whether we get presuppositional judgments in cases like these seems to depend on many factors, and to vary with speakers. (I find all the judgments marginal, and (34b) is markedly worse than (34a).) Again, focus, together with the right discourse settings, can help bring out the non-presuppositional readings. For instance, compare:

- (35) a. We are a group of royalty-chasers, who play a game of trying to kiss various monarchs. Mary has kissed the prince of Monaco, John has kissed the queen of England, Jane has kissed the king of Norway, but *#NOBODY has kissed the king of France.*
- b. We were at a wild party. All sorts of crazy things were going on, and lots of my friends kissed strange people, but at least, *?Nobody has kissed the KING OF FRANCE.*

In contexts, (35a) appears to be a presupposition failure, while (35b) appears at least marginally acceptable. At least (and perhaps more neutrally), there is marked contrast between the two, with (35b) significantly better than (35a). One of the reasons we seem to find cases like (34) acceptable (if and when we do) is that we naturally read them with focus in the right places, as if they were in the right contexts.

What we see here, I suggest, is simply the same phenomenon of presupposition obviation we saw in Section 4. I tried to be cautious there about just what this phenomenon is, but the role of focus and context gives us some evidence that it is at work, as does the rather delicate nature of the judgments. If it is right that what we have is presupposition obviation for the definite description, then we no longer have any reason to appeal to scope interactions between the definite description and the negative quantifier. The sentences in (34) are not ambiguous with respect to scope. There is one reading for each, when it comes to scope, but in some settings, there is also a phenomenon of presupposition obviation, which can affect the truth-value judgments associated with the sentences. Again, we have an illusion that the definite description might take scope, induced by focus, and a truth-value judgment supported by presupposition obviation.

We can get some further evidence that scope is not at work here by looking at cases where there are non-trivial scope interactions with monotone decreasing quantifiers. The scopal properties of these quantifiers is actually a somewhat difficult area, but we can find examples where a monotone decreasing quantifier takes *narrow* scope with respect to another quantifier. For instance:

- (36) a. SOME MONARCH trusts nobody.
 b. MOST MONARCHS trust nobody.
 c. EVERY MONARCH trusts nobody.

In all these cases, the reading where the subject takes scope over the object is the salient one. The foci correspond to answering a question like *Who trusts nobody?* I

have marked the foci, as we will want to make a comparison with a focused definite description. But in fact, the readings with *nobody* taking wide scope do not seem to be available, regardless of focus. It may very well be that there are no such readings.²⁸

When we consider improper definite descriptions in this configuration, we can encounter whatever phenomenon is at work in (35b):

- (37) What royalty trust anybody?
 ?THE KING OF FRANCE trusts nobody.

Again, the judgment here is delicate, but this appears to be as good as (32b), and I think, about as good as (35b). For those who find it acceptable, scope between the description and *nobody* will not explain it. Those who find this acceptable get the judgment that it is *true*. But treating the definite description as scope-taking does not predict this. The pattern in (36) gives us good reason to suppose that the quantifier *nobody* would take narrow scope here, if there were any scope to take. But in that case, the definite description would have to take wide scope. The result would be that the sentence, on the quantificational reading, would imply that the king of France exists, and so would be false. Thus, the scope-taking treatment of this case gets the wrong answer. Insofar as we have a truth-value judgment to work with here, the scope-taking account gets the wrong truth value. Insofar as we have a presupposition failure, the scope-taking treatment fails to predict it.

To stress, I find the judgments here somewhat dubious, and certainly too delicate to place much weight on. My point is that insofar as we find cases of negative quantifiers like (35b) or (37) acceptable, we should not account for them via scope interactions. If they are acceptable, then scope interaction gets the wrong answer for (37). More generally, we already have good reason to see these as cases of presupposition obviation, which indeed goes naturally with the delicate judgments at work. I already argued that presupposition obviation is not a scope phenomenon. The point that scope-based treatments get the wrong answers in some cases reinforces this conclusion. Negative quantifiers give no independent evidence of scope with definite descriptions, and again, no reason to prefer the quantificational over the Fregean approach.

6 Conclusion

A close look at negation has shown that the initial observations of Section 1 bear up well under scrutiny. We saw there, with examples like (4), that we do not

²⁸ The syntax literature offers some explanations of these sorts of effects, which might predict that we cannot get inverse scope readings in (36). It is commonly observed, for instance, that objects tend not to take wide scope in negative environments. See for instance, Aoun and Li (1993), for one approach to this issue. The division of quantifiers into types in Beghelli and Stowell (1997) predicts that monotone decreasing quantifiers in object position will generally not take wide scope.

generally have data to indicate scope ambiguities involving definite descriptions and negation. I have argued in the rest of this paper that those observations are correct. Furthermore, I have tried to explain away the appearance of scope interactions between definite descriptions and negation in some cases as an illusion, due to focus. I have also shown that judgments surrounding improper definite descriptions are best explained by something other than scope interaction. In some cases, it is metalinguistic negation that leads to these judgments, in others (perhaps), it is presupposition obviation. Neither of these is a scope phenomenon, though each interacts with focus in ways that reinforce the illusion of scope. I thus conclude, we do not have evidence of scope interaction when it comes to definite descriptions and negation.

I have suggested that this conclusion undercuts a familiar Russellian argument for the quantificational approach, of the sort I sketched in (1). To some extent, it offers comfort to the Fregean approach as well. The Fregean approach predicts that there are no truth-conditional scope ambiguities produced by definite descriptions, and at least in the case of negation, we find none.²⁹ This is not enough to establish the correctness of the Fregean approach, or to decide if definite descriptions really do take scope, but it leaves the Fregean approach viable. Perhaps to an even greater extent, my conclusion underscores the note of caution I sounded at the beginning of this paper. Scope itself is a highly theoretical matter, whose relations to our basic data can be complex. We have seen here another way in which the data can be complex, as we have to distinguish illusions of scope from data really supporting scope hypotheses.

Acknowledgment Thanks to Kent Bach, Chris Barker, Paul Elbourne, John MacFarlane, Robert May, Paul Pietroski, Daniel Rothschild, Rob Stainton, the Bay Area Philosophy of Language Discussion Group, and an anonymous referee for very helpful comments on earlier versions of this paper. A preliminary version of some of this material was presented in my seminar at UC Davis in 2005. Thanks to the participants there, especially Brian Bowman, Josh Parsons, and Paul Teller.

References

- Abbot, B. (2004). Definiteness and indefiniteness. In L. R. Horn & G. Ward (Eds.), *Handbook of pragmatics* (pp. 122–149). Oxford: Blackwell.
- Acquaviva, P. (1993). *The Logical form of negation*. Ph.D. dissertation, Scuola Normale Superiore, Pisa. Published by Garland, New York, 1997.
- Aoun, J., & Li, Y.-H. A. (1993). *Syntax of scope*. Cambridge: MIT Press.
- Barwise, J., & Cooper, R. (1981). Generalized quantifiers and natural language. *Linguistics and Philosophy*, 4, 159–219.
- Beaver, D. I. (2001). *Presupposition and assertion in dynamic semantics*. Stanford: CSLI Publications.

²⁹ I reach similar conclusions for cases of interactions between definite descriptions and quantifiers in my (2007). Other kinds of cases, notably interactions with modals, have been discussed by Elbourne (2005) and Heim (1991), but still raise a number of important issues.

- Beghelli, F., & Stowell, T. (1997). Distributivity and negation: The syntax of *each* and *every*. In A. Szabolcsi (Ed.), *Ways of scope taking* (pp. 71–107). Dordrecht: Kluwer.
- Büring, D. (1997). The great scope inversion conspiracy. *Linguistics and Philosophy*, 20, 174–195.
- Chierchia, G., & McConnell-Ginet, S. (1990). *Meaning and grammar*. Cambridge: MIT Press.
- Chomsky, N. (1976). Conditions on rules of grammar. *Linguistic Analysis*, 2, 303–351. Reprinted in Chomsky (1977).
- Chomsky, N. (1977). *Essays on form and interpretation*. Amsterdam: North-Holland.
- Diesing, M. (1992). *Indefinites*. Cambridge: MIT Press.
- Donnellan, K. S. (1981). Intuitions and presuppositions. In P. Cole (Ed.), *Radical pragmatics* (pp. 129–142). New York: Academic Press.
- van Eijck, J. (1993). The dynamics of descriptions. *Journal of Semantics*, 10, 239–267.
- van Eijck, J., & Kamp, H. (1997). Representing discourse in context. In J. van Benthem & A. ter Meulen (Eds.), *Handbook of logic and language* (pp. 179–237). Cambridge: MIT Press.
- Elbourne, P. D. (2005). *Situations and individuals*. Cambridge: MIT Press.
- Fara, D. G. (2001). Descriptions as predicates. *Philosophical Studies*, 102, 1–42.
- von Fintel, K. (2004). Would you believe it? The King of France is back! In M. Reimer & A. Bezuidenhout (Eds.), *Descriptions and beyond* (pp. 315–341). Oxford: Oxford University Press.
- Fodor, J. D. (1979). In defense of the truth value gap. In C.-K. Oh & D. A. Dineen (Eds.), *Presupposition*, vol. 11 of *syntax and semantics* (pp. 199–224). New York: Academic Press.
- Frege, G. (1879). *Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*. Halle: Nebert. References are to the translation as ‘Begriffsschrift, a Formal Language, Modeled upon that of Arithmetic, for Pure Thought’ by S. Bauer-Mengelberg in Van Heijenoort (1967).
- Frege, G. (1891). *Function und Begriff*. Jena: Pohle. References are to the translation as ‘Function and Concept’ by P. Geach in Frege (1984).
- Frege, G. (1893). *Grundgesetze der Arithmetik*, vol. I. Jena: Pohle. References are to the partial translation as ‘The Basic Laws of Arithmetic’ by M. Furth, University of California Press, Berkeley, 1964.
- Frege, G. (1984). *Collected papers on mathematics, logic, and philosophy*. Oxford: Basil Blackwell. Edited by B. McGuinness.
- Geurts, B. (1998). The mechanisms of denial. *Language*, 74, 274–307.
- Geurts, B., & van der Sandt, R. (2004). Interpreting focus. *Theoretical Linguistics*, 30, 1–44.
- Glanzberg, M. (2002). Context and discourse. *Mind and Language*, 17, 333–375.
- Glanzberg, M. (2005a). Focus: A case study on the semantics–pragmatics boundary. In Z. G. Szabó (Ed.), *Semantics versus pragmatics* (pp. 72–110). Oxford: Oxford University Press.
- Glanzberg, M. (2005b). Presuppositions, truth values, and expressing propositions. In G. Preyer & G. Peter (Eds.), *Contextualism in philosophy* (pp. 349–396). Oxford: Oxford University Press.
- Glanzberg, M. (2006). Quantifiers. In E. Lepore & B. Smith (Eds.), *Oxford handbook of philosophy of language* (pp. 794–821). Oxford: Oxford University Press.
- Glanzberg, M. (2007). Definite descriptions and quantifier scope: Some Mates cases reconsidered. *European Journal of Philosophy*, 3, 133–158.
- Groenendijk, J., & Stokhof, M. (1984). *Studies in the semantics of questions and the pragmatics of answers*. Ph.D. dissertation, University of Amsterdam.
- Groenendijk, J., & Stokhof, M. (1991). Dynamic predicate logic. *Linguistics and Philosophy*, 14, 39–100.
- Gundel, J. K. (1974). *The role of topic and comment in linguistic theory*. Ph.D. dissertation, University of Texas at Austin. Published by Garland, New York, 1988.
- Hamblin, C. L. (1973). Questions in Montague English. *Foundations of Language*, 10, 41–53.
- van Heijenoort, J. (Ed.) (1967). *From Frege to Gödel: A source book in mathematical logic, 1879–1931*. Cambridge: Harvard University Press.
- Heim, I. (1982). *The semantics of definite and indefinite noun phrases*. Ph.D. dissertation, University of Massachusetts at Amherst. Published by Garland, New York, 1989.

- Heim, I. (1991). Artikel und definitheit. In A. von Stechow & D. Wunderlich (Eds.), *Semantics: An international handbook of contemporary research* (pp. 487–535). Berlin: de Gruyter.
- Heim, I. (1992). Presupposition projection and the semantics of attitude verbs. *Journal of Semantics*, 9, 183–221.
- Heim, I., & Kratzer, A. (1998). *Semantics in generative grammar*. Oxford: Blackwell.
- Herburger, E. (2000). *What counts*. Cambridge: MIT Press.
- Higginbotham, J., & May, R. (1981). Questions, quantifiers and crossing. *Linguistics Review*, 1, 41–79.
- Horn, L. R. (1986). Presupposition, theme and variations. In A. M. Farley, P. T. Farley, & K.-E. McCollough (Eds.), *CLS 22: Papers from the parasession on pragmatics and grammatical theory* (pp. 168–192). University of Chicago: Chicago Linguistics Society.
- Horn, L. R. (1989). *A Natural history of negation*. Chicago: University of Chicago Press.
- Jackendoff, R. S. (1972). *Semantic interpretation in generative grammar*. Cambridge: MIT Press.
- Jacobson, P. (2002). The (dis)organization of grammar: 25 years. *Linguistics and Philosophy*, 25, 601–626.
- de Jong, F., & Verkuyl, H. (1985). Generalized quantifiers: The properness of their strength. In J. van Benthem & A. ter Meulen (Eds.), *Generalized quantifiers in natural language* (pp. 21–43). Dordrecht: Foris.
- Kadmon, N. (1990). Uniqueness. *Linguistics and Philosophy*, 13, 273–324.
- Kadmon, N. (2001). *Formal pragmatics*. Oxford: Blackwell.
- Kadmon, N., & Roberts, C. (1986). Prosody and scope: The role of discourse structure. In A. M. Farley, P. T. Farley, & K.-E. McCollough (Eds.), *CLS 22: Papers from the parasession on pragmatics and grammatical theory* (pp. 16–28). University of Chicago: Chicago Linguistics Society.
- Kamp, H. (1984). A theory of truth and semantic representation. In J. Groenendijk, T. Janssen, & M. Stokhof (Eds.), *Truth, interpretation, and information* (pp. 1–41). Dordrecht: Foris.
- Karttunen, L. (1977). Syntax and semantics of questions. *Linguistics and Philosophy*, 1, 3–34.
- Keenan, E. L., & Stavi, J. (1986). A semantic characterization of natural language determiners. *Linguistics and Philosophy*, 9, 253–326. Versions of this paper were circulated in the early 1980s.
- Kempson, R. M., & Cormack, A. (1981). Ambiguity and quantification. *Linguistics and Philosophy*, 4, 259–309.
- Kratzer, A. (1989). An investigation of the lumps of thought. *Linguistics and Philosophy*, 12, 607–653.
- Kratzer, A. (1991). The representation of focus. In A. von Stechow & D. Wunderlich (Eds.), *Semantics: An international handbook of contemporary research* (pp. 825–834). Berlin: de Gruyter.
- Krifka, M. (1991). A compositional semantics for multiple focus constructions. *Proceedings of Semantics and Linguistic Theory*, 1, 127–158.
- Lackey, D. (Ed.) (1973). *Essays in analysis*. New York: George Braziller.
- Ladd, D. R. (1996). *Intonational phonology*. Cambridge: Cambridge University Press.
- Ladusaw, W. A. (1996). Negation and polarity items. In S. Lappin (Ed.), *Handbook of contemporary semantic theory* (pp. 321–341). Oxford: Blackwell.
- Lappin, S., & Reinhart, T. (1988). Presuppositional effects of strong determiners: A processing account. *Linguistics*, 26, 1021–1037.
- Laserson, P. (1993). Existence presuppositions and background knowledge. *Journal of Semantics*, 10, 113–122.
- Levinson, S. C. (1983). *Pragmatics*. Cambridge: Cambridge University Press.
- Marsh, R. C. (Ed.) (1956). *Logic and knowledge*. London: George Allen and Unwin.
- May, R. (1985). *Logical form: Its structure and derivation*. Cambridge: MIT Press.
- Montague, R. (1973). The proper treatment of quantification in ordinary english. In J. Hintikka, J. Moravcsik, & P. Suppes (Eds.), *Approaches to natural language* (pp. 221–242). Dordrecht: Reidel. Reprinted in Montague (1974).

- Montague, R. (1974). *Formal philosophy*. New Haven: Yale University Press. Edited by R. Thomason.
- Neale, S. (1990). *Descriptions*. Cambridge: MIT Press.
- Pietroski, P., & Hornstein, N. (2002). Does every sentence like this display a scope ambiguity? In W. Hinzen & H. Rott (Eds.), *Belief and meaning*. Frankfurt: Hansel-Hohenhausen.
- Potts, C. (2005). Lexicalized intonational meaning. *University of Massachusetts Occasional Papers*, 30, 129–146.
- Progovac, L. (1994). *Negative and positive polarity*. Cambridge: Cambridge University Press.
- Reinhart, T. (1979). Syntactic domains for semantic rules. In F. Guenther & S. J. Schmidt (Eds.), *Formal semantics and pragmatics for natural languages* (pp. 107–130). Dordrecht: Reidel.
- Reinhart, T. (1981). Pragmatics and linguistics: An analysis of sentence topics. *Philosophica*, 27, 53–94.
- Reinhart, T. (1983). *Anaphora and semantic interpretation*. Chicago: University of Chicago Press.
- Reinhart, T. (1995). Interface strategies. *OTS Working Papers in Linguistics*.
- Reinhart, T. (1997). Quantifier scope: How labor is divided between QR and choice functions. *Linguistics and Philosophy*, 20, 335–397.
- Roberts, C. (1996). Information structure in discourse: Towards an integrated formal theory of pragmatics. *Ohio State University Working Papers in Linguistics*, 49, 91–136.
- Roberts, C. (2003). Uniqueness in definite noun phrases. *Linguistics and Philosophy*, 26, 287–350.
- Rooth, M. (1985). *Association with focus*. Ph.D. dissertation, University of Massachusetts at Amherst.
- Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics*, 1, 75–116.
- Rooth, M. (1996). Focus. In S. Lappin (Ed.), *Handbook of contemporary semantic theory* (pp. 271–297). Oxford: Blackwell.
- Rooth, M. (1999). Association with focus or association with presupposition? In P. Bosch & R. van der Sandt (Eds.), *Focus: linguistic, cognitive, and computational perspectives* (pp. 232–244). Cambridge: Cambridge University Press.
- Russell, B. (1905). On denoting. *Mind*, 14, 479–493. Reprinted in Marsh (1956).
- Russell, B. (1919). *Introduction to mathematical philosophy*. London: George Allen and Unwin.
- Russell, B. (1957). Mr. Strawson on referring. *Mind*, 66, 385–389. Reprinted in Lackey (1973).
- Sainsbury, R. M. (1979). *Russell*. London: Routledge and Kegan Paul.
- van der Sandt, R. A. (1992). Presupposition projection as anaphora resolution. *Journal of Semantics*, 9, 333–377.
- Sellars, W. (1954). Presupposing. *Philosophical Review*, 63, 197–215.
- Sharvey, R. (1969). Things. *The Monist*, 53, 488–504.
- Soames, S. (1987). Review of Hornstein's *Logic as Grammar*. *Journal of Philosophy*, 84, 447–455.
- von Stechow, A. (1991). Current issues in the theory of focus. In A. von Stechow & D. Wunderlich (Eds.), *Semantics: An international handbook of contemporary research* (pp. 804–825). Berlin: de Gruyter.
- Strawson, P. F. (1950). On referring. *Mind*, 59, 320–344. Reprinted in Strawson (1971).
- Strawson, P. F. (1952). *Introduction to logical theory*. London: Methuen.
- Strawson, P. F. (1964). Identifying reference and truth-values. *Theoria*, 30, 96–118. Reprinted in Strawson (1971).
- Strawson, P. F. (1971). *Logico-linguistic papers*. London: Methuen.
- Szabolcsi, A. (2004). Positive polarity—negative polarity. *Natural Language and Linguistic Theory*, 22, 409–452.
- Whitehead, A. N., & Russell, B. (1927). *Principia mathematica*, vol. I. Cambridge: Cambridge University Press, 2nd edn.
- Wilson, G. (1978). On definite and indefinite descriptions. *Philosophical Review*, 87, 48–76.
- Winter, Y. (1997). Choice functions and the scopal semantics of indefinites. *Linguistics and Philosophy*, 20, 399–467.
- van der Wouden, T. (1997). *Negative contexts*. London: Routledge.

Evidentials: Some Preliminary Distinctions

James Higginbotham

Abstract I raise several questions about the semantic interpretation of evidentials, supposing throughout that they are only through grammaticalization distinguished from main Verbs, and thus amenable to abstract study, even through languages that do not support evidential morphology. These questions point to distinctions that are often not made, and sometimes not even considered, in the important and growing literature on the topic. A major question is: does a person who asserts an evidential sentence say one thing, or two? Some links to well-known philosophical topics, such as first-person authority, are also explored.

Keywords Evidentials · binary constructions · functional heads · first-person authority

The now large, and growing, literature on evidentials relies in considerable measure upon arduous and accurate fieldwork, undertaken from different points of view, and aiming at the classification, and the syntactic and semantic effects, of evidential constructions in the varieties of languages in which they occur. At the same time, one is free to conjecture evidentials of one sort or another in more familiar languages. In consequence, there is some controversy over how to formulate, out of the raw material, a proper cross-linguistic concept. In this article, I will note some distinctions that seem to me essential to concept formation in this area. These distinctions are indeed in some measure noted here and there in the literature on evidentials; but they are relatively neglected in much of it. In many cases, crucial data (often very hard to elicit, even if envisaged) are lacking. The semantic points can, however, be illustrated in any language, English included.

Besides the classificatory questions that I will raise here, there are some larger issues involving the semantic relations between host elements and subordinate elements, which invite inquiry into relations between propositional content and

J. Higginbotham

School of Philosophy, University of Southern California, 3709 Trousdale Parkway, Los Angeles, CA 90089-0451, USA
e-mail: higgy@usc.edu

speech act, or in Fregean terms between *sense* and *force*. These issues appear in connection with evidentials, but also in a wider domain, taking in even the question of the proper syntactic description of the combination of host and subordinate elements.

In this article, following a brief introduction to the topic, I consider a family of semantic questions concerning evidentials; questions that, so far as I am aware, have not received definitive answers in the literature. My discussion will include a preliminary exploration of English data, and these data in turn will be put to use in an attempt to clarify some of the issues with evidentials.

A note that may help orient the reader to the perspective taken here: I understand the notion *evidential* in such a way that syntactic appearance of an evidential, e.g. as an affix, or as a main Verb, or in another guise, is not a defining characteristic of the notion. (In this respect, if I understand him correctly, I agree with Rooryck (2001a).) Evidential affixes are presumably functional heads, taking clausal complements (with different effect in view of the distinction between singular and binary constructions, discussed below). From the point of view taken here, they might as well be main Verbs.

1 Evidential Constructions

In a number of languages, not in general grouped together by historical affinity, evidential morphemes appear optionally or obligatorily on Verbal (also in some cases Nominal) heads. They generally function so as to indicate something about the epistemic background of an assertion, but may also have other implications, for instance for scope of quantifiers (Lecarme (2003)). I will concentrate here on the Verbal case, where a typical sentence from a language with evidentials might be as in (1), with ‘ ϵ ’ representing the evidential:

- (1) A horse ran+ ϵ through the village.

(1) might be asserted on a variety of grounds: I saw (or thought I saw) a horse; I saw some tracks that I take to be hoofprints (spaced far apart, as if made by a horse galloping); horses run through the village all the time in this season; somebody told me a horse passed by; and so on. The choice of ϵ selects amongst these. The selection may be from 2–4 categories (Willett (1988)), not by any means uniform across languages (see Faller (2002), Speas and Tenny (2004), and references cited there). Viewing (1) as asserted, the speaker is indicating something about the background leading to the belief asserted, its ‘source of information,’ the ‘strength of evidence,’ or something of the sort.

From a syntactic point of view, evidentials are high-up verbal affixes that signify a speaker’s relation to her grounds for the assertion of *S* (or, in the case of at least some interrogatives, ask the respondent’s relation to the grounds upon which he asserts an answer to the question whether *S*). I use ‘grounds for assertion’ as neutral between two popular interpretations, namely ‘source of information’ and ‘nature of evidence.’ Willett (1988) proposed a hierarchy of up to four evidential categories:

(i) personal experience; (ii) direct sensory evidence; (iii) indirect evidence; (iv) hearsay (his terms). From this description, the epistemic background could take any of several forms.

2 Evidence versus Source

There appears to be a divergence between evidentials that have to do with strength of evidence, as opposed to those that make reference to source of information. The two notions will inevitably be hard to disentangle in some cases, but can be separated directly in cases of what is commonly referred to in philosophical discussions as *first-person authority*.¹

The domain of first-person authority comprises assertions, or occurrences of belief, advanced without anything that could count as evidence (though there may be an experiential basis for them), where (in the normal case) the truth of the (sincere) assertion or belief goes without saying. Examples: 'My finger hurts;' 'I'm thirsty;' 'I doubt John will come;' "'The child seems sleeping" isn't a sentence for me;' and the like. Naturally, viewed as utterances, they involve the first person, and the present tense. Thus 'My finger hurt yesterday' does not carry first-person authority (perhaps I am misremembering); the assertion 'His finger hurts' is always, if reasonable, based on evidence (his behavior, for instance). But I don't look at my own behavior to determine that my finger hurts, or that I am thirsty, or that I doubt John will come. And when I pronounce that so-and-so is or is not a sentence of my native language, I don't produce evidence, even though I may be wrong.

Of the writers I have seen, only Garrett (2001) (discussing Tibetan) observes any cases involving first-person authority, mentioning for instance 'I am hungry' (where in the normal case it is nonsense to ask, 'What's your evidence?'). This assertion takes what he calls the Tibetan 'immediate knowledge' evidential. That label nicely straddles the distinction between source of information (that feeling of hunger in the speaker's stomach) and any question of evidence. However, I am not aware that the full domain of first-person authority has been exploited for data. Besides the examples above, that domain would include statements 'I want ...', or 'I firmly believe ...', and the like, and would depend upon the first person for the relevant evidential. In this and in other respects to be noted below, the morphosemantics of the evidential appears to be underdetermined by what is now known.

The above purely reflective semantic diagnostics of first-person authority can be associated with the strangeness of the use of examples normally displaying it when they are put in conjunction with English epistemic *must*. Now, much of the literature on epistemic *must* takes it to be a modal attaching to the whole of a declarative, as in (2):

- (2) Max must be in his office.

indicating at the same time that evidence is in some sense 'indirect' (see for instance Izvorski (1997) and references cited there). I find this assumption doubtful, because it does not explain why (2) is not a stronger assertion than the simple (3):

(3) Max is in his office.

(a property that epistemic *must* shares with English qualifiers such as *doubtless*, or *bound to be*).² In any case, epistemic *must* does indeed indicate that the situation is evidentially odd with respect to contexts that normally fall under first-person authority. Thus the somewhat strained interpretations of assertions of (4) and (5), each of which suggests that the speaker has evidence for something that normally is known without it:

(4) I must be hungry.

(5) I must believe that I am a Communist.

(with the modal understood epistemically). It would be interesting to know whether the ‘indirect’ evidentials give rise to similar interpretations in languages with evidential morphemes, or whether they are simply rejected out of hand.

To sum up: evidentials that invariably signify *evidence*, rather than *source*, should be incompatible with normal cases of first-person authority (though of course syncretism remains a possibility). Conversely, if there are evidentials that can be shown only to characterize sources, but never evidence, these should be incompatible with the counterparts of assertions such as (4) and (5), where normal first-person authority is specifically denied.

3 The Object of Evidentiality

Do the complements of evidentials refer to (or quantify over) propositions, or events? It has been noted that some evidentials have historically developed out of, or as extensions of, perception Verbs (Botne (1997)). Garrett (2001) counts one evidential form as that of what he calls *direct perception*. Now, the perception verbs (in English, particularly ‘see’ and ‘hear’, but also ‘feel’ in the strict sense of apprehension by touch, and including some uses of ‘remember’ and ‘imagine’) may be understood epistemically or non-epistemically. The epistemic case is illustrated by (6), the non-epistemic by (7):

(6) John saw/heard [that Mary was crying]

(7) John saw/heard [Mary cry/crying]

In (6), ‘see’ signifies roughly *come to know by using one’s eyes* (or metaphorical extensions thereof), and ‘hear’ signifies *come to learn that it was said or rumoured*. The construction is referentially opaque, and so does not admit substitutivity of identity for singular terms, or existential generalization. In (7), however (following Higginbotham (1983), the extension to *remember* and *imagine* being given in Higginbotham (2003)), the perception verbs appear *in propria persona*, alleging the seeing or hearing of things; i.e., events or states. The tenseless and complementizerless objects serve as existential quantifiers over events *e*, classified in our examples as events of Mary’s crying. Hence, substitutivity of identity holds without exception, and existential generalization is allowed with respect to

all argument positions. But more than this, the truth conditions of the constructions shown in (7) have nothing to do with the agent's epistemic situation. Thus it might be true that John heard Mary crying even though he thought it was the wind in the trees. Again, if the agent asserts (8), and the complement is perceptual, then the evidential must modify the main clause (as the object does not express a proposition, but rather describes an event):

(8) I saw+direct evidential Mary cry

The evidential would attach to the whole of (8), the truth of *Mary cried* being an immediate logical consequence of its truth.

Now, Garrett (2001) remarks that the Tibetan direct evidential does not readily occur with complement predicates outside what he, following Kratzer (1995), calls 'stage-level' (or, perhaps equivalently, to the active or transient in the sense of Higginbotham (1983)), meaning by this those that are, or are normally, understood to apply to objects from time to time, rather than being permanently included or excluded. This property is characteristic of the complements to perception V as in (8), and thus provides some evidence that the complements are descriptions of events, rather than propositional. It should be possible to test the matter further, but so far as I am aware this task has not been undertaken.

4 Singular or Binary?

I introduce, or anyway give a name to, a very general distinction. Consider a construction ' $\dots\dots$ ', having a certain *content* (which may be more or less complex, as specified further below), and with nothing at all in the position marked by ' $_$ '. Let X be an element introduced in this position, so that we have constructions

$\dots\dots$, and $\dots X \dots$

our question is whether the introduced element X contributes to the *content* of ' $\dots X \dots$ ', or rather constitutes a further remark, on the side as it were, made in addition to that made by ' $\dots\dots$ '. In the first case, I shall say that the construction ' $\dots X \dots$ ' is *singular* with respect to the augmentation of ' $\dots\dots$ ' by X , and in the second case that it is *binary*, with X constituting further material, to be adjoined to whatever content and force attaches to ' $\dots\dots$ '. (Ambiguity is possible, so that it may be in principle that a construction is either singular or binary; in which case the question would arise whether the phrase structure is univocal.)

The full characterization of singular and binary constructions is more complex than that indicated by the sketch of the last paragraph: for we must consider not only the possibility of ambiguity as above, but also whether the host construction ' $\dots\dots$ ' is itself binary; and if so, how X contributes to the whole. However, even the broad classification just given allows for illustrations of the distinction.

The most obvious example of a binary construction is that in which the element X is an appositive relative clause. Consider (9):

(9) Bill, who is my friend, is coming to dinner.

In asserting (9), I have asserted two things: (i) that Bill is coming to dinner, and (ii) that Bill is my friend. The assertions are not equal in weight: intuitively, (ii) is by way of a ‘side remark’ to the main assertion (i). But (ii) is likewise asserted. Suppose the appositive appears with interrogatives or imperatives, said with appropriate force, as in (10) or (11):

(10) Is Bill, who is your friend, coming to dinner?

(11) Let Bill, who is your friend, come to dinner if he wants to!

The force of (10) is that of *asking*, we may suppose, and the force of (11) that of *exhorting*. But the appositive forms no part either of the question or the exhortation; rather, it constitutes an assertion, as it did in (9). These latter observations underscore the correctness of taking the appositive outside the main assertion in (9).

The abstract picture suggested by our example is that in a singular construction, viewed as uttered with a certain force F , the element X of ‘... X ...’ falls under the propositional content uttered with that force; whereas in a binary construction X appears outside ‘... X ...’ under F , and is uttered with a force F' , which may be the same as or different from F . We thus have, for the singular case, something we may put as (12), and for the binary (13):

(12) $F[\dots X \dots]$

(13) $F[\dots \dots]; F'[X]$

Using ‘ A ’ for assertive force, ‘ Q ’ for asking, and ‘ I ’ for the various speech acts that fall under the imperative mood, the abstract forms of (9)–(11) are as in (14)–(16), respectively:

(14) $A[p]; A[q]$

(15) $Q[p]; A[q]$

(16) $I[r]; A[q]$

Besides the appositive relative, various parenthetical remarks clearly form binary constructions with their hosts. Consider (17):

(17) John is going to Hollywood, by the way.

In asserting (17) one asserts that John is going to Hollywood, and that this fact is ‘by the way of’ whatever is under discussion. The parenthetical may be embedded on the surface, say within the antecedent of a conditional; but it is asserted with a force wholly outside the conditional, say as in (18):

(18) If John goes to Hollywood, by the way, he’ll become a star.

Here it is the whole conditional that is ‘by the way.’ Similar remarks apply to cases that Rooryck (2001b) classes as parenthetical, for example (19) and (without the complementizer) (20):

(19) The doors open at 8.00, very likely.

(20) (That) John is a nice fellow, I admit.

With the complementizer, (20) is just a case of inversion, or topicalization. Without it, the construction is necessarily binary. Likewise, (19) must be sharply distinguished from (21):

(21) Very likely, the doors open at 8.00.

inasmuch as (21) is singular, and freely embeddable, for instance as in (22):

(22) John knows that, very likely, the doors open at 8.00.

In (23), however, the adverb attaches preferably, and perhaps only, to the entire assertion:

(23) John knows that the doors open at 8.00, very likely.

The binary nature of appositive relative constructions, and parentheticals such as ‘by the way’, is not undermined by cases where the content of the appositive or parenthetical interacts with that of the main clause. The following example is due to Frege (1892):

(24) Napoleon, recognizing the danger to his right flank, himself led the guards against the enemy position.

Frege worries a bit about whether the main clause and the appositive are to be taken as just separately asserted, or whether, as intuitive understanding would suggest, the truth of the appositive is to be taken as supplying a reason for the truth of the main clause. (In Frege’s system, the latter would imply a shift from direct to indirect reference.) Ultimately, he concludes that the speaker merely ‘hints’ at the connection, so that (24) must be distinguished from, say, (25):

(25) Napoleon, because he recognized the danger to his right flank, himself led the guards against the enemy position.

Abbreviating the main clause by ‘*p*’ and the complement to ‘because’ by ‘*q*’, the structure of (25) is as in (26):

(26) $A[p]$; $A[p \text{ because } q]$.

In (24), however, we have from the linguistic structure just (27):

(27) $A[p]$; $A[q]$.

This circumstance allows incorporation of what Frege calls a ‘hint,’ which we would generally understand today as speaker’s implication, as in the last assertion in (28):

(28) $A[p]$; $A[q]$; $A[p \text{ because } q]$.

(which would, incidentally, fully justify Frege’s further remark that the clauses should be taken ‘twice over,’ once for their ordinary reference, as in the two assertions, and once for their indirect reference, as in the speaker’s implication).

An example of (the very beginning of) more complex recursion is (29):

(29) John, who is my friend by the way, is coming to dinner.

In this case, I believe, the parenthetical can serve to dismiss the implication that John's coming to dinner has anything to do with his being a friend of mine. The sequence of assertions is then as in (30):

- (30) $A[p]; A[q]; A[q$ is by the way of $p]$.

5 Application to Evidentials

I turn now to the question whether evidentials are singular or binary with respect to their host clauses, building upon and endeavoring to put in a general perspective some remarks of Faller, Garrett, and Rooryck in work cited above. Consider again our abstract example (1), repeated here:

- (1) A horse ran+ ε through the village.

and let the evidential ε represent (say), 'Somebody told me.' If (1) is a singular construction, then the evidential falls under the assertive force of the whole, so that we have what we may give as (31):

- (31) $A[\varepsilon [A$ horse ran through the village]]

entirely comparable to the English, 'Somebody told me/They say a horse ran through the village.' If it is binary, however, we expect (32):

- (32) $A[a$ horse ran through the village]_{*i*}; $A[\text{Somebody told me that}]_i$

with anaphoric reference from the object 'that' to the main clause.

The two possibilities above may be discriminated. Garrett (2001) in particular notes, following other studies, that the Tibetan direct evidential (alluding to perceptual evidence, and *pace* the interpretive question noted in Section 5 above) accommodates itself to questions, so that in what we may represent as (33):

- (33) A horse ran+direct-evidential through the village?

the interpretation is: did *you* (the addressee) directly observe a horse run through the village? But this is the conclusion that would be expected if the direct evidential is singular, a higher V or functional head taking the clause as complement.

By contrast, again taking the data from Garrett's study, the indirect evidential appears to be binary. In particular, like the *by the way* of (18) above, it may not be embedded in the antecedent of a conditional. Examples of the following sort are given as ungrammatical:

- (34) *If a horse ran+indirect-evidential through the village, I shall be surprised.

But if the indirect evidential were a singular contribution to the antecedent, these should be fine, and interpretable roughly as in (35):

- (35) If they say/report/tell me that a horse ran through the village, I shall be surprised.

To explain the ungrammaticality of (34), we should have to eliminate two further possibilities: (i) that the indirect evidential could be a binary contribution to the antecedent alone, and (ii) that it could be a binary contribution to the entire conditional. But (i), as in the English cases discussed above, is presumably not coherent; and (ii) may follow from the fact that the evidential is a verbal affix, and so, unlike free-standing remarks, is not syntactically separable from its host clause.

6 Recursion in Binary Constructions

Thus far the binary constructions that we have considered have seen the element *X* of ‘...*X*...’ as constituting the content of a speech act of its own, quite apart from the context provided by ‘..._...’. The question arises, however, whether *X* may be understood interior to some element in the host ‘..._...’, and if so what its effect may be. As we shall see, the matter is moderately complicated, even to the point of possible idiosyncrasy amongst the hosts to the binary construction.

Let us say that a complement-taking predicate *V* (or *be+A*; I use ‘*V*’ for short) is *performatively transparent* if clauses

I [*V*-(object)-complement]

(in the present tense) have the property that ‘saying them makes them true.’ The performatively transparent elements include the classical performatives such as *promise*, but also other *V* that range over acts of speech, and some of those *V* whose first-person present-tense use may reflect first-person authority. Their particular semantic properties perhaps arise because first-person, present-tense uses are accommodated (in the sense of David Lewis (1979)). The following examples illustrate the division:

- (36) *V* that are performatively transparent, and range over acts of speech: *state, say, ask, demand, announce, allow* (as in *I allow that John is a wise man*), *deny*.
- (37) *V* that are performatively transparent, but do not range over acts of speech: *think, guess, suppose, presume, wonder, be afraid, believe*.
- (38) *V* that are not performatively transparent: *know, be sorry, be convinced, regret, demonstrate*.

In general, *V* that are performatively transparent, and only these, take binary constructions in their complements. However, that a *V* is performatively transparent does not imply that it can take any old binary construction in its complement; rather, it is limited to those that are compatible with the kind of events or situations over which it ranges (the others are something like category mistakes; they will be marked with the hash-mark ‘#’).

- (39) Admittedly/Frankly, I don’t give a damn.
- (40) John said/#knew/#denied/announced/concluded that admittedly/frankly, he didn’t give a damn.

(39) is obviously a binary construction, as in (41):

(41) A[I don't give a damn]_i; A[I admit/am being frank in saying that]_i

As (40) indicates, the speech-act V *say* and *announce* accept (39) in indirect discourse, showing that, on the assumption that the V is understood as ranging over assertions (not merely their contents), the complement can split those assertions into parts, just as direct discourse can. On the other hand, *know*, whose complements range only over contents, cannot accept material such as *frankly* or *admittedly*, which are understood uniquely as modifying the assertion:

(42) #John knew that frankly, he didn't give a damn.

We return below to the cases of *deny* and *conclude*.

As noted above, there are V that are performatively transparent, but do not involve speech acts. Consider (43):

(43) I guess/believe I do the cooking/I'll have a cup of coffee.

The V in (43) can be used in either of two ways: (i) as avowing a guess or belief on the part of the subject, or (ii) as expressing (with reservations) an intention of the subject, whose content is given by the complement. To my ear, the presence of the complementizer *that* as in (44) effectively destroys the second reading:

(44) I guess/believe that I do the cooking/I'll have a cup of coffee.

(see further the discussion of 'Slifting' below.) The uses as in (ii) of *guess* and *believe* are not confined to first-person or simple present-tense uses. Thus we can easily have (45), (46) and the like:

(45) Entering the bar, Tex thought he would have a cup of coffee.

(46) Mary guessed she would take the examination in February.

It is characteristic of the use of performatively transparent constructions that one looks so to speak past the higher V to its complement, as if only the complement were uttered. Thus, for instance, the difference between (47) and (48), properly exploited in Faller (2002), exemplifies this property:

(47) I guess/believe I'll do the cooking—No you don't.
(can only deny that you do the cooking)

(48) I know I'll do the cooking—No you don't.
(can (only?) deny knowledge)

Faller observes that one of the evidentials of Cuzco Quechua is like *guess* or *believe* in (47) (the other, derived from a 'past tense' marker, signifying, in Faller's words, that 'the described eventuality is located outside the speaker's perceptual field,' is presumably fully embeddable).

Returning to the speech-act V, an extreme example of the phenomenon of looking past the higher V is found with V such as *state*, *deny*, *affirm*, and the like. These V give rise to the 'performadox' of Lycan (1984): a person who testifies as in (49)

counts as speaking truly or falsely depending just on whether he has ever been a member of the Communist Party; but what she *says* is only that she denies it!

(49) I deny that I have ever been a member of the Communist Party.

In partial explanation, we may say: the speaker of (49) thereby speaks truly (by accommodation); so in saying that she denies that *p*, she has denied it; thus she has represented herself as believing (perhaps even as knowing) that not-*p*; so she speaks truly if not-*p*, falsely otherwise. The explanation carries over to third-person reports of statements such as (49), as in (50):

(50) He denied that he had ever been a member of the Communist Party (by saying, “I deny etc.”); but he lied.

The two types of performative transparency just examined—that with *guess* and *believe* on the one hand, and that with *state* or *deny* on the other—are syntactically and semantically distinct. The latter, with the ‘performadox’ V, shows no syntactic peculiarities, in that the complementizer is present; moreover, the basis for the transparency lies in the semantics of the V itself. The former, however, involve special uses of the higher V, and although exhibiting communication of the normal sort between the higher clause and its complement (as witness the sequence of tense in (46), for instance, or the possibility of quantifying into the complement), effectively detaches the complement from the host. But now it follows, assuming that such syntactic communication as binding and sequence of tense apply only to structures under one roof, so to speak, that the apparently binary structure must after all be unified under one head (i.e., it is not merely a sequence of assertions, but a structure in its own right). Then (47) would be as in (51):

(51) *T*[I’ll do the cooking]_i; *A*(that_i is my guess)

with some force, call it *T* for tentative advancement, attaching to the complement (the speaker of (47) does not *assert* that he will do the cooking). A curious consequence, which supports (51), is that, whereas ordinary complements can be replaced freely by nominals or demonstratives, indicating the argument provided by a host clause, such replacement is not possible with *I guess*, *he believes* and similar locutions, in the interpretation in question. To give but one example, note the likeness of the interpretations of (52) and (53), the first of which evidently involves *guess* only in the interpretation *make a guess*:

(52) Mary said that there were 3,172 jellybeans in the jar, and John guessed that (too).

(53) Mary said that John would do the cooking, and John guessed that (too).

The only interpretation of (53) is that John *made a guess* that he would do the cooking. The alternative interpretation, that John ‘guessed he would do the cooking’ in the sense, e.g., of having resignedly said to himself, ‘I guess I do the cooking,’ is excluded. This consequence follows if (a) *guess* in the intended interpretation takes a proposition *tentatively advanced* as its complement, to whose content the object position is anaphoric, as in (51), and (b) the demonstrative anaphor *that* cannot carry

any element of force; for then the absence of any such element rules out the intended interpretation of the host V *guess*.³ The larger syntax of these constructions, which would be expected to host recursion to arbitrary depth, is a matter for further investigation.

7 Slifting Diagnostics

Some information about the nature of binary constructions may come from cases called *Slifting*, where the complement, minus complementizer, appears on the left, with the host in final position, as in (54):

(54) There would be blood on the moon, well he knew.

Rooryck (2001b) considers all of these to be ‘parentheticals,’ and he thinks of them as within the orbit of evidential constructions. But, as (54) already illustrates, they need not involve V of saying or thinking. Also, they need not involve overt subjects:

(55) The book would sell well, it was widely assumed.

Moreover, Slifting is not only a root phenomenon, inasmuch as a Slifted construction can appear as the complement (only) to a performatively transparent host:

(56) John said/#believed that Mary was a great singer, he would have to admit.

Many cases of Slifting are binary constructions, but (for example) (55) is not, as the speaker is not asserting that the book would sell well. On the other hand, the hosts for Slifting are limited:

(57) There were books on the table, he said/*denied.

(58) The earth went around the sun, it was believed/*known.

I conclude, then, that these matters are mostly orthogonal to consideration of evidentials.

8 Final Questions

Besides the question whether the objects of certain evidentials are proposition-like or event-like, and how far they may interact with first-person authority, the outstanding question, I believe, is the extent to which they are singular or binary constructions. I have illustrated very sketchily some binary constructions in English, combining these with some observations in the literature on evidentials. Issues arise even about the syntax of these, on which I have here remained agnostic; even so, the variety of data suggests that it will not be so simple a matter to explain their nature, and especially the nature of the hosts to embedded binary constructions. I conclude by summarizing the lines of research thus far suggested.

If an evidential is *singular*, then we expect it to function as a kind of grammaticalized main V. So we might have

John feels (that he's) dizzy (personal experience).
 Mary saw a horse run through the village.
 John determined that a horse ran through the village (he saw hoofprints).
 Mary heard that a horse ran through the village.

In that case, the evidential clause should be embeddable under anything whatever.

If an evidential is *binary*, then we expect it to be embeddable only under the right kind of host V, if at all. As mentioned above, the impossibility of embedding within the antecedent of a conditional is taken as a diagnostic for 'illocutionary' status by Faller (in Cuzco Quechua). Garrett classifies the Tibetan 'indirect' evidential as 'performative' on similar grounds. Perhaps we may indicate the relevant contrast through these examples:

(I'm) Dizzy!
 A horse ran through the village, from what I saw.
 A horse ran through the village, I have determined.
 A horse ran through the village, (so) I understand.

In the last three cases the speaker asserts that a horse ran through the village. All three are embeddable under the right host, but not (with interpretation in the embedded clause) under 'I am convinced that'.

Almost all of the literature on evidentials concentrates upon ordinary assertions about what's going on in the world. I have already remarked cases of first-person authority, which should in some ways be special. One could add to these: mathematical facts; generally known empirical matters; and so on. The crucial point, however, is whether (or which) evidentials are singular, functioning both syntactically and semantically as heads taking complements, which function perhaps syntactically but not semantically in that way (perhaps because the head is performatively transparent), and which are analogous to adverbials, or to the Slifting cases. Only as this is sorted out do 'evidentials' form a category.⁴

Notes

1. The phenomenon is illustrated below. Recent literature has concentrated on the basis for it, and whether that basis is compatible with a generally anti-Cartesian view of our knowledge of mental states, and of the contents of those states. Important sources include Davidson (1984), Heal (2001), and references cited in those works. For our purposes, these controversies are not critical: it will be sufficient to note that the cases falling under first-person authority are all of them cases of knowledge without evidence.
2. It would take us too far afield to examine the various cases here. But I note that epistemic *must* retains its significance even when its host clause is not asserted, as in (i):
 - (i) John believes that Mary must be in her office.
 Furthermore, (ii) is an absurd, apparently self-contradictory, assertion:
 - (ii) Mary must be in her office, but perhaps she isn't.

It is doubtful that any pragmatic explanation will suffice to explain the status of (ii), as the absurdity carries over to the unasserted complement of (iii):

(iii) John believes [that Mary must be in her office, but perhaps she isn't].

(iii) has it that John's beliefs are contradictory, so it would appear. For these and other reasons, it may be preferable to assign to epistemic *must* simply the role of characterizing the proposition to which it attaches as a conclusion drawn from some premisses available to the relevant agent (as in (4)–(5) below). Contrary to the view that *must* is an epistemic modal restricting the circumstances of evaluation of the proposition, it would follow that, for the agent, the falsehood of the proposition is no more epistemically or doxastically possible than in the simple case. This revision would allow, for instance, epistemic *must* to be used in drawing the conclusion of a mathematical proof, properly I believe.

3. The interpretation (51) of (47) is *paratactic* in the classical sense; i.e., it detaches the complement content from the host clause. At the same time, as noted above, these elements must be united insofar as such constraints as sequence of tense are observed.
4. The article has grown from a series of presentations, first as part of the Workshop *Language Under Uncertainty* at Kyoto University, January 2005, and later from talks at the Rutgers University Center for Cognitive Science, and at the University Ca' Foscari, Venice. I am grateful to Professor Yukinori Takubo for the opportunity in Kyoto, to Ernest Lepore at Rutgers and Guglielmo Cinque at Venice for their invitations, and to the various audiences for discussion. Above all it's a pleasure to be able to have it appear in a volume for Ernie, who has done so much to advance the hybridization, and consequent cross-fertilization, of Linguistic and Philosophical issues and interests.

References

- Botne, R. (1997). Evidentiality and epistemic modality in lega. *Studies in Language*, 21, 509–29.
- Davidson, D. (1984). First person authority. *Dialectica*, 38, 101–111.
- Faller, M. (2002). Remarks on evidential hierarchies. In D. I. Beaver et al. (Eds.), *The construction of meaning*. Stanford: CSLI Publications. pp. 89–112.
- Frege, G. (1892). “Über Sinn und Bedeutung.” *Zeitschrift für Philosophie und philosophische Kritik*, 100. pp. 25–50. Translated as “On Sense and Reference” by Max Black in Peter Geach and Max Black (eds. and trans.), *Translations from the Philosophical Writings of Gottlob Frege*. Third edition. Oxford: Blackwell, 1980.
- Garrett, E. (2001). *Evidentiality and assertion in tibetan*. Unpublished doctoral dissertation, UCLA.
- Heal, J. (2001). On first-person authority. *Proceedings of the Aristotelian Society*, 102(1), 1–19.
- Higginbotham, J. (1983). The logic of perceptual reports: An extensional alternative to situation semantics. *The Journal of Philosophy*, 80(2), 100–127.
- Higginbotham, J. (2003). Remembering, imagining, and the first person. In A. Barber (Ed.), *Epistemology of language* (pp. 496–533). Oxford: Oxford University Press.
- Izvorski, R. (1997). The present perfect as an epistemic modal. In A. Lawson (Ed.), *Proceedings of Semantics and Linguistic Theory (SALT) VII*. Ithaca, NY: CLS Publications, Cornell University. pp. 171–188.
- Kratzer, A. (1995). Stage-Level and Individual-Level Predicates. G. Carlson et al. (Eds.), *The generic book* (pp. 125–175). Chicago: The University of Chicago Press.
- Lecarme, J. (2003). Nominal tense and evidentiality. In J. Guéron & L. Tasmowski (Eds.), *Tense and point of view* (pp. 277–299). Paris: Presses de l'Université Paris X-Nanterre.
- Lewis, D. (1979). Scorekeeping in a language game. *Journal of Philosophical Logic*, 8, 339–359.
- Lycan, W. (1984). *Logical form in natural language*. Cambridge, Massachusetts: The MIT Press. A Bradford Book.

- Rooryck, J. (2001a). Evidentiality, Part I. *Glott International*, 5(4), 125–133.
- Rooryck, J. (2001b). Evidentiality, Part II. *Glott International*, 5(5), 161–168.
- Speas, M. (with Carol Tenny) (2004). Evidential paradigms, world variables, and person agreement features. *Italian Journal of Linguistics*, to appear.
- Willett, T. (1988). A cross-linguistic survey of the grammaticalization of evidentiality. *Studies in Language*, 12, 51–97.

The Direct Expression of Metaphorical Content

Marga Reimer

Abstract According to a direct expression account of metaphor (Bezuidenhout, 2001), “what is said” by a metaphorical utterance of a sentence is different from what is said by a literal utterance of that same sentence. Thus, what is said by a metaphorical utterance of “Jesus was a carpenter” is different from what is said by a literal utterance of that same sentence. The linguistic contents of the two utterances are (in other words) different. Further, on a direct expression account of metaphor, what is said by a metaphorical utterance of a sentence is expressed *directly*; it is not *inferred from* some other, putatively more fundamental, content. Such a view is at odds with the semantic minimalism argued for by Cappelen and Lepore in their (2005) *Insensitive Semantics*. For it entails the view that (contra Cappelen and Lepore) the linguistic content of an utterance depends (inter alia) on how the uttered sentence is being used: literally or non-literally. In this paper, I do two things. *First*, I argue that certain arguments that Cappelen and Lepore would likely raise against a direct expression view of metaphor won’t work. *Second*, appealing to considerations used in support of my first point, I argue that one of Cappelen and Lepore’s central arguments against radical contextualism (the view of linguistic communication that motivates a direct expression account of metaphor) is based on an uncharitable interpretation of that view. The net result is, I hope, a better understanding of both the direct expression account of metaphor and the more general view of linguistic communication that underlies it.

Keywords Metaphor · radical contextualism · semantic minimalism

1 Preliminaries

Today metaphor is widely recognized among philosophers of language as a pervasive feature of natural language. However, until the publication of Max Black’s “Metaphor” (1962), philosophical discussions of metaphor, few that there were, focused almost exclusively on literary metaphor. Philosophers were apparently

M. Reimer
Department of Philosophy, University of Arizona, Tucson, Arizona 85721-0027, USA
e-mail: reimer@u.arizona.edu

unaware of, or simply uninterested in, the extent to which metaphor pervades ordinary discourse. Yet that metaphor can be found in the ordinary conversations of ordinary people is, or at least should be, uncontroversial. We've all heard about exams that are "nightmares," meetings that "drag on" for an eternity, and "narrow-minded" administrators who make their decisions "with blinders on." Perhaps less well-known is the fact that metaphor can also be found in discourse that is scientific ("forces" and "strings"), mathematical ("rational" and "irrational" numbers), political ("right" and "left"), economic ("bull" and "bear" markets), technological ("virus" and "mouse"), and even philosophical ("correspondence"¹ and "coherence" theories of truth, "wide" and "narrow" content).

Metaphor is not only a ubiquitous linguistic phenomenon, it is also a philosophically interesting one. It raises questions concerning (inter alia) content, communication, and truth.² Such questions include:

- (i) What is the *content* of a metaphorical utterance? In particular, is the content of such an utterance any different from the content of a literal utterance of the same sentence?
- (ii) How do speakers *communicate* via metaphor, given that the content of a metaphor presumably differs from the literal content of the words uttered?³
- (iii) Are the "insights" ("visions" or "viewpoints") characteristically prompted by especially effective metaphors reducible to ordinary *propositional truths*, or are they different in kind from such truths?

My concern in this paper will be with questions (i) and (ii) and, more specifically, with a *direct expression* response to these questions. A direct expression theorist about metaphor (such as Bezuidenhout, 2001) makes two distinct claims. *First*, the content expressed by a metaphorical utterance – "what is said" by such an utterance – is the (propositional) content the speaker intends the utterance to express. Thus, the content of such an utterance is *not* the content that would have been expressed by a literal utterance of the very same sentence. *Second*, the content of a metaphorical utterance is expressed *directly*; it is not inferred from some other, putatively more fundamental, content.⁴ Consider, for example, Eliot's (1917) "I have measured out my life with coffee spoons". Suppose that what Eliot intended to convey via that particular line is something to the effect that the protagonist has led an excessively cautious life. According to the direct expression theorist, such is the content of the metaphorical utterance, a content *directly expressed* by the utterance. Nevertheless, the words uttered have their usual lexical meanings,

¹ For an account of the metaphor of correspondence as used by correspondence theorists, see Reimer (2006).

² For details, see Reimer and Camp (2006).

³ Donald Davidson (1978), who argues that metaphors have no meanings beyond their literal meanings, would claim that this question has a false presupposition. For a defense of Davidson, see Reimer (2001).

⁴ For similar views, see Carston (2002), and Sperber and Wilson (1985/86, 1986).

meanings the interpreter draws on in arriving at the utterance's content. How is the interpretative process effected? Through a combination of lexical underdetermination and contextual supplementation. For it is not only indexicals (like "you" and "me") and demonstratives (like "this" and "that") that require contextual supplementation to determine a content. The same is true for linguistic expressions across the board, including those used metaphorically. So says the direct expression theorist of metaphor.

My central aim in this paper is a simple, if ambitious, one: to understand the account of linguistic content/communication that underlies a direct expression account of metaphor. In attempting to achieve this aim, I consider some data involving indirect speech reports of both literal and metaphorical utterances. I suggest, contra recent claims by Cappelen and Lepore (2005), that such data do nothing to undermine the sort of "contextual" approach to linguistic content/communication that motivates a direct expression account of metaphor. My intention, however, is not so much to *defend* a direct expression account of metaphor as it is to get clear on the general approach to linguistic content/communication that underlies it. Only when such an approach is properly understood will it be possible to provide a fair assessment of the direct expression account of metaphor. In the final section of the paper, I move away from the particular case of metaphor to linguistic communication more generally. Specifically, I consider Cappelen and Lepore's claim that contextualist approaches to linguistic communication are at odds with the fact that such communication is, by and large, a remarkably successful enterprise. Drawing on the preceding discussion of the direct expression account of metaphor, I argue that Cappelen and Lepore's claim is based on a fundamental misunderstanding of contextualism. Charitably interpreted, contextualism is quite compatible with the impressive success of ordinary everyday linguistic communication.

2 Some Data

What follows is a bit of data relevant to the illustration, motivation, and evaluation of a direct expression account of metaphor. I should warn the reader that the description that follows is quite detailed, although no more so than necessary given the goals of the present paper.

There are seven members of the Creative Writing Program at the local community college: Alwood, Babs, Cappy (the head), Dana, Everett, Freda, and Gregor. Not surprisingly, they all have a wide variety of beliefs about one another. And while they differ with respect to many of their beliefs about their fellow colleagues, there is some degree of overlap, just as one would expect.

Let's begin by considering the first of two cases, which I will call **Case 1**. Both Alwood and Cappy know Babs as the eccentric owner of seven laptops who always brown-bags her lunch. They both regard her poetry as surprisingly naive for that of a 50 year-old woman who was once CEO of a major computer corporation. Against this context, the following exchange takes place.

Alwood: Do you think that Babs' poetry is sophisticated enough for college students?

Cappy: Are you kidding? Babs scribbles her poems with the Crayolas she keeps stashed in her Mickey Mouse lunch box.

Dana, Everett, and Freda are all in an adjoining room when they overhear Cappy's utterance. They do not, however, hear the question that prompts the utterance. Dana knows about Babs' laptops and lunch bags, and also knows about Cappy's opinion of Babs' poetry. Neither Everett nor Freda know about any of these things. Dana correctly interprets Cappy's utterance as *metaphorical*; Everett incorrectly interprets it as *literal*; Freda has no idea whether the utterance was meant literally or metaphorically, but is certain that it was intended as some sort of insult. Gregor, who did not overhear Cappy's remark, knows about Babs' laptops and lunch bags, and suspects that Cappy finds her poetry to be excessively naive (as he himself does). Later that day, the following three exchanges take place:

Gregor: Does Cappy think that Babs' poetry is mature enough for an audience of young adults?

Dana: Maybe for an audience of Kindergartners! He said that Babs scribbles her poems with the Crayolas she keeps stashed in her Mickey Mouse lunch box.

Gregor: Do you have any idea whether Babs might like this 1930s Disney poster for her office? I notice that Walt himself signed it on the back with a red crayon.

Everett: Definitely. She collects that sort of thing. Cappy said that Babs scribbles her poems with the Crayolas she keeps stashed in her Mickey Mouse lunch box.

Gregor: Does Cappy ever talk about people behind their backs?

Freda: Yeah. The other day he said to Alwood that Babs scribbles her poems with the Crayolas she keeps stashed in her Mickey Mouse lunch box.

Here's the intuitive data: Dana's and Freda's speech reports are *true*. Everett's speech report seems neither clearly true nor clearly false, although there is something misleading about it. It's certainly *infelicitous*.

Now let's consider a second case, which I will call **Case 2**. It's 10 years later and Babs has changed considerably. She has become known for her esoteric, sexually charged, poetry. She has amassed an impressive collection of Disney memorabilia, which includes a Mickey Mouse lunch box that has replaced the brown bags. She is no longer seen using her laptops and insists that cell-phones are the "scourge" of modern society. In fact, she is suspected by both Alwood and Cappy of having become techno-phobic. Against this context, the following exchange takes place.

Alwood: Is Babs still writing out her poems out manually, or has she finally started using her laptops again?

Cappy: The former, alas. Babs scribbles her poems with the Crayolas she keeps stashed in her Mickey Mouse lunch box.

As in **Case 1**, Dana, Everett, and Freda overhear Cappy's utterance, but not the question that prompts it. And again, Gregor overhears neither part of the exchange. Dana incorrectly interprets Cappy's utterance as *metaphorical*. This interpretation seems only natural to her, as she is unaware of the changes Babs has gone through. Everett correctly interprets Cappy's utterance as *literal*, for he is aware of Babs' changes and knows that Cappy is as well. Freda, who knows comparatively little about Babs or the changes she has undergone, has no idea whether the utterance was meant literally or metaphorically, but has no doubt it was intended as some sort of insult. Later the same day, the following four exchanges take place:

Gregor: Does Babs still spew out that childish drivel about cotton candy, lost puppies, and doe-eyed children?

Dana: Sure does. Cappy said that Babs scribbles her poems with the Crayolas she keeps stashed in her Mickey Mouse lunch box.

Gregor: Does anyone have a red pen or marker I could borrow?

Everett: Not me, but I'll bet I know who could lend you a red crayon. Cappy said that Babs scribbles her poems with the Crayolas she keeps stashed in her Mickey Mouse lunch box.

Gregor: Does Cappy have a problem with Babs?

Freda: Yeah. The other day he said to Alwood that Babs scribbles her poems with the Crayolas she keeps stashed in her Mickey Mouse lunch box.

Gregor: Has Cappy always had a problem with Babs?

Freda: Yeah. For years he's been saying that Babs scribbles her poems with the Crayolas she keeps stashed in her Mickey Mouse lunch box.

Intuitively, Everett's report as well as Freda's reports are *true*. Although neither clearly true nor clearly false, there is something misleading about Dana's report. It's certainly *infelicitous*.

3 A Direct Expression Account of Cappy's Metaphorical Utterance

How should Cappy's metaphorical utterance be analyzed? An advocate of a direct expression account (of the sort proposed by Bezuidenhout, 2001) might provide the following sort of analysis. As uttered by Cappy, the sentence "Babs scribbles her poems. . ." expresses one linguistic content in **Case 1** and a different linguistic content in **Case 2**. This is not to say that the sentence itself exhibits any sort of lexical or structural ambiguity. Nevertheless, what is "said" in **Case 1** is quite different from what is "said" in **Case 2**. Cappy's metaphorical utterance (his first utterance) expresses a content (which is asserted) to the effect that Babs' poetry is child-like, perhaps inappropriately so (i.e., childish). It does this by way of what are termed "ad hoc" concepts, concepts constructed "on the spot," as it were.⁵

⁵ For details, see Bezuidenhout (2001).

Ad hoc concepts are constructed by way of pragmatic processes like enrichment and loosening, processes that affect the content expressed, not just implicatures. “Enrichment” occurs when expressed content results from a *narrowing* of semantically encoded content; “loosening” occurs when expressed content results from a *broadening* of semantically encoded content.

Let’s consider some examples from Bezuidenhout (2001). While the interpretation of an utterance of (1),

(1) I went to the bank.

would likely call for enrichment, the interpretation of an utterance of (2),

(2) It is silent in here.

would likely call for loosening. For typical utterances of (1) involve tacit reference to some particular time (such as the day of the utterance), while typical utterances of (2) could easily be true even in circumstances where subtle sounds are audible. Thus, while the content expressed by an utterance of (1) might be something like (3),

(3) I went to the bank today.

that expressed by an utterance of (2) might be something like (4),

(4) It is unusually quiet in here.

On Bezuidenhout’s (2001) direct expression account of metaphor, metaphorical content results from the formation of ad hoc concepts generated by way of *loosening*.⁶ Let’s consider Cappy’s metaphorical utterance about Babs. As all three of those who overhear Cappy’s utterance know, someone who “scribbles with . . . Crayolas” engages in the one of the most child-like creative processes imaginable. In interpreting Cappy’s utterance metaphorically, Dana (in **Case 1**) is prompted to “loosen” the semantically encoded content of the predicate phrase. The result is the formation of an ad hoc concept, a more *general* concept of a creative process that (like childish scribbling) is remarkable in its total lack of sophistication. It is this contextually broadened concept that is being used by Cappy (in his first utterance) to characterize Babs’ poetry. He employs the ad hoc concept *scribbling with. . . Crayolas**⁷ to characterize Babs’ poetry as having been produced by a remarkably unsophisticated process.⁸ Given the ability of natural language speakers to invoke context in the construction of ad hoc concepts from semantically encoded content, there is no need to assume that the interpreter *first* processes and *then* rejects a literal interpretation of the entire utterance. Indeed, there is no need to assume that the utterance expresses, in *any* theoretically important sense, what a literal utterance of the same sentence would have expressed.

⁶ For similar views, see Carston (2002) and Sperber and Wilson (1985/86, 1986).

⁷ Following current convention, an asterisk indicates that the specified phrase refers to an ad hoc concept generated by some pragmatic process(es).

⁸ Bezuidenhout (2001) provides several, more detailed, illustrations of the interpretation of live metaphor by way of the formation of ad hoc concepts.

4 A Reason for Rejecting a Direct Expression Account?

Although Cappelen and Lepore (2005) ('C&L' from here on) do not address metaphor in their work, their twin doctrines of "semantic minimalism" and "speech act pluralism" appear to be incompatible with a direct expression account of metaphorical content/communication. According to the former, only those expressions *standardly* treated as context-sensitive (i.e., indexicals and demonstratives) are in fact context-sensitive. According to the latter, even utterances of sentences *not* containing genuinely context-sensitive expressions typically result in the production of a variety of different speech acts (depending upon, inter alia, the communicative intentions of the speaker). Thus, while a sentence used metaphorically might well result in a metaphorical *assertion*, its only genuinely *semantic content* is its "minimal" content, a compositionally determined content arrived at by treating only a very narrow class of linguistic expressions as context-sensitive.

In opposition to a direct expression account of metaphor, C&L might argue that lexically/syntactically unambiguous expressions, used metaphorically, do not pass their various tests for context-sensitivity, thereby undermining such an account. However, my goal in this section is to show that such expressions, used metaphorically, arguably pass at least the first of C&L's three tests for context-sensitivity.

Before applying the test in question to the data described in Section 2, something needs to be said about both the target and overall strategy of C&L's context-sensitivity tests. Concerning their target, the tests are designed to undermine "contextualism," in both its moderate and radical forms. Contextualism is the view that context-sensitive expressions include more than those standardly treated as context-sensitive, they include more than just indexicals and demonstratives. According to *radical* contextualism, extra-linguistic context *invariably* plays a role in the determination of linguistic content. The claim here is a categorical one; it's not that some "special" linguistic expressions, not widely recognized as context-sensitive, are in fact context-sensitive. Such is the claim of *moderate* contextualists. The claim made by radical contextualists is that *all* linguistic expressions are context-sensitive.⁹ Why such pervasive context-sensitivity? Because (according to the radical contextualist) the sorts of features capable of influencing linguistic content by combining with essentially indeterminate linguistic (type) meaning, are ubiquitous. As noted by C&L, such features include (inter alia) "perceptual inputs, accompanying activities, previous conversational context, purpose of conversation, nature of audience, and assumptions shared by conversational participants."¹⁰ My interest is in radical contextualism for it is this form of contextualism that underpins a direct expression account of metaphor.

The rhetorical strategy used by C&L in applying their tests for context-sensitivity is to consider what initially appear to be the *most* promising cases for contextualism

⁹ None of which is to say that all linguistic expressions are context sensitive *in just the same way*. See Bezuidenhout (2002, 2006).

¹⁰ Cappelen and Lepore (2005), p. 93.

and then argue that a contextualist construal of even these cases is highly implausible (if not flat-out false). My strategy here effectively reverses that of C&L. I will consider what initially appear to be the *least* promising cases for contextualism, namely, utterances involving live metaphor. In cases of live metaphor, it seems almost obvious that what the speaker *says*, what her words directly express, is not what the metaphor means. What the speaker says, in such cases, is a matter of ordinary literal content rather than constructed metaphorical content. Or so it appears. However, I will argue that C&L's first test for context-sensitivity does nothing to undermine a contextualist account of such cases. It does nothing to undermine the view that live metaphors directly express their context-sensitive metaphorical contents. And if the test doesn't undermine a contextualist account of the *least* promising cases for such an account, it's arguably not going to undermine a contextualist account of the *most* promising cases for such an account. However, I will not rely on rhetoric to make the latter point, but will argue for it independently.¹¹

So much for background, let's now look at the details of the test in question, which is C&L's first test for context-sensitivity. This test relies on the notion of an "inter-contextual disquotational indirect report." C&L define such a report as follows:

Take an utterance *u* of a sentence *S* by a speaker *A* in context *C*. An Inter-Contextual Disquotational Indirect Report of *u* is an utterance *u'* in a context *C'* [where *C'* is relevantly different from *C*] of 'A said that *S*'.¹²

According to what I will call the "Indirect Report Test" (or IRT):

If the occurrence of an expression *e* in a sentence tends to block disquotational reports (i.e., render such reports false), then you have evidence that *e* is context-sensitive.¹³

Now let's apply the test to the predicate phrase "scribbles her poems. . .". Let this phrase be *e*. Let *u* refer to either of Cappy's two utterances. Then, the reports of Dana, Everett, and Freda would all appear to be inter-contextual disquotational indirect reports of *u*. If these reports are *false*, then there is evidence that *e*, the phrase "scribbles her poems. . .", is context-sensitive. If the reports are *true*, the test fails to provide evidence for the context-sensitivity of *e*.

Before looking critically at IRT, let me make four provisional points regarding its application to the data in question, **Cases 1** and **2**. (The first two of these points will be challenged below.) The *first* point concerns the contexts of utterance. The contexts in which Dana, Everett, and Freda make their reports are *relevantly different* from those in which Cappy says what he says about Babs. Additionally, the contexts of Cappy's two utterances are *relevantly different* from one another. To say that the contexts are "relevantly different" (as C&L put it) is simply to say

¹¹ See discussion of John's utterance of "I'm ready" in Section 5 below.

¹² Cappelen and Lepore (2005), p. 88.

¹³ Cappelen and Lepore (2005), p. 88.

that the sorts of contextual features that are thought, *by the contextualist*, to influence linguistic content (including those just specified) vary across the contexts in question. All of this is just to say that the reports in question appear to satisfy one of the conditions of the test: they appear to be made in contexts that are relevantly different from those of the original utterances. The *second* point is that the reports made by Cappy's colleagues should all be *false*, if the contextualist is right. They should all be false because the sentence "Babs scribbles her poems. . ." is alleged (by the contextualist) to express *different* contents in each of the (relevantly different) contexts in which it is uttered. So what Cappy says via either of his utterances of the sentence "Babs scribbles her poems. . ." should be different from what Cappy's colleagues express by that very same sentence when they use it to report what he said. The *third* point is that, with the possible exception of the infelicitous reports of Everett and Dana, the reports of Cappy's colleagues are *true*, just as a minimalist account would predict. For according to the minimalist, the sentence used by Cappy in saying what he says about Babs expresses *the same minimal semantic content* in the contexts in which it is initially uttered (by Cappy) and subsequently reported (by Dana, Everett, and Freda). What content is that? The compositionally determined content arrived at by treating only a very narrow class of expressions as context-sensitive, namely: *Babs scribbles HER poems with the Crayolas SHE keeps stashed in HER Mickey Mouse lunch box*. The *fourth* point is that the first report of Everett and the second report of Dana are *infelicitous*. Recall that while Everett mistakes Cappy's metaphorical utterance for a literal one, Dana mistakes Cappy's literal utterance for a metaphorical one.

5 A Direct Expression Response

There are any number of responses that the direct expression theorist might make here.¹⁴ I'll give what I take to be the most promising response. It has three parts to it. *First*, I argue that the contexts of the intuitively true reports of Dana and Everett are *not* "relevantly different" from those of Cappy's utterances. This means that those reports do not satisfy the conditions of IRT. *Second*, I claim that Freda's intuitively true reports are in fact *quotational* rather than *disquotational* reports. For this reason, Freda's reports do not satisfy the conditions of IRT either. *Third*, I tentatively suggest that the contexts of the infelicitous reports of Dana and Everett *are* relevantly different, and that those reports are arguably *false*, just as contextualism predicts.

The first question that needs to be asked of the contextualist is: What makes two contexts "relevantly different" from one another? To answer this question, we must first answer the question: What are the relevant features of a given context of utterance? In other words, what features of a context are relevant to determining the

¹⁴ See C&L, pp. 97–99 for details.

linguistic content of the expressions used in that context? Here's a tentative proposal, couched in terms of necessary conditions:

The relevant features of a given context include only those features *actually exploited* by the speaker in her effort to be understood as intended.

The features "exploited" by the speaker are simply those that she *relies on* in her attempt to be understood as intended. Her reliance on such features needn't be something that she is actually *conscious* of. For instance, when my 10 year-old daughter says to me, "I'd love a piece of juicy red watermelon," she is relying, in her communicative endeavors, on the fact that the color of a watermelon is typically identified with the color of its *flesh*, rather than with the color of its *skin*. However, it is doubtful that she is *consciously aware* of relying on this particular fact.

I regard the foregoing analysis of "relevant features" as a plausible one on two grounds. First (and most importantly), it is consistent with much of the contextualist literature that I have read, where the mental state of the speaker (broadly construed) is *critical* in determining linguistic content. Second, it renders contextualism a reasonably plausible approach to linguistic content/communication. It thus conforms to the often disregarded "principle of charity." But if the proposed interpretation is accurate, it suggests that Cappelen and Lepore are confused when they claim that, according to the contextualist, "relevant" features of the context include, "perceptual inputs, accompanying activities, previous conversational context, purpose of conversation, nature of audience, and assumptions shared by conversational participants."¹⁵ These features are arguably relevant only in cases where they are *actually exploited* by the speaker in her effort to be understood as intended. For the contextualist, perceptual inputs and the like are perhaps best described as *potentially relevant* contextual features. Thus, suppose Cappelen and Lepore talk about the weather in Tucson before proceeding with philosophical matters. Then, Lepore's remark that "Tucson is the sunniest city in the nation" is not relevant to the content of his subsequent philosophical exchanges with Cappelen unless it is somehow exploited in the communicative endeavors that constitute those exchanges. The remark would, for instance, be relevant if immediately followed with: Now how many indexicals does *that* sentence contain? It would not be relevant if followed, 20 minutes later by: "Let's talk about relevance theory today". This suggests that perhaps the only way to arrive at an accurate answer to the question: What contextual features are relevant to linguistic content?, is to relativize the question to a particular case and then probe the cognitive state of the speaker. One would ask the speaker, post facto: What contextual factors, in such and such a case, were you relying on in your attempt to be understood as intended? (Of course, the ordinary speaker might need a crash course in context and communication before being in a position to provide an informed response to the theorist's query.)

Given the foregoing construal of "relevant" features, "relevantly different" contexts might be tentatively defined as follows:

¹⁵ Cappelen and Lepore (2005), p. 93.

Two contexts are “relevantly different” just in case exploitation of their respective features results in the communication of different contents by way of the same (lexically/syntactically unambiguous) expression.

Although less than precise, this proposal is clear enough to pare down considerably those contexts that count as “relevantly different.” Thus, if Cappelen and Lepore observe a man in a red shirt during one of their exchanges and then observe a woman in a green shirt during another of their exchanges, this does not mean that the contexts of those exchanges are relevantly different. It does not mean that (e.g.) “I enjoyed Anne’s review” as uttered by Cappelen would communicate one content in the first context and another in the second. Yet there certainly appear to be cases where context *does* influence content in much the way the contextualist claims. Consider a case where John’s utterance of “I’m ready” is preceded by “Are you ready for the exam?” Contrast this case with one where John’s utterance of “I’m ready” is preceded by “Are you ready to leave the apartment?” Here, the contexts would indeed be relevantly different, as John would be exploiting features of those contexts to communicate different contents: *I’m ready for the exam*, *I’m ready to leave the apartment*. In both cases, it is the immediately preceding utterance that is exploited. Contra Cappelen and Lepore, it is not obvious that one could correctly report either of John’s utterances via “John said he was ready”. To see this, just suppose that “John said he was ready” is uttered in response to “Is John ready for the exam?” Suppose further that the reporter mistakenly thinks that John meant he was ready for the exam when he said “I’m ready”. Suppose, however, that John had meant he was ready to leave the apartment and had earlier made clear he was *not* ready for the exam by saying, quite explicitly, “I am *not* ready for the exam”. Then intuitively, the report would be false, just as a contextualist analysis of that report would predict.¹⁶ (A parallel scenario could be constructed for a case where John *did* mean that he was ready for the exam, but the reporter mistakenly thought he meant he was ready to leave the apartment.)

Let’s illustrate the foregoing proposal concerning “relevantly different” contexts by applying it to the data described in Section 2, beginning with the intuitively true reports of Dana and Everett. Although the contexts of the reports are indeed different from those of the original utterances, they are not *relevantly* different. They are not relevantly different because the reporters do not exploit features of their respective contexts in the relevant sorts of ways. They do not exploit those features in order to use the predicate phrase “scribbles her poems” to communicate a content different than the one that Cappy uses the phrase to communicate. These reports, the contextualist might claim, are true because the speakers are using the phrase in question to communicate the same content that Cappy uses that phrase to communicate. What content is that? Where Cappy’s utterance is metaphorical, the content is something like *writes with the naiveté of a child*; where Cappy’s utterance is literal, the content

¹⁶ Intuitions to the effect that the report would be true could be explained on the hypothesis that the interpreter is reading the report as quotational. See the discussion below regarding Freda’s intuitively true utterances.

is *scribbles her poems with the Crayolas*. . . So, contrary to what Cappelen and Lepore suggest, radical contextualism (as presented above) does not contradict, but instead predicts, the truth of the intuitively true reports of Dana and Everett.

What about Freda's reports, all three of which seem to be true? This is easily explained by the contextualist. In none of her reports does Freda intend to use the predicate phrase "scribbles her poems. . ." to communicate a content different than either of the two contents communicated by Cappy. This is because, by hypothesis, Freda does not (fully) disambiguate either of Cappy's utterances and so she presumably does not intend to use the phrase in question to express *any* content at all – let alone one that is different than the one expressed by Cappy. She doesn't intend to indicate that content because she doesn't know what it is. Thus, the contexts of Freda's reports are *not* "relevantly different" from those of Cappy's original utterances. What, then, is Freda's communicative intent with regard to the predicate phrase, if not to report what Cappy said?

This brings me to my second point: Freda's reports are not *disquotational* speech reports, they are *quotational* speech reports, reports that might be explicated as:

Cappy said that Babs 'scribbles her poems with the Crayolas she keeps stashed in her Mickey Mouse lunch box'.

Because Freda does not (fully) interpret Cappy's utterance, her reports should arguably be read with tacit quotation marks around the content sentence, or at least around the predicate phrase in that sentence. Freda is reporting on the words Cappy employed in insulting Babs; she is not reporting on the content of the insult itself. This hypothesis would explain the intuition that those reports are true: they *are* true. They are true because Cappy did indeed use the quoted phrase in saying what he did about Babs.¹⁷

Freda herself would probably regard all three reports as equivalent to quotational reports. The hypothesis that the reports are quotational is suggested by Freda's likely unwillingness to paraphrase them. (If you think that she might be willing to do so, then try to specify those paraphrases that she might accept – keeping in mind that she has no idea whether the reported utterances were intended literally or metaphorically.) Analogous cases are easy to construct. Imagine the following two dialogues:

Alwood: What made you finally decide against becoming a pilot?

Cappy: Flying planes can be dangerous. Even the best pilots sometimes crash.

Alwood: What's the chief environmental danger to the countless insects that buzz around airports – is it the exhaust fumes from the planes?

Cappy: Actually, it's not. The chief danger is flying into the windshield of a plane flying in the opposite direction. To a mosquito, flying planes can be dangerous.

¹⁷ An alternative, though not clearly incompatible analysis, is provided by Bezuidenhout (2006). Bezuidenhout accounts for intuitions of truth in such cases by suggesting that "said that" is interpreted as meaning something like *used such-and-such form of words*.

Suppose Freda overhears just a fragment of each exchange – Cappy’s utterances of “. . . flying planes can be dangerous”. Suppose further that she doesn’t know how she ought to disambiguate the sentence in either case, although she recognizes its ambiguity. She might still report that, *in both scenarios*, Cappy said the very same thing, namely, *that flying planes can be dangerous*. Any such report would *sound* true. But how can this be, given that the sentence clearly expresses different propositions in the two different contexts – as even a minimalist would concede? The answer is simple. Since Freda does not disambiguate either utterance, she arguably intends her reports to be interpreted as quotational reports, as equivalent to: Cappy said that “flying planes can be dangerous”. That this is correct is suggested by her likely unwillingness to offer paraphrases of either of her reports – for to paraphrase would be to do what, by hypothesis, she fails to do; it would be to disambiguate Cappy’s utterance.

How, then, does one go about finding the *right* sort of test cases to apply to the thesis that expressions used metaphorically are context-sensitive? One tries to construct cases of *disquotational* speech reports where the reporter uses the content sentence to communicate a content *other than* that communicated by the agent of the original utterance. One then asks whether such reports would be true or false. Examples of such cases are (contrary to what C&L seem to suggest) easy enough to imagine and would arguably include (i) Everett’s infelicitous report of Cappy’s first (metaphorical) utterance and (ii) Dana’s infelicitous report of Cappy’s second (literal) utterance. In (i), Everett intends the content sentence to be read literally, though it was intended metaphorically by Cappy; in (ii), Dana intends the content sentence to be read metaphorically, though it was intended literally by Cappy. In both these cases, the reporter intends to use the content sentence to express a content which is (unbeknownst to her) different than that expressed by the reported utterance. How is the reporter able to do this? By exploiting certain features of the context of utterance. To see this, consider again the relevant reports, together with the immediately surrounding linguistic context.

Gregor: Do you have any idea whether Babs might like this 1930s Disney poster for her office? I notice that Walt himself signed it on the back with a red crayon.

Everett: Definitely. She collects that sort of thing. Cappy said that Babs scribbles her poems with the Crayolas she keeps stashed in her Mickey Mouse lunch box.

Gregor: Does Babs still spew out that childish drivel about cotton candy, lost puppies, and doe-eyed children?

Dana: Sure does. Cappy said that Babs scribbles her poems with the Crayolas she keeps stashed in her Mickey Mouse lunch box.

A moment’s reflection is sufficient to show that the immediately preceding linguistic context is exploited (that is, relied on) by Everett and Dana in their attempts to be interpreted as intended. Everett’s report would constitute a relevant response only if the content sentence were interpreted *literally*; Dana’s report would constitute a relevant response only if the content sentence were interpreted *metaphorically*. The

contexts of both reports are relevantly different, not only from one another, but also from the contexts of Cappy's original utterance. For both reporters exploit features of their respective contexts in ways that enable them to use the same phrase "scribbles her poems. . ." to communicate different contents: one metaphorical, one literal.

There are two additional points that I would like to make about the infelicitous reports of Everett and Dana. *First*, they are disquotational, in contrast to Freda's reports. *Second*, they are arguably false, just as contextualism predicts. Regarding the first point: In contrast to Freda, it is likely that both Everett and Dana would be willing to paraphrase their reports. Everett might offer this as a paraphrase of his report: "Cappy said that Babs jots down her poems with some crayons she stores in her lunch box". And Dana might offer this as a paraphrase of her report: "Cappy said that Babs writes childish poetry". That they might have reported Cappy's utterances in these alternative ways is consistent with the hypothesis that their original reports were disquotational rather than quotational.

Now for my second point: Everett's and Dana's infelicitous reports are arguably false, just as contextualism predicts. A minimalist might concede that the reports are *infelicitous*, but claim that the reports are, strictly speaking, *true*. The minimalist might further claim that intuition supports this construal. In response, a contextualist might claim that such misleading intuitions are easily explained from the point of view of contextualism. Consider again the two exchanges concerning flying planes, and suppose that Freda reports both of Cappy's utterances by way of the sentence "Cappy said that flying planes can be dangerous". Intuitively, you might think, both reports are *true* – and indeed *obviously* true. But suppose that (unbeknownst to the interpreter of her reports) Freda disambiguated Cappy's two utterances and in so doing, misinterpreted the first as meaning *planes that fly can be dangerous* and misinterpreted the second as meaning *the flying of planes can be dangerous*. If these are in fact the meanings she intends to convey in her reports, then those reports are surely false, even though they might *sound* true. Because the false disquotational reports are homonymous with true quotational reports, the interpreter might charitably (though incorrectly) interpret them as true.

The contextualist could provide the same sort of account of the intuitions regarding the infelicitous reports of Everett and Dana.¹⁸ Any intuition to the effect that these reports are, strictly speaking, true has at least two possible sources. Either they are interpreted as quotational reports or the interpreter (here, the theorist) charitably but falsely assumes that Everett's and Dana's interpretations of the reported utterances are accurate. Recall, however, that the contexts of the two reports make it clear that the reporters have *misinterpreted* Cappy's communicative intent. Thus, the most plausible explanation of the intuition that the reports are true is that *the theorist who intuits them as true* interprets them as quotational. They are being interpreted as quotational reports to the effect that Cappy said that Babs "scribbles her poems with the Crayolas she keeps stashed in her Mickey Mouse lunch box". What drives such

¹⁸ Of course, it is not my intention to suggest that syntactic ambiguities are similar in kind to the ambiguities that result from the sort of context-sensitivity that, according to the contextualist, pervades all forms of discourse. My point concerns interpretation only.

interpretation? Why (in other words) aren't the reports interpreted as *disquotational* reports, which is (by hypothesis) how the speaker intends that they be interpreted? Perhaps it is the fact that the disquotational reports, though false, are homonymous with true quotational reports. A deeply entrenched commitment to charity might then prompt the theorist to "hear" the true, if unintended, quotational reading of the report. If this sort of charity-motivated misinterpretation can occur in the "flying planes" scenario, I do not see why it cannot occur in the "scribbles her poems. . ." scenario.

The conclusion should now be clear: Cappelen and Lepore's IRT does nothing to suggest that either radical contextualism, or the direct expression account of metaphor it motivates, are false or in any way misguided. Of course, this is not to say radical contextualism is problem-free and, as we'll see in the following section, Cappelen and Lepore believe that there are other, perhaps more serious, problems with radical contextualism.

6 Linguistic Communication: Miraculous or Merely Incompletely Understood?

After arguing against radical contextualism on the grounds that expressions it claims to be context-sensitive fail to pass their tests for context-sensitivity, Cappelen and Lepore continue their assault. They argue that radical contextualism is at odds with the undeniable success of ordinary linguistic communication; they also argue that radical contextualism is internally inconsistent.

In concluding, I would like to look briefly at the first of these two arguments. Oversimplifying only slightly, the point seems to be this. The radical contextualist claims that linguistic content – the content directly expressed by the uttered words – can be influenced by an enormous variety of extra-linguistic factors. Bezuidenhout (2002) mentions a number of these, several of which are cited by Cappelen and Lepore. They include:

- (i) knowledge that has already been activated from the prior discourse context (if any);
- (ii) knowledge that is available based on who one's conversational partner is and on what community memberships one shares with that person;
- (iii) knowledge that is available through observation of the mutual perceptual environment.

According to Cappelen and Lepore, if radical contextualism were correct, linguistic communication would be a miracle. It would be a miracle because it would require the interpreter to know far more than she could reasonably be expected to know. But linguistic communication is not a miracle; it is commonplace. So radical contextualism is false.

Following Cappelen and Lepore, let's consider two sorts of situations, one in which the interpreter shares a context with the speaker and another in which she

does not. We have already considered both sorts of situations. Let's start with cases involving a shared context: the exchanges between Alwood and Cappy in **Cases 1** and **2**. Here, there is no miracle in Alwood's successful interpretations of Cappy's utterances. This is because, as suggested above, Alwood need be aware of only those features of the context *actually exploited* by Cappy in his attempt to be understood as intended – whether metaphorically or literally. These features include shared beliefs about Babs and the questions which respectively prompt Cappy's responses. They do not include (e.g.) a previous conversation between Alwood and Cappy about Cappy's cancer-stricken mother. Nor do they include the fact that Alwood and Cappy are members of the same rollerblading team. Nor do they include the fact that the aroma of freshly brewed coffee is wafting through the perceptual field shared by Alwood and Cappy. For none of these facts (which are instances of (i)–(iii) above) are exploited in the communicative endeavors that constitute the exchanges in question. Of course, they *might have been exploited* had Alwood opened with a different remark, such as:

- (i) How's the chemo going?
- (ii) After last night's practice, I'm beginning to think we're the two worst players on the team.
- (iii) Apparently, you didn't listen to my request for Vienna Roast; that smells like something you'd get at a truck stop!

Of course, had Alwood opened with any of the foregoing, Cappy's understanding of the remark would require that Cappy be aware of the contextual features being exploited, features which might include a previous conversation regarding Cappy's ailing mother, membership on the same rollerblading team, the aroma of coffee wafting through the air at the time of the utterance.

But what about a case where the interpreter does *not* share a context with the speaker? We looked at several such cases, and the results were just as contextualism would predict. Misinterpretation becomes a very real possibility. This possibility becomes an actuality when Everett misinterprets Cappy's first utterance as literal and when Dana misinterprets his second utterance as metaphorical. These misinterpretations occur because the interpreters are unaware of the features of the context being exploited by the speaker in her attempt to be understood as intended. In particular, they are unaware of the question that precedes the utterance they misinterpret; they are also unaware of mutual beliefs concerning the subject of the utterance (Babs) that are shared by Alwood and Cappy. Of course, even when contexts are *not* shared, misinterpretation isn't inevitable. But where's the argument, from contextualist premises, that it is?

Now there may well be other concerns with explaining linguistic communication on a contextualist model, such as explaining how, exactly, speakers "create" or "construct" linguistic content from contextually-embedded semantically encoded content. How this happens is somewhat of a mystery perhaps, one involving the speaker's truly impressive ability to exploit extra-linguistic context in the attempt to be understood as intended, as well as the hearer's equally impressive ability to interpret the speaker as she intends to be interpreted. But the possession of such abilities,

however impressive, is not miraculous and, in fact, seems in perfect keeping with what is known about the astounding power and breadth of human cognitive capacities generally. Moreover, contextualists (including Bezuidenhout) have provided at least partial explanations of the psychological processes that take the interpreter from indeterminate linguistic (type) meaning and context to communicated content. Perhaps such speculations, whatever their initial plausibility, must await confirmation from the relevant sorts of empirical studies.

However, at least this much should be clear: It is simply a mistake to suppose, as do Cappelen and Lepore, that radical contextualism requires that interpreters know far more than they could reasonably be expected to know. Successful interpretation does *not* require the simultaneous awareness of every *potentially* exploitable feature of a context of utterance. However, if an interpreter is unaware of the contextual features that a speaker is *actually* exploiting in her attempt to be understood as intended, then misinterpretation is certainly a possibility, just as contextualism predicts. C&L are therefore mistaken in claiming that contextualism entails that ordinary linguistic communication is miraculous, as opposed to merely incompletely understood.

Acknowledgments I would like to thank Rob Stainton, Liz Camp, Catherine Wearing, Anne Bezuidenhout, and an anonymous reviewer for helpful comments on earlier drafts of this paper.

References

- Bezuidenhout, A. (2001). Metaphor and what is said: A defense of a direct expression view of metaphor. *Midwest Studies in Philosophy*, 25, 156–186.
- Bezuidenhout, A. (2002). Truth-conditional pragmatics. *Philosophical Perspectives*, 16, 105–134.
- Bezuidenhout, A. (2006). The coherence of contextualism. *Mind and Language*, 21(1), 1–10.
- Black, M. (1962). Metaphor. In *Models and metaphors*. Ithaca: Cornell University Press.
- Cappelen, H., & Lepore, E. (2005). *Insensitive semantics*. Oxford: Oxford University Press.
- Carston, R. (2002). *Thoughts and utterances*. Oxford: Oxford University Press.
- Davidson (1978). What metaphors mean. Reprinted in *Inquiries into truth and interpretation* (1984). Oxford: Oxford University Press.
- Eliot, T. S. (1917). *Prufrock and other observations*. London: The Egoist Ltd.
- Reimer, M. (2001). Davidson on Metaphor. *Midwest studies in philosophy*, 25, 142–155.
- Reimer, M. (2006). The metaphor of correspondence. *Protosociologie*, 23, 93–111.
- Reimer, M., & Camp, E. (2006). Metaphor. In E. Lepore, & B. Smith (Eds.), *The Oxford Handbook of Philosophy of Language*. Oxford: Oxford University Press.
- Sperber, D., & Wilson, D. (1985/86). “Loose Talk,” *Proceedings of the Aristotelian Society*, LXXXVI, 153–171.
- Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Oxford: Blackwell.

The Empirical Case for Bare Demonstratives in Vision

Zenon Pylyshyn

Abstract This chapter draws some parallels between the study of language and the study of visual perception. In both cases there is evidence for modularity and for the productivity of the representations to which they give rise. In both cases it has been recognized that the purely conceptual representations that have been discussed in language and in computational vision are not the whole story. What is missing is some direct connection with token individuals in the world that they describe. In language this connection may be established by a demonstrative term (e.g., this). In the case of vision the parallel would be a singular term that has reference but no descriptive content. I develop a theory of such direct (or demonstrative) reference in vision called Visual Index or FINST theory. I motivate the need for a mechanism of direct reference in vision by describing a number of empirical phenomena. These include the phenomena of single-object advantage, detecting patterns by selecting relevant parts (and executing “visual routines”), and keeping track of individual objects that move among identical distractors (in Multiple Object Tracking experiments). I also discuss the need for a mechanism of direct reference to solve such classical problems of vision as the correspondence problem (computing when several proximal tokens correspond to the same distal object) and the binding problem (establishing when several properties that occur in a visual scene are properties of the same object).

Keywords Visual tracking · multiple object tracking · demonstrative reference · FINST · Visual Index · binding problem · individuals · modularity

1 Background: Representation in Language and Vision

One of the most important ideas that developed in the late twentieth century, and for which Chomsky, Fodor and Newell/Simon can take much of the credit, is realism

Z. Pylyshyn

Rutgers Center for Cognitive Science, Rutgers, the State University of New Jersey, New Brunswick, NJ 08854, USA

about mental representations. In the human sciences, realism about theoretical (and especially mental) entities had fallen out of fashion in the middle third of the twentieth century. It seems to me that there were two things that made the difference in bringing cognitivism, back into psychological science. One was the work that began with Hilbert and was developed by Turing and Church and Markov and others who formulated the abstract notions of mechanism and of what we now call “information processing.” This is the lineage that led to Cybernetics and later to Artificial Intelligence, though a very large proportion of the field would now probably dissociate itself with that “logicist” part of the family tree, just as earlier Logicians like Frege dissociated themselves with psychological pursuits. The other development that brought mentalism back was the discovery that it was possible to treat some aspects of the human capacity for language in a way that made it at least appear to be compatible with mechanism. These developments encouraged many people to hope that one day we might have an explanatory theory of some of the mechanisms of linguistic competence, not just a taxonomic description of a corpus of linguistic utterances. The specific results achieved in transformational grammar, coupled with the generative or procedural aspect of the theoretical mechanisms (which, after all, wore the formal garb of Post Production systems and of Markov Algorithms) gave us hope that we were on the track of a theory of language understanding and language production.

Well, we were wrong about a lot of things, and especially about how a theory of grammar might be incorporated into a theory of comprehension/production (recall, for example, the decisive failure of the “derivational theory of complexity”). Many of the early ideas of psycholinguistics were later abandoned. What remained, however, was the basic belief that both rules, which included “rules of grammar”, and formal structures (of sentences) would play a central role in the theory of not only the language capacity, but also of cognition more generally. Moreover, ever since those developments in the late 50s and early 60s, talk about rules and representations no longer meant we were describing a corpus of behavior; rather when we spoke of rules we were referring to an internal property of some system or mind. We now routinely spoke of rules and the structures that they generate as being “internally represented”.

What was meant by the phrase “internally represented,” however, was far from clear – even to those of us who spoke that way. And it does not get any clearer if one adopts Chomsky’s way of putting it when, for example, he says that a theory of the speaker/hearer “involves rules”, or that the theory postulates a certain rule R “as a constituent element of [the speaker/hearer’s] initial state” or “attributes to . . . [the speaker/hearer] a mental structure . . . that includes the rule R and explains his behavior in terms of this attribution” (Chomsky, 1986, p. 243); or when he says that a speaker is “equipped with a grammar” or “internalizes a system of rules”. Yet, despite the uncertainties, none of us doubted that what was at stake in all such claims was nothing less than an empirical hypothesis about *how things really are inside the head of a human cognizer*. We knew that we were not speaking metaphorically nor were we in some abstract way describing the form of the data.

The way the story has gone within the study of language, including psycholinguistic studies of human performance, is now familiar to cognitive scientist, at least in broad outline if not in detail. But there is another area of cognitive science, quite different from the study of language, which has also made considerable progress: That is the area of visual perception. Under the important influence of David Marr (Marr, 1982), who saw the parallels between his enterprise and Chomsky's; visual perception, like language, was seen as being essentially modular (Pylyshyn, 1999), as amenable to the sort of competence-performance distinction that made progress in linguistics possible; and as fundamentally concerned with questions of representation. There has probably been more progress in the study of visual perception – and more interaction between the evidence of psychophysics, phenomenology and neuroscience – than in any other area of cognitive science. At the same time, there has been nearly as much misunderstanding and ideological dispute in the study of visual perception as there has been over the years in the study of language. In what follows I will discuss one recent line of work in which I have been involved that concerns the nature of the representations underlying visual perception (including one major shortcoming of the received view).

In addition to the broad methodological point that we need to distinguish between competence and performance, which informs both linguistics and vision science, the two fields share other properties, both methodological and substantive.

2 Some Parallels Between the Study of Vision and Language

The study of language and linguistic processes (learning, parsing, understanding and generating) developed in parallel with a clarification of basic goals of the science of linguistics. The study of visual perception has evolved in a similar way, as we developed a clearer view of its goals and to such questions as the following.

- (1) Is vision a distinct process or is it continuous with cognition? If the former, then how can we draw the boundary between vision and cognition?
- (2) Are the sources of evidence used in the study of vision special in any way? Do they, for example, include the equivalent of the sort of “judgments” (of grammaticality and ambiguity) used routinely by linguistics?
- (3) What function is computed by the visual system? Can we characterize the inputs and the outputs of vision – i.e., the representations that vision computes – in a perspicuous way, and in a way that shows its connection with general cognition?
- (4) What form of representation is computed by visual processes? Is the form of representation similar to the form of representation computed for language (e.g., Logical Form) or must it be different in fundamental ways?

In what follows I will focus primarily on the last item (4). Before I do that, however, I would like to point out some considerable similarities between vision and language processing as well as similarities of methodology faced by vision science and linguistics. On the face of it there are many similarities between language and vision. They are both productive so there is no limit on how many patterns can be

generated or recognized and in both cases similarities among patterns require appeal to the structure of the stimuli (i.e., both achieve their paradigmatic structure – the similarities and differences among distinct stimuli – by virtue of differences in the syntagmatic or syntactic structure among elements within each stimulus). Another way to put this is that in both vision and language there is syntactic structure which must be expressed by structure-dependent rules. Recognition of the type of each linguistic stimulus proceeds by the reconstruction of its structure through a process called parsing. In vision recognition also proceeds by a form of parsing (as developed, for example, in the recognition-by-components theory, Biederman, 1987). Also both language and vision reveal a substantial amount of innate structure and what rule-learning there is has to deal with the poverty of the stimulus – the fact that a finite set of samples of patterns is logically insufficient for inferring the structural rules. Moreover, determining the structure of individual patterns (parsing) must deal with missing parts: stimuli in both language and vision contain unexpressed parts that are filled-in by the observer: language structures contain gaps, deletions and traces that are not expressed in the physical signal, and vision routinely deals with partially-occluded patterns which are completed and filled in by the visual system (by a process called amodal completion, as illustrated by the many Kanizsa figures, see e.g., Kanizsa, 1979). The filling-in in both cases is done by modular processes, as opposed to being inferred from general knowledge. These general similarities suggest that the processes in both cases may be similar even though they are independent of one another.

Now consider the 4 questions set out above. The first question (#1 above), whether language and vision are distinct modules (or, as Chomsky puts it, different organs) has, I believe, been answered in the affirmative in both domains (I have argued the case for a visual module in Pylyshyn, 1999). Although there remain border skirmishes, as there always are at borders, it seems clear that vision and language both involve distinct functions and even distinct areas of the brain. The debate ultimately turns on the question of where the boundary is and that awaits the development of better and more general theories because ultimately it is the theory that tells you how to deal with the gray areas.

The same might be said of the second question (2). Linguistics has always used intuitions of native speakers regarding such phenomena as grammaticality, ambiguity and paraphrase. But these were subject to considerable argument in early years of generative grammar because one can't just ask someone whether a sentence is grammatical or ambiguous or whether two sentences mean the same thing. The very notions of grammaticality, ambiguity and sameness of meaning are *theory laden* (to use a term from Hanson, 1958). The sentence that Chomsky used in his earliest writing to illustrate the difference between grammaticality and acceptability (“Colorless green ideas sleep furiously”) was the subject of criticism and many people produced interpretations of the sentence to show it was meaningful. Intuitions of grammaticality are always problematic.¹ Yet in recent times the use of intuition in linguistics has not disappeared – it continues to play a central role in linguistic theory-building. But now it is used to answer well-posed questions derived from the theories. There is a similar problem in vision science where the

appeal to “how things look” or to the contents of conscious experience is similarly problematic. I need not list the many ways that conscious experience is misleading nor the well-known cases where vision is unaccompanied by any conscious experience at all. In fact when people report on what they experienced, or “what things look like” their reports may be guided by their own folk theories and expectations. As I have recently claimed (Pylyshyn, 2007, chap. 4) although we cannot stop using “what things look like” as a source of evidence, we need to use this kind of evidence in conjunction with evolving theories, just the way linguistic intuitions have been tamed by theories.

Question (3) is more problematic. In the case of language the input is either an acoustical stream or a string of linguistic units or formatives: phonemes or morphemes or lexical items, depending on whether the theory is to accommodate phonology, morphology or only syntax. It is widely held that these are independent levels of description that can be addressed separately. In the case of vision one might think of the input as consisting of an image, such as found on the retina.² But a case can be made that vision is an active process so the input might be better described in terms of what Gibson called the *ambient optical array* through which the organism moves and explores. There is also the question of the output of vision (assuming that vision really is an independent module). It has generally been assumed that the output of vision is much like the output of the language analyzer – logical forms. In any case few people think of the output of vision as anything but a symbolic description since without that vision would not inform the organism and lead to belief fixation (the exception being people who have advocated a theory of mental imagery that claims it uses the mechanisms of vision because in that case vision and mental images both generate displays in the brain, as opposed to logical forms).

The question can be raised of whether Logical Form, such as discussed in language, or some other essentially *descriptive* form of representation is adequate for representing visual percepts. The answer I am offering is that it is not. But I am *not* about to suggest that visual percepts should be thought of as *pictorial* or *analogue* or any other sort of ill-understood formats that many writers have proposed (Kosslyn, 1994). I find such proposals to be either hopelessly underspecified and metaphorical or else clearly false, although this is not the place to say why (Pylyshyn, 2002, 2007). What I claim is that the representations underlying visual percepts are mostly symbolic conceptual descriptions of roughly the classical sort. But I will also argue that notwithstanding the need for a logical form to allow perception to inform thoughts, this form of representation is incomplete in at least one critical respect – it lacks resources for picking out and referring to particular unique token individuals in the world. This logical form of representation lacks the special power that a *demonstrative* has of picking out a unique individual qua individual, as opposed to selecting whatever satisfies a certain description (which may sometimes be a unique individual, though even then it is selected as something that satisfies a certain description).

As linguistic terms demonstratives play a very important role in communication because they refer to token individuals. It is there that they come essentially into contact with perception; demonstratives pick out individual tokens in the perceptual

field, both in communication and in thought. They are, as Perry and Kaplan have argued, indispensable in language and thought (Almog, Perry, & Wettstein, 1989; Perry, 1979). One question that is controversial among philosophers of language is whether there are *bare* demonstratives or only complex demonstratives. A bare demonstrative refers to an individual without at the same time referring to it as something that falls under some conceptual category or other (as when we think “this”) whereas a complex demonstrative works like a descriptive noun phrase to pick out an individual that has the properties mentioned or implied by the referring expression (as in “this brown dog”). Ernie Lepore has been one of the defenders of the position that there are bare demonstratives, and moreover that the interpretation of complex demonstratives relies on the prior identification made by a bare demonstrative. Thus the reference fixing by the bare demonstrative in a complex demonstrative phrase has a prior status (Lepore & Ludwig, 2000). This is exactly the position that I have taken with respect to visual demonstratives. Since one of the functions that demonstrative reference plays (either in spoken language or the language of thought) is that of grounding conceptual representations in perception, then at least some of the things that perception picks out must be picked out without regard to the conceptual category it falls under. This is precisely the role that demonstratives play in language and I have argued that there must be a mental tokens in Language of Thought that play a similar role. We called these tokens FINST indexes, for historical reasons (it once stood for FINgers of INSTantiation).

2.1 Augmenting the Language of Thought to Include Demonstratives

I have defended the appropriateness of what I call here the classical symbolic view of visual representation on a number of different grounds (Pylyshyn, 2003). For example I have cited such properties as the abstractness and variability in definiteness of our visual representations (the way sentences can be abstract and variable in the sorts of details they encode) and the necessity that the system of representations meet the usual requirements of productivity and systematicity that Fodor and I discussed in connection with our critique of connectionist proposals (Fodor & Pylyshyn, 1988). I believe that compositional symbolic representations are the only form of representation that even come close to having the sort of requisite expressive power for visual percepts, even though they remain incomplete in a number of ways, such as their inability to conveniently encode magnitudes and the inability to individuate and reference tokens of visual objects. It is the latter shortcoming that I will discuss in this essay. A more extensive argument, with empirical evidence to support the detailed assumptions, is presented in (Pylyshyn, 2001a, 2003, 2007).

Theories of visual perception attempt to give an account of how a proximal stimulus (presumably a pattern impinging on the retina) can lead to a rich representation of a distal three-dimensional world and thence to either the recognition of known objects or to the coordination of actions with visual information. Such theories

typically provide an effective (i.e., computable) mapping from a 2D pattern to a representation of a 3D scene, usually in the form of a symbol structure. But such a mapping, though undoubtedly one essential purpose of a theory of vision, leaves at least one serious problem. The problem is that of connecting visual representations with the world in a certain critical way. This problem occurs for a number of reasons, but for our purposes I will emphasize just one such reason: the mapping from the world to our visual representation is not arrived at in one step, but incrementally. We know this both from empirical observations (e.g., percepts are generally built up by scanning attention and/or one's gaze) and also from theoretical analysis — e.g., Ullman (1984) has provided good arguments for believing that some relational properties, such as the property of being inside or on the same contour, have to be encoded serially by scanning a display. But then one problem arises immediately: If the representation is built up incrementally, we need to know that a certain part of our current representation refers to a particular individual object in the world. The reason is quite simple. As we elaborate the representation by uncovering new properties of a scene that we have partially encoded we need to know where (i.e., to which part of the representation) to attach the new information. In other words we need to know when a certain token in the existing representation should be taken as corresponding to the same (real, physical, individual) object as a particular token in the new representation, so that we can append newly noticed properties to the representation of the appropriate individual objects.

A possible way in which a purely descriptive representation could pick out individuals is by using definite descriptions. It could, for example, assert things like “the object x that has property P ” where P uniquely picks out a particular object x . In that case, in order to add new information, such as that this particular object also has property Q one would add the new predicate Q and also introduce an identity assertion, thus asserting something like $P(x) \wedge Q(y) \wedge x = y$ (and, by the way, adding this new compound descriptor to memory so that the same object might be relocated in this way when a further new property of that object is later noticed).³ But this is almost certainly not how the visual system adds information. This way of adding information would require adding a new predicate Q to the representation of an object that is *picked out by a certain descriptor*. To do that would require first recalling the description under which x was last encoded, and then conjoining to it the new descriptor and identity statement. Each new description added would require retrieving the description under which the object in question was last encoded.

The alternative to this unwieldy method is to allow the descriptive apparatus to make use of singular terms such as names or demonstratives. If we do that, then adding new information would amount to adding the predicate $Q(a)$ to the representation of a particular object a , and so on for each newly noticed property of a . Empirical evidence that we will review below suggests that the visual system's Q -detector recognizes instances of the property Q as a *property of a particular visible object*, such as object a , this is the most natural way to view the introduction of new visual properties by the sensorium. In order to introduce new properties in that way, however, there would have to be a non-descriptive way of picking out

a, such as a singular term or a name or a demonstrative. This is, in effect, what labeling objects in a diagram does through external means and what demonstrative terms like “this” or “that” do in natural language.⁴ This alternative is *prima facie* the more plausible one since it is surely the case that when we detect a new property we detect it as applying to *that* object, rather than as applying to some object in virtue of its being the object with a certain (recalled) property.⁵ Such intuitions, however, are notoriously unreliable so later in this chapter I will examine empirical evidence which suggests that this view is indeed more likely to be the correct one. For example, I will describe studies involving multiple-object tracking that make it very unlikely that objects are tracked by regularly updating a description that uniquely picks out the objects. In these studies the only unique descriptor available is location, and under certain plausible assumptions the evidence shows that it is very unlikely that the coordinates of the points being tracked are being regularly updated so that tracking is based on maintaining identity by updating descriptions.

There are a number of other reasons why a visual representation needs to be able to pick out individuals the way demonstratives do (i.e., independent of their properties or locations). For example, among the properties that are extracted (and presumably encoded in some way) by the visual system are a variety of relational predicates, such as **Collinear**(X1, X2, . . . Xn) or **Inside**(X1, C1) or **Part-of**(F1, F2), and so on. But these predicates apply over distinct individual objects in the scene independent of what properties these individuals have. So in order to recognize a relational property involving several objects we need to specify which objects are involved. For example, we cannot recognize the **Collinear** relation without picking out which objects are recognized as collinear. If there are many objects in a scene only some of them may be collinear so we must associate the relation with the objects in question. This is quite general since properties are predicated of things, and relational properties (like the property of being “collinear”) are predicated of several things. So there must be a way, independent of the process of deciding which property obtains, of specifying which objects (in our current question-begging sense) have that property. Ullman, as well as a large number of other investigators (Ballard, Hayhoe, Pook, & Rao, 1997; Watson & Humphreys, 1997; Yantis & Jones, 1991) talk of the objects in question as being “tagged” (indeed, “tagging” is one of the basic operations in Ullman’s theory of visual routines). The notion of a tag is an intuitive one since it suggests a way of *marking objects* for reference purposes. But the operation of tagging only makes sense if there is some thing on which a tag can literally be placed. It does no good to tag an internal representation (unless one assumes that it is an exact copy of the world) since the relation we wish to encode holds in the world and may not hold in the representation. But how do we tag parts of the world? What we need is what labels gave us in the previous example: A way to name or refer to individual parts of a scene *independent of their properties or their locations*.

What this means is that the representation of a visual scene must contain something more than descriptive or pictorial information in order to allow re-identification of particular individual visual elements. It must provide what natural language provides when it uses names (or labels) that uniquely pick out particular individuals, or when it embraces demonstrative terms like “this” or “that”. Such terms are used

to indicate particular individuals. This assumes that we have a way to *individuate*⁶ and keep track of particular individuals in a scene even when the individuals change their properties, including their locations. Thus what we need are two functions that are central to our concern: (a) we need to be able to pick out or individuate distinct individuals (following current practice, we will call these individuals *objects*) and (b) we need to be able to refer to these objects as though they had names or labels. Both these purposes are served by a primitive visual mechanism that I call a *visual index*. So what remains is for me to provide an empirical basis for the claim that the visual system embodies a primitive mechanism of the sort I call a *visual index* or a FINST. I begin with a description of the first of the two functions it provides, that of *individuating primitive visible objects*.

2.2 Primitive Visual Objects

Let me first provide a sketch of how the notion of an object has come into general use in the study of vision and visual attention. I will first describe a number of experiments that suggest that the detection of certain properties, such as color or shape or location, are perceptually separate from the detection of the individuals that bear them, and that the detection of objects likely precedes the detection of their properties. Then I will describe some experiments that further show that what the visual system detects when it is said to detect objects is not a proximal feature-cluster, but something that persists despite certain sorts of changes in its properties, including its location. By then we will see that the application of the term *object*, while still insufficient to bear the load of what is required of a real *individual*, as philosophers understand this term, begins to be much more interesting. In fact it offers a construct that I will call a *primitive visible object* that will be the building block for a story of how certain thoughts can be grounded on basic perceptual processes — i.e., how we can think about something for which we have no concept.

2.2.1 Evidence of Independent Recognition of Objects and Their Properties in Early Vision

Interest in what is now referred to as object-based attention may have begun with the observation that under certain conditions there appears to be a dissociation between the perception of certain properties and the perception of which objects have those properties. In fact it seems as though attention is required in order to bind properties to their bearers. For example, Anne Treisman and her colleagues showed that when properties of items not under direct attentional scrutiny were reported from a visual display there were frequent errors in which properties were assigned to the wrong items, resulting in what are called “illusory conjunctions”. For example, (Treisman & Gelade, 1980) found that if attention was distracted by a subsidiary visual task (such as naming digits at the center of the display), subjects frequently reported seeing the correct shape and color of items but in the wrong combinations resulting in erroneous *conjunctions* of color and shape (e.g., they reported that the display

contained a red X and a green O when in fact it had contained a green X and a red O). The illusory conjunctions appear with a large variety of properties of objects (Treisman & Gelade, 1980; Treisman & Sato, 1988). For example, illusory conjunctions occur for shape properties so that a display with right oblique lines, L-shaped and S-shaped figures led to the misperception of triangles and dollar signs. There is also evidence that certain object properties can be detected while their locations are either misidentified or unknown.⁷ Thus you might see that a display contains the letter X but fail to detect *where* it was located, or “see” it to be at the wrong location (Chastain, 1995; Treisman, 1986). There has also been considerable interest in recent years in the so-called “two visual systems” view (Ungerleider & Mishkin, 1982) which claimed that there are two streams of visual processing in the brain: A dorsal stream that encodes *where* a thing is and a ventral stream that encodes *what* it is (its identity).⁸ These and related studies (including demonstrations that people can attend to large random shape embedded within other shapes which they must ignore – Rock & Gutman, 1981) suggested that attention is allocated to what are called *objects* (or individuals) rather than to particular *places, regions, or properties*. There is even evidence from the clinical syndrome known as unilateral neglect that what is neglected must be described in relation to perceptual objects rather than locations in space (Tipper & Behrmann, 1996).

2.2.2 Evidence that Extracting Several Pieces of Information from a Display is Easier if they are Part of One Object

The notion that objects are detected and then visual properties are bound to them at a very early stage in visual perception has also received support from studies showing that it is faster to find (and identify) several features or a properties if they are associated with the same object (and also features that are part of different objects interfere less in a search task). For example, (Duncan, 1984) and later (Baylis & Driver, 1993) showed that access to relational properties of two features (such as “larger than”) is faster when the features in question belong to the same perceptual object than when they are parts of different objects which nonetheless are objectively in the same relative relation (e.g., the same distance apart). These studies all point to the idea that objects are selected first and *then* properties of these objects may be encoded and available for judgments.

2.2.3 Evidence for Access to Multiple Objects

In order to detect such relational properties as that a number of points are collinear or that a point is inside a closed contour the visual system must have a way to refer to the individuals over which these predicates are supposed to apply. In general, to evaluate $P(x,y)$ both x and y need to be bound to the individuals in question. Yet attention has generally been assumed to be unitary: you can devote attention to only one thing at a time (not one *place* at a time⁹). Since we can move attention from object to object there must be some way to specify which object to move it to next. We must have some pre-attentional access or variable binding mechanism.

So the mechanism for binding mental variables to objects must be more primitive than and precede the allocation of focal attention. Visual Index Theory (Pylyshyn, 2001b) claims that prior to the allocation of focal unitary attention visual indexes (or FINSTs) must be “grabbed” by portions of the visual landscape. The function of these indexes is to provide a way to access objects on demand, or to bind parts of the cognitive representation to objects. How many objects? Empirically we have found the number to be around 4 or 5 over a wide variety of experimental paradigms.

Several properties of the indexing process are illustrated by a series of studies we have performed involving selecting a subset of items in a visual search task. The search task we used was adapted from one originally introduced by (Treisman & Gelade, 1980). In a series of studies, Jacquie Burkell and I (Burkell & Pylyshyn, 1997), used the sudden-onset of new objects (which we called “late-onset placeholders”) to control search. The empirical question was whether the search would be confined to the subset defined by the late-onset objects – those that we assumed had been indexed. The answer was unequivocal: Only indexed objects constituted the search set. Moreover, it made no difference how far apart the indexed objects were, showing that they did not have to be searched out before being matched against the search criteria. (For more details on these and a number of other studies see, Pylyshyn, 2003; Pylyshyn et al., 1994.)

2.3 Individuating and Tracking Primitive Visual Objects: Multiple Object Tracking Studies

Perhaps the clearest way to see what is being claimed when I say there is a primitive mechanism in early vision that picks out and maintains the identity of visible objects is to consider a set of experiments, carried out in my laboratory, to which the ideas of visual individuation and identity maintenance were applied. The task is called the *Multiple Object Tracking (MOT) Task*.

In a typical experiment, subjects are shown a screen containing anywhere from 12 to 24 simple identical objects (points, spheres, plus signs, figure-eight shapes) which move across the entire visual field in unpredictable ways without colliding. A subset of these objects is briefly rendered distinct (usually by flashing them on and off a few times). The subject’s task is to keep track of this subset of objects (called “targets”). At some later time in the experiment (say 10 seconds into the tracking trial) one of the objects is again flashed on and off. The subject must then indicate whether or not the flashed (probe) figure was one of the targets. A large number of experiments, beginning with studies by (Pylyshyn & Storm, 1988), have shown clearly that subjects can indeed track up to 5 independently moving identical objects. Moreover, we were able to argue that the motion and dispersion parameters of the original Pylyshyn and Storm experiment were such that tracking could not have been accomplished using a serial strategy in which attention is scanned to each figure in turn, storing its location, and returning to find the figure closest to that location on the next iteration, and so on. Based on some weak assumptions about how fast focal attention might be scanned and based on actual data on how fast the

objects actually moved and how close together they had been in this study, we were able to conclude that such a serial tracking process would very frequently end up switching to the wrong objects in the course of its tracking. This means that the moving objects could not have been tracked using a unique stored description of each figure, inasmuch as the only possible descriptor that was unique to each figure at any particular instant in time was its location. If we are correct in arguing from the nature of the tracking parameters that stored locations cannot be used as the basis for tracking, then all that is left is the figure's *numerical identity* or its persistence as the same individual. This is exactly what I claim is going on – tracking by maintenance of a primitive perceptual individuality.

Recently a large number of additional studies in our laboratory have replicated these multiple object tracking results, confirming that subjects can successfully track several independently moving objects.¹⁰ Moreover, performance in detecting changes to elements located inside the convex hull outline of the set of targets was no better than performance on elements outside this region, contrary to what would be expected if the area of attention were simply widened or shaped to conform to an appropriate outline (Pylyshyn, et al., 1994). Using a different tracking methodology, Intriligator and Cavanagh (2001) also failed to find any evidence of a “spread of attention” to regions between targets. It appears, then, that items can be tracked despite the lack of distinctive properties (and, indeed when their properties are changing) and despite constantly changing locations and unpredictable motions. Taken together these studies implicate a notion of primitive visible object as a category induced by the early visual system, preceding the recognition of properties and preceding the evaluation of any visual predicate.

The multiple object tracking task exemplifies what I mean by “tracking” and by “maintaining the identity” of objects. It also operationalizes the notion of “primitive visible object” – a primitive visible object is whatever attracts a FINST index and allows multiple-object tracking. Note that this is a highly mind-dependent definition of objecthood. Objecthood and object-identity are defined in terms of a causal perceptual mechanism. A certain sequence of object-locations will count as the movement a single object if the early (pre-attentive) visual system groups it this way – i.e., if it is so perceived – whether or not we can find a physical property that is invariant over this sequence and whether or not there exists a psychologically-plausible description that covers this sequence. The visual system may also count as one individual object certain kinds of disappearances and reappearances of visual objects. For example, Scholl and Pylyshyn (1999) have shown that if the objects being tracked in the MOT paradigm disappear and reappear in certain ways they are tracked as though they had a continuous existence. If they disappear and reappear by deletion and accretion along a fixed contour, the way they would have if they were moving behind an occluding surface (even if the edges of the occluder are not invisible), then they are tracked as though they were continuously moving objects. Performance in the MOT task does not deteriorate if targets disappear in this fashion although it suffers dramatically if targets suddenly go out of existence and reappear, or if they slowly shrink away and then reappear by slowly growing again at exactly the same place as they had accreted in the occlusion condition.

2.4 A Theory of Visual Indexing and Binding: The FINST Mechanism

The basic motivation for postulating Visual Indexes is that, as we saw at the beginning of this essay, there are a number of reasons for thinking that individual objects in the field of view must first be *picked out* from the rest of the visual field and the identity of these objects *qua individuals* must be maintained or tracked despite changes in the individual's properties including its location in the visual field. Our proposal claims that this is done *primitively* without identifying the object through a unique descriptor. The object in question must be segregated from the background or picked out as an individual (the Gestalt notion of making a figure-ground distinction is closely related to this sort of "picking out"). Until some piece of the visual field is segregated and picked out, no visual operation can be applied to it since it does not exist as something distinct from the entire field.

In its usual sense (at least in philosophy), picking out an individual requires having criteria of individuation – i.e., requires having a sortal concept. How can we track something without re-recognizing it as the same thing at distinct periods of time, and how can we do that unless we have a concept or a description of it? My claim is that just as the separation of figure from ground (the "picking out") is a primitive function of the architecture of the visual system, so also is this special sort of preattentive tracking. What I am proposing is not a full-blooded sense of identity-maintenance, but a sense that is relativized to the basic character of the early visual system. The visual system cannot in general re-recognize objects as being the same without some descriptive apparatus, but it can track in a more primitive sense, providing certain conditions are met (several of these conditions were mentioned earlier in discussing the Yantis and the Pylyshyn and Scholl results cited above).

What this means is that our theory is concerned with a sense of *picking out* and *tracking* that are not based on top-down *conceptual* descriptions, but are given pre-conceptually by the early visual system, and in particular by the FINST indexing mechanism. Moreover, the visual system treats the object so picked-out as distinct from other individuals, independent of what properties this object might have. If two different objects are individuated in this way they remain distinct as far as the visual system is concerned. Moreover, they remain distinct despite certain changes in their properties, particularly changes in their location. Yet the visual system need not know (i.e., need not have detected or encoded) any of their properties in order to implicitly treat them as though they were distinct and enduring visual tokens. Of course there doubtless are properties, such as being in different locations or moving in different ways or flashing on and off that allow indexes to be assigned to these primitive objects in the first place. But none of these properties define the objects – they are not essential properties. What constitutes the relevant property is that it attracts an index and the set of properties that do that need not form a natural kind. My claim is that to index *x*, in *this primitive sensory sense*, there need not be any concept, description or sortal that picks out *x*'s by type.¹¹

The basic idea of the FINST indexing and binding mechanism is that a causal chain leads from certain kinds of visible events, via primitive mechanisms of the early visual system, to certain conceptual structures (which we may think of as

symbol structures in Long Term Memory). This provides a mechanism of reference between a visual representation and what we have called primitive visible objects in the world. The important thing here is that the inward effects are purely causal and are instantiated by the non-conceptual apparatus of what I have called early vision (Pylyshyn, 1999). This apparatus guarantees that under certain conditions the link will maintain a certain continuity, thus resulting in its counting as the *same link*. It is tempting to say that what makes it continuous is that it keeps pointing to the same thing, but according to our view this is circular since the only thing that makes it the *same thing* is the very fact that the it the index references it. There is no other sense of “sameness” so that “primitive visible object” as we have defined it is thoroughly mind dependent.

By virtue of this causal connection, the conceptual system can *refer to* any of a small number of primitive visible objects. It can, for example, interrogate them to determine some of their properties, it can evaluate visual predicates (such as *Collinear*) over them, it can move focal attention to them, and so on. The function that I am describing is extremely simple and only seems complicated because ordinary language fails to respect certain distinctions (such as the distinction between individuating and recognizing, indexing and knowing where something is, and so on). Elsewhere (Pylyshyn, 2003) I provide an extremely simple network, based on the Koch and Ullman (1984) winner-take-all neural net, which implements such a function.

3 What does all this have to do with Connecting Vision and the World?

What we have described is a mechanism for picking out, tracking and providing cognitive access to what we call an object (or, more precisely, a *primitive visible object*). The notion of an *object* is ubiquitous in cognitive science, not only in vision but much more widely. I might also note that it has been a central focus in developmental psychology where people like Susan Carey and Fei Xu have studied “a child’s concept of object” (Xu, 1997), and in clinical neuroscience, where it has been argued that deficits such as unilateral neglect must be understood as a deficit of object-based attention rather than space-based attention. Space does not permit me to go into any of these fields although I am engaged in a larger project where I do examine the connections among these uses of the term “object”. But I would like to draw your attention to the fact that giving objects the sort of central role in vision that I have described suggests a rather different ontology. Just as it is natural to think that we apprehend properties such as color and shape as *properties of objects*, so it is also natural to think that we apprehend objects as a kind of property that particular *places* have. In other words we usually think of the matrix of space-time as being primary and of objects as being occupants of places and times. Everyone from Kant to modern cognitive scientists takes this for granted – that’s (in part) why it is so natural to think of mental images as having to be embedded in real space in the brain. Yet the findings I have described in the study of visual attention (as

well as other areas of psychological research which I cannot describe here, but see, Pylyshyn, 2003) suggests an alternative and rather intriguing possibility. It is the notion that *primitive visible object* is the primary and more primitive category of early (preattentive) perception, so that we perceive objecthood first and determine location the way we might determine color or shape – as a property associated with objects. If this is true then it raises some interesting possibilities concerning the nature of the mechanisms of early vision. In particular it suggests what we argued is independently needed – a mechanism for directly referring to objects in a way that does not rely on having a unique description under which that object falls. This is the function (of “demonstrating”) served by the hypothesized visual index mechanism.

Notice that when I am careful I hedge my use of the term *object* in making this claim, as I must because what I have been describing is not the notion of an object in the usual sense of a physical object or individual. *Object* or *individual* are sortal concepts whose individuation depends on assuming certain conceptual categories. But our notion does not assume any concepts. The individuals that are picked out by the visual system and tracked primitively are something less than full blooded individuals. Yet because they are what our visual system gives us through a brute causal mechanism – because that is its nature – it serves as the basis for all real individuation. As philosophers like (Wiggins, 1979) and (Hirsch, 1982) have argued, you cannot individuate objects in the full blooded sense without a conceptual apparatus – without sortal concepts. But similarly you cannot individuate them with *only* a conceptual apparatus. Sooner or later concepts must be grounded in a primitive causal connection between thoughts and things. The project of grounding concepts in sense data has not fared well and has been abandoned in cognitive science. However the principle of grounding concepts in perception remains an essential operation if we are not to succumb to an infinite regress (or worse, to a behaviorist eliminative understanding of concepts). Visual indexes provide a putative grounding for basic objects and we should be grateful because without them (or at any rate something like them) we would be lost in thought without any grounding in causal connections with the real-world objects of our thoughts. With indexes we can think about things (I am sometimes tempted to call them *FINGs* since they are interdefined with *FINSTs*) without having any concepts of them: One might say that we can have *demonstrative thoughts*. And nobody ought to be surprised by this since we know that we can do this: I can think of this here thing without *any description* under which it falls. And, perhaps even more important, because I can do that I can reach for it.

If this analysis is correct – if people do select visual objects before they represent their properties – then treating demonstrative terms as consisting of bare demonstratives (plus additional properties based on the rest of the complex), rather than complex demonstratives that pick out objects-with-specified-properties – makes sense. It makes sense for all the reasons that (Lepore & Ludwig, 2000) have given together with the empirically-motivated grounds suggested here – namely, that attentive selection (or *FINST* selection) at its initial and most primitive nonconceptual stage picks out visual objects before it encodes their properties. The property

encoding places *conceptual* logical forms into the Object Files which were created empty after a new object came into view or was noticed.

Well I have probably waded deep enough into philosophy for the modest purposes of this essay. Needless to say there are some details to be worked out as this is a work-in-progress. But I hope I have at least made the point that there is a real problem to be solved in connecting visual representations to the world that is different in principle from the representations of sentences referred to as Logical Form. Whatever the eventual solution to the problem of visual representation turns out to be, it will have to respect a collection of facts some of which I have sketched here. Moreover any visual or attentional mechanism that might be hypothesized for this purpose will have far reaching implications, not only for theories of situated vision, but also for grounding the content of visual representations and perhaps for grounding perceptual concepts in general.

Notes

1. Consider the following two sentences. Which one (if any) is grammatical? "I am having trouble deciding between/among P" where P is some numerical predicate. The choice, according to some grammars taught in school, depends on whether P yields exactly two alternatives. But that is not decidable in general. Does that mean that grammar contains undecidable rules? Clearly not: What is shown is that a rule one believed to be a rule of grammar turns out not to be part of grammar at all, something that intuition is powerless to decide.
2. This obvious point took hundreds of years to appreciate. Only after Kepler's seminal analysis of how an image can be focused by a lens was the role of the retinal image appreciated (Lindberg, 1976).
3. Strictly speaking the definite description that uniquely picks out a certain object at a particular time is a quantified expression of the form: $\exists xP(x)$, where P is the unique property of the object in question. When an additional predicate Q that pertains to the same object is to be added, the unique descriptor is retrieved and the new stored expression added: $(\exists x\exists y\{P(x) \perp Q(y) \perp x=y\})$. If a further property R of the same object is detected at some later time, the last expression must be matched to the object at which R is discovered and its descriptor updated to the expression $\exists x\exists y\exists z\{P(x) \perp Q(y) \perp R(z) \perp x=y \perp y=z\}$. This continual updating of descriptors capable of uniquely picking out objects is clearly not a plausible mechanism for incrementally adding to a visual representation. It demands increasingly large storage and retrieval based on pattern matching.
4. Notice that the need for demonstratives remains even if the representation were picture-like instead symbolic, so long as it was not an exact and complete copy of the world but was built up incrementally. If the picture depicts some state of affairs in the world we still have the problem of deciding when two pictorial bits are supposed to depict the same object. We still need to decide when two picture-fragments are supposed to depict the same object (even though they may look different) and when they are supposed to depict different objects. This is the same problem we faced in the case of symbolic representations. We don't know whether the thing in the picture that is depicted as having the property P is the thing to which we must now add the depiction of the newly-noticed fact that it also has property Q. Without a solution to that puzzle we don't know to which part of the picture to add newly noticed properties.
5. There is another alternative for picking out objects that I will not discuss here because the evidence I will cite suggests that it is not the correct option for visual representations. This alternative that assumes the existence of demonstratives, as we have done, except the demonstratives in question are *place demonstratives* or *locatives*, such as "this place". Such an

apparatus would allow the unique picking out of objects based on their location and would overcome the problem with the pure descriptivist story that we have been describing. That alternative is compatible with the view presented here although, as we will argue, the idea that object individuation is mediated by location alone (or location alone) does not seem to be supported by the empirical data.

6. As with a number of terms used in the context of early vision (such as the term “object”), the notion of *individuating* has a narrower meaning here than in the more general context where it refers not only to separating a part of the visual world from the rest of the clutter (which is what we mean by individuate here), but also providing identity criteria for recognition instances of that individual. As is the case with *objecthood* and other such notions, we are here referring primarily to primitive cases – i.e. ones provided directly by mechanisms in the early vision system (in the sense of Pylyshyn, 2007) and not constructed from other perceptual functions.
7. This claim is contentious. There have been a number of studies (reviewed in Pashler, 1998) showing that in those cases where an object is correctly identified, its location generally can be correctly reported. However, what these studies actually show is that for objects whose shapes (or in some cases color) can be correctly reported, their location can usually also be reported. From our perspective this only shows that there is a precedence ranking among the various properties of an object that are recorded and reported and that rough location may be higher on the ranking than other properties. What the experiments do not show (contrary to some claims) is that *in order to detect the presence of an object one must first detect its location*. The studies described below (dealing with multiple Indexing) suggest ways to decide whether an object has been detected in the relevant sense (i.e., individuated and indexed, though not necessarily recognized). The theoretical position to be developed here entails that one can *index* an object without encoding its location. There are, so far as I know, no data one way or another regarding this prediction.
8. More recent studies have shown that the what-where dichotomy is not quite the right way to distinguish the two visual systems. Rather it appears that one of the systems (the ventral system) specializes in recognition while the other (the dorsal system) specializes in visual-motor control (Goodale & Milner, 2004).
9. Location-based attention is not ruled out by these studies. It still remains possible that a “spotlight of attention” can be scanned across a display in search of objects of interest. However, these studies do show that at least *some* forms of attention are directed to whole objects irrespective of their location in space.)
10. Some published research includes (Allen, McGeorge, Pearson, & Milne, 2004; Alvarez, Arsenio, Horowitz, & Wolfe, 2005; Alvarez & Cavanagh, 2005; Alvarez & Scholl, 2005; Bahrami, 2003; Blaser, Pylyshyn, & Domini, 1999; Blaser, Pylyshyn, & Holcombe, 2000; Cavanagh, 1992; Cavanagh & Alvarez, 2005; Cavanagh, Labianca, & Thornton, 2001; Culham et al., 1998; Fougny & Marois, 2006; Intriligator & Cavanagh, 2001; Jovicich et al., 2001; Keane & Pylyshyn, 2006; Ogawa & Yagi, 2002; O’Hearn, Landau, & Hoffman, 2005; Pylyshyn, 1989, 1994, 2004; Pylyshyn et al., 1994; Pylyshyn & Storm, 1988; Saiki, 2003; Scholl & Pylyshyn, 1999; Scholl, Pylyshyn, & Feldman, 2001; Scholl, Pylyshyn, & Franconeri, 1999; Sears & Pylyshyn, 2000; Slemmer & Johnson, 2002; Suganuma & Yokosawa, 2002; Trick, Perl, & Sethi, 2005; vanMarle & Scholl, 2003; Viswanathan & Mingolla, 1998, 2002; Yantis, 1992).
11. I am claiming that there is a mechanism in early (pre-conceptual) vision that latches onto certain entities for purely causal reasons, not because those entities meet conditions provided by a cognitive predicate – i.e., not because they constitute instances of a certain concept. In other words if $P(x)$ is a primitive visual predicate of x then the x is assumed to have been independently and causally bound to what I have called a primitive visible object. Although this sort of latching or seizing by primitive visible objects is essentially a bottom-up process, this is not to say that it could not in some cases be guided by intentional processes, such as perhaps scanning one’s attention until a latching event is located or an object meeting a certain description is found. For example, it is widely assumed (Posner, Snyder, & Davidson, 1980) that people can scan their attention along some path (by simply moving it continuously

through space like a spotlight beam) and thereby locate certain sorts of objects. A possible consequence of such scanning is that an index may get assigned to some primitive objects encountered along the way.

References

- Allen, R., McGeorge, P., Pearson, D., & Milne, A. B. (2004). Attention and expertise in multiple target tracking. *Applied Cognitive Psychology, 18*, 337–347.
- Almog, J., Perry, J., & Wettstein, H. (Eds.). (1989). *Themes from Kaplan*. New York: Oxford University Press.
- Alvarez, G. A., Arsenio, H. C., Horowitz, T. S., & Wolfe, J. M. (2005). Are multielement visual tracking and visual search mutually exclusive? *Journal of Experimental Psychology: Human Perception and Performance, 31*(4), 643–667.
- Alvarez, G. A., & Cavanagh, P. (2005). Independent attention resources for the left and right visual hemifields. *Psychological Science, 16*(8), 637–643.
- Alvarez, G. A., & Scholl, B. J. (2005). How does attention select and track spatially extended objects? New effects of attentional concentration and amplification. *Journal of Experimental Psychology: General, 134*(4), 461–476.
- Bahrami, B. (2003). Object property encoding and change blindness in multiple object tracking. *Visual Cognition, 10*(8), 949–963.
- Ballard, D. H., Hayhoe, M. M., Pook, P. K., & Rao, R. P. N. (1997). Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences, 20*(4), 723–767.
- Baylis, G. C., & Driver, J. (1993). Visual attention and objects: Evidence for hierarchical coding of location. *Journal of Experimental Psychology: Human Perception and Performance, 19*, 451–470.
- Biederman, I. (1987). Recognition-by-components: A theory of human image interpretation. *Psychological Review, 94*, 115–148.
- Blaser, E., Pylyshyn, Z. W., & Domini, F. (1999). Measuring attention during 3D multielement tracking. *Investigative Ophthalmology and Visual Science, 40*(4), 552 (Abstract).
- Blaser, E., Pylyshyn, Z. W., & Holcombe, A. O. (2000). Tracking an object through feature-space. *Nature, 408*(Nov 9), 196–199.
- Burkell, J., & Pylyshyn, Z. W. (1997). Searching through subsets: A test of the visual indexing hypothesis. *Spatial Vision, 11*(2), 225–258.
- Cavanagh, P. (1992). Attention-based motion perception. *Science, 257*, 1563–1565.
- Cavanagh, P., & Alvarez, G. A. (2005). Tracking multiple targets with multifocal attention. *Trends in Cognitive Sciences, 9*(7), 349–354.
- Cavanagh, P., Labianca, A. T., & Thornton, I. M. (2001). Attention-based visual routines: Sprites. *Cognition, 80*(1–2), 47–60.
- Chastain, G. (1995). Location coding with letters versus unfamiliar, familiar and labeled letter-like forms. *Canadian Journal of Experimental Psychology, 49*(1), 95–112.
- Chomsky, N. (1986). *Knowledge of Language: Its Nature, Origin, and Use*. New York: Praeger.
- Culham, J. C., Brandt, S. A., Cavanagh, P., Kanwisher, N. G., Dale, A. M., & Tootell, R. B. H. (1998). Cortical fMRI activation produced by attentive tracking of moving targets. *Journal of Neurophysiology, 80*(5), 2657–2670.
- Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General, 113*(4), 501–517.
- Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition, 28*, 3–71.
- Fougnie, D., & Marois, R. (2006). Distinct capacity limits for attention and working memory. Evidence from attentive tracking and visual working memory paradigms. *Psychological Science, 17*(6), 526–534.
- Goodale, M., & Milner, D. (2004). *Sight unseen*. New York: Oxford University Press.

- Hanson, N. R. (1958). *Patterns of discovery*. Cambridge: Cambridge University Press.
- Hirsch, E. (1982). *The concept of identity*. Oxford, UK: Oxford.
- Intriligator, J., & Cavanagh, P. (2001). The spatial resolution of attention. *Cognitive Psychology*, 4(3), 171–216.
- Jovicich, J., Peters, R., Koch, C., Braun, J., Chang, L., & Ernst, T. (2001). Brain areas specific for attentional load in a motion-tracking task. *Journal of Cognitive Neuroscience*, 13, 1048–1058.
- Kanizsa, G. (1979). *Organization in vision: Essays on Gestalt perception*. New York: Praeger.
- Keane, B. P., & Pylyshyn, Z. W. (2006). Is motion extrapolation employed in multiple object tracking? Tracking as a low-level, non-predictive function. *Cognitive Psychology*, 52(4), 346–368.
- Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: towards the underlying neural circuitry. *Human Neurobiology*, 4, 219–227.
- Kosslyn, S. M. (1994). *Image and Brain: The resolution of the imagery debate*. Cambridge, MA: MIT Press.
- Lepore, E., & Ludwig, K. (2000). The semantics and pragmatics of complex demonstratives. *Mind*, 109, 199–240.
- Lindberg, D. C. (1976). *Theories of vision from al-Kindi to Kepler*. Chicago: University of Chicago Press.
- Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco: W.H. Freeman.
- Ogawa, H., & Yagi, A. (2002). The effect of information of untracked objects on multiple object tracking. *Japanese Journal of Psychonomic Science*, 22(1), 49–50.
- O’Hearn, K., Landau, B., & Hoffman, J. E. (2005). Multiple Object tracking in people with Williams syndrome and in normally developing children. *Psychological Science*, 16(11), 905–912.
- Perry, J. (1979). The problem of the essential indexical. *Noûs*, 13, 3–21.
- Posner, M. I., Snyder, C., & Davidson, B. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General*, 109, 160–174.
- Pylyshyn, Z. W. (1989). The role of location indexes in spatial perception: A sketch of the FINST spatial-index model. *Cognition*, 32, 65–97.
- Pylyshyn, Z. W. (1994). Some primitive mechanisms of spatial attention. *Cognition*, 50, 363–384.
- Pylyshyn, Z. W. (1999). Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences*, 22(3), 341–423.
- Pylyshyn, Z. W. (2001a). Connecting vision and the world: Tracking the missing link. In J. Branquinho (Ed.), *The Foundations of Cognitive Science* (pp. 183–195). Oxford, UK: Clarendon Press.
- Pylyshyn, Z. W. (2001b). Visual indexes, preconceptual objects, and situated vision. *Cognition*, 80(1/2), 127–158.
- Pylyshyn, Z. W. (2002). Mental imagery: In search of a theory. *Behavioral and Brain Sciences*, 25(2), 157–237.
- Pylyshyn, Z. W. (2003). *Seeing and visualizing: It’s not what you think*. Cambridge, MA: MIT Press/Bradford Books.
- Pylyshyn, Z. W. (2004). Some puzzling findings in multiple object tracking (MOT): I. Tracking without keeping track of object identities. *Visual Cognition*, 11(7), 801–822.
- Pylyshyn, Z. W. (2007). *Things and places: How the mind connects with the world (Jean Nicod Lectures Series)*. Cambridge, MA: MIT Press.
- Pylyshyn, Z. W., Burkell, J., Fisher, B., Sears, C., Schmidt, W., & Trick, L. (1994). Multiple parallel access in visual attention. *Canadian Journal of Experimental Psychology*, 48(2), 260–283.
- Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, 3(3), 1–19.
- Rock, I., & Gutman, D. (1981). The effect of inattention on form perception. *Journal of Experimental Psychology: Human Perception and Performance*, 7, 275–285.
- Saiki, J. (2003). Feature binding in object-file representations of multiple moving items. *Journal of Vision*, 3(1), 6–21.

- Scholl, B. J., & Pylyshyn, Z. W. (1999). Tracking multiple items through occlusion: Clues to visual objecthood. *Cognitive Psychology*, *38*(2), 259–290.
- Scholl, B. J., Pylyshyn, Z. W., & Feldman, J. (2001). What is a visual object: Evidence from target-merging in multiple-object tracking. *Cognition*, *80*, 159–177.
- Scholl, B. J., Pylyshyn, Z. W., & Franconeri, S. L. (1999). When are featural and spatiotemporal properties encoded as a result of attentional allocation? *Investigative Ophthalmology & Visual Science*, *40*(4), 4195.
- Sears, C. R., & Pylyshyn, Z. W. (2000). Multiple object tracking and attentional processes. *Canadian Journal of Experimental Psychology*, *54*(1), 1–14.
- Slemmer, J. A., & Johnson, S. P. (2002). *Object tracking in ecologically valid occlusion events*. Paper presented at the Vision Sciences 2002, Sarasota, FL.
- Suganuma, M., & Yokosawa, K. (2002). *Is multiple object tracking affected by three-dimensional rigidity?* Paper presented at the Vision Sciences Society, Sarasota, FL.
- Tipper, S. P., & Behrmann, M. (1996). Object-centered not scene-based visual neglect. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(5), 1261–1278.
- Treisman, A. (1986). Properties, parts, and objects. In K. Boff, L. Kaufmann, & J. Thomas (Eds.), *Handbook of perception and human performance*. New York: Wiley.
- Treisman, A., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, *12*, 97–136.
- Treisman, A., & Sato, S. (1990). Conjunction search revisited. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(3), 459–478.
- Trick, L. M., Perl, T., & Sethi, N. (2005). Age-related differences in multiple-object tracking. *Journals of Gerontology: Series B: Psychological Sciences & Social Sciences*, *2*, 102.
- Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In J. Ingle, M. A. Goodale, & R. J. W. Mansfield (Eds.), *Analysis of visual behavior* (pp. 549–586). Cambridge, MA: MIT Press.
- vanMarle, K., & Scholl, B. J. (2003). Attentive tracking of objects versus substances. *Psychological Science*, *14*(4), 498–504.
- Viswanathan, L., & Mingolla, E. (1998). Attention in depth: Disparity and occlusion cues facilitate multi-element visual tracking (Abstract). *Investigative Ophthalmology and Visual Science*, *39*(4), 634.
- Viswanathan, L., & Mingolla, E. (2002). Dynamics of attention in depth: Evidence from multi-element tracking. *Perception*, *31*(12), 1415–1437.
- Watson, D. G., & Humphreys, G. W. (1997). Visual marking: prioritizing selection for new objects by top-down attentional inhibition of old objects. *Psychological Review*, *104*(1), 90–122.
- Wiggins, D. (1979). *Sameness and Substance*. London: UK: Blackwell.
- Xu, F. (1997). From Lot's wife to a pillar of salt: Evidence that physical object is a sortal concept. *Mind and language*, *12*, 365–392.
- Yantis, S., & Jones, E. (1991). Mechanisms of attentional selection: temporally modulated priority tags. *Perception and Psychophysics*, *50*(2), 166–178.
- Yantis, S. (1992). Multielement visual tracking: Attention and perceptual organization. *Cognitive Psychology*, *24*, 295–340.

Index

A

Abbot, B., 201, 202
Acquaviva, P., 196
Allen, R., 271
Almog, J., 260
Alternative semantics, 205
Alvarez, G.A., 271
Amended Assign and Combine Model (AACM), 83, 90
Anderson, D.R., 135
Aoun, J., 216
Argument
 embeddings, from, 151–152
 overdetermination, from, 151
Arsenio, H.C., 271
Artificial intelligence, 256
Asher, N., 170, 177, 178
Attitude sentences, properties of, 55
Audi, R., 64
Auxiliary verbs, 53
Availability Principle (AP), 131–132

B

Bach, K., 73, 80, 99, 102, 127
Bahrami, B., 271
Bain, A., 151, 155
Ballard, D.H., 262
Bamber, D., 135
Barwise, J., 25, 198, 201
Baylis, G.C., 264
Beaver, D.I., 202
Beghelli, F., 216
Begriffsschrift, 150, 153
Behrman, M., 264
Bentham, E., 154
Bever, T., 140
Bezuidenhout, A., 80, 237, 238, 241–243, 248, 251, 253
Biconditional disquotation principle, 169, 170

Biederman, I., 258
Binary constructions, recursion in, 229–232
Blair, D., 31–56
Blaser, E., 271
Boole, G., 154, 155
Borg, E., 100
Botne, R., 224
Botterell, A., 32, 37
Bradley, F., 158
Brandom, R., 68
Burge, T., 35
Büring, D., 39, 196, 204, 207, 208
Burkell, J., 265
Burnham, K., 135

C

Camp, E., 238
Cantor's Theorem, 169
Cappelen, H., 65, 66, 71, 79–82, 84, 85, 88–102, 129, 237, 239, 243, 244, 246–248, 251, 253
Cappy's metaphorical utterance, 241–242
Carston, R., 80, 84, 238, 242
Cavanagh, P., 266, 271
Chastain, G., 264
Chierchia, G., 195
Chomsky, N., 134, 137, 138, 140, 141, 144, 207, 255, 256, 258
Clapp, L., 79–103
Coady, C., 64
Cognitive science, 257
Cohen, S., 105, 119, 126
Complementizer phrase (CP), 44
 properties of, 55
 of tensed infinitives, 55
 types of clauses, 55
Conditional Excluded Middle (CEM) and Modalized Restrictive Account, 16

- principle of, 3, 4
 - and theory of conditionals, 9–12
 - Context dependence
 - uncontroversial kinds of, 112–114
 - Context-of-judgment-invariant, 87
 - Context Shifting Arguments (CSAs), 79, 80, 82
 - main argument against, 89–90
 - second premise of, 91–97
 - vagueness and *reductio* of, 90–91
 - Contextualism
 - controversies in, 114
 - in epistemology, 105
 - examples of, 119–122
 - forms of, 243
 - knowledge and belief in, 109–111
 - and objectivity, 122–126
 - and skeptic, 114–118
 - syntactic or semantic complexity in, 118
 - virtues of, 126
 - Contextual muddying
 - effects of, 96
 - phenomenon of, 92
 - Contextual salience absorption, 92
 - problem of, 93
 - Cooper, R., 25, 198, 201
 - Coppee, H., 155
 - Cormack, A., 195
 - Creative Writing Program, 239
 - Crimmins, M., 84
 - Culham, J. C., 271
 - Cybernetics, 256
- D**
- Davidson, B., 271
 - Davidson, D., 32, 34–37, 40–43, 51, 54–56, 79, 149, 172, 233, 238
 - analysis of ‘paratactic theory,’ 32–36
 - ‘paratactic theory’ of indirect discourse sentences, 31
 - T-theory, 98
 - Definite descriptions
 - approaches to, 197
 - interpreting, 198–200
 - presupposition, 200–203
 - theories of, 197
 - de Jong, F., 201
 - DeRose, K., 105, 119–121, 126
 - Determinate entailment, 178
 - Dever, J., 147–189
 - Diesing, M., 201
 - Direct expression account
 - reason for rejecting, 243–245
 - Disunity of truth* approach, 149
 - Disunity of Truth Hypothesis (DOT), 157
 - Domini, F., 271
 - Donnellan, K. S., 203
 - Dretske, F., 126
 - Driver, J., 264
 - Dugundji, J., 159
 - Dummett, M., 171, 172, 189
 - Duncan, J., 264
 - Dynamic semantics, 84, 187
 - and omniscience argument, 182–186
- E**
- Elbourne, P.D., 199, 217
 - Eliot, T.S., 238
 - Epistemic intuitions, 73–77
 - Evidential constructions, 222–223
 - application to, 228–229
 - singular or binary, 225–228
 - vs. source, 223
 - Evidentiality, object of, 224–225
 - Evil-making characteristics, 66
- F**
- Faller, M., 222, 228, 230, 233
 - Fara, D.G., 177, 178, 198, 199
 - Feldman, J., 271
 - Feldman, R., 126
 - Fine, K., 148, 174
 - FINST indexes, 260
 - and binding mechanism, 267
 - FINST theory, 255
 - von Fintel, K., 3–6, 8, 9, 13, 19, 21, 22, 24, 26, 27, 201, 213
 - account of “unless,” 19–20
 - uniqueness clause formulation, 22
 - First-order language, 174
 - Focal presupposition, 206
 - Focus, phenomenon of, 194
 - Fodor, J., 148
 - Fodor, J.A., 260
 - Fodor, J.D., 213
 - Force, terminality of, 152
 - Forster, M., 135
 - Fougnie, D., 271
 - Franconeri, S.L., 271
 - Frege, G., 79, 149, 150, 152–158, 198, 199, 227, 256
 - Frege’s terminality argument, 150–152
 - Fricker, E., 64, 73
- G**
- Geach, P., 162
 - Geis, M., 18, 19, 26
 - Gelade, G., 263–265

Geurts, B., 14, 206, 212
 Glanzberg, M., 193–217
 Gleitman, L., 142
 Godel, K., 159
 Goldberg, S., 62, 63, 65, 68–70, 72, 74
 Goodale, M., 271
 Graham, P., 64, 72
 Grammatical structure
 kinds of, 52
 logical form and, 51
 rules of, 256
 theory of, 47
 Grice, P., 79, 82, 126
 Groenendijk, J., 199, 205
 Gutman, D., 264

H

Hamblin, C.L., 205
Handbook of Logic and Language, 4
 Hand, M., 54
 Hanson, N. R., 258
 Hayhoe, M. M., 262
 Heal, J., 233
 Heim, I., 12, 48, 196, 199, 201, 202, 217
 Henderson, D., 63
 Herburger, E., 206, 207, 209
 Hidden indexicals, 84
 Higginbotham, J., 3–6, 9–11, 16, 18, 19, 36,
 39, 55, 198, 221–234
 Higginbotham's puzzle, 29
 Hirsch, E., 269
 Hoffman, J.E., 271
 Holcombe, A.O., 271
 Horn, L.R., 211–213
 Hornstein, N., 195
 Horowitz, T.S., 271
 Horwich, P., 172
 Human linguistic practices, 130
 Humphreys, G.W., 262
 Hunter, D., 105–127

I

Iatridou, S., 3–6, 8, 9
 Illusory conjunctions, 263
 Impoverished Context Shifting Arguments
 (ICSAs), 90, 93
 Independent recognition of objects, evidence
 of, 263–264
 Indirect discourse sentences, 43
 Indirect Report Test (IRT), 244
 Inferential disquotation principle, 169, 170
 Information processing, 256
Insensitive Semantics, 237
 Intensionality, structural theory of, 56

Inter-contextual disquotation (ICD), 89–90,
 97, 244
 Internal alignment tests, 170
Inter-sentential phenomena, 42
Intra-sentential phenomena, 42
 Intuitive truth conditions, 131
Ipsa facto, 68
 Irrelevance, problem of, 131–133, 140
 Izvorski, R., 223

J

Jackendoff, R.S., 206, 208
 Jacobson, P., 196
 Janssen, T.M.V., 4
 Johnson, K., 129–144
 Johnson, S.P., 271
 Jones, E., 65–67, 110–112, 262
 Jovicich, J., 271

K

Kadmon, N., 199, 201, 202, 204, 205, 207,
 208, 213
 Kamp, H., 150, 199
 Kanizsa, G., 258
 Kant, I., 156
 Karttunen, L., 205
 Kayne, R., 48
 Keane, B.P., 271
 Keenan, E.L., 198
 Kempson, R.M., 195
 Knowledge communication, 62
 Koch, C., 268
 Kosslyn, S.M., 259
 Kratzer, A., 12, 13, 48, 196, 199, 201, 205,
 209, 225
 Krifka, M., 207
 Kripke models
 intuitionistic logic, for, 164
 Kripke, S., 148, 163, 164, 168, 180
 Kripke semantics, 168
 Kripke structure, 148

L

Labianca, A.T., 271
 Lackey, J., 72
 Ladd, D.R., 204
 Ladusaw, W.A., 212
 Landau, B., 271
 Langford, C.H., 167
 Language analyzer, logical forms of, 259
 Lappin, S., 201
 Larson, R., 33, 48
 Lasersohn, P., 213
 Lecarme, J., 222

Leibnitz' law, formulation of, 33
 Lepore, E., 32, 65, 66, 71, 79–82, 85–102, 129, 148, 149, 172, 234, 237, 239, 243, 244, 246–248, 251, 253, 260, 269
 Leslie, S.-J., 3–30
 Levinson, S.C., 201
 Lewis, C.I., 167
 Lewis, D., 12, 13, 106, 126, 229
 Lewisian account, of modality, 180
 Lexical semantics, 162
 Liberman, M., 142
 Lindberg, D.C., 270
 Linguistic communication, 251–253
 Linguistic expressions, iterated concatenation of, 46
 Linguistic structures, chain of, 51
 Linguistic theories
 scientific theories, as, 133–140
 theoretical aspects of, 194
 Linguistic Theory and Davidson's Program in Semantics, 31
 Linguistic utterances, 256
 Literary metaphor, 237
 Liu, C., 144
 Li, Y.-H.A., 216
 Loewer, B., 32, 172
 Logical form, theory of, 51
 Longobardi, G., 34
 Long term memory, symbol structures in, 268
 Lotze, H., 154
 Ludwig, K., 32, 260, 269
 Lycan, W., 230
 Lynch, M., 159

M

McCloskey, J., 48
 MacColl, H., 159
 McConnell-Ginet, S., 195
 McGeorge, P., 271
 Markov algorithms, 256
 Marois, R., 271
 Marr, D., 257
 May, R., 196, 198
 Metalinguistic negation, 211, 214
 Metaphysical modality, 171
 Milne, A. B., 271
 Milner, D., 271
 Mingolla, E., 271
 Mishkin, M., 264
 Modal language, emergence of, 164–167
 Montague, R., 79, 198
 Morgan, A.D., 155
 Motan Principle, 182, 185

Multi-clausal constructions, interpretation of, 56
 Multiple objects, evidence for access to, 264–265
 Multiple object tracking (MOT) task, 255, 265
 Multi-valued logics, 147–148

N

Natural language
 adequacy of AACM for, 83
 analysis of, 197
 grammar of, 32
 knowledge-transmitting use of, 73
 logical forms of, 40
 philosophical theories of, 32, 77
 propositional attitude and indirect discourse, properties of, 31
 semantically compositional, 3
 semantics for, 55, 129
 structures of, 41
 Neale, S., 193, 197, 198, 201–203, 214
 Negation and focus, 203–204
 background on, 204–206
 descriptions, 209–210
 improper descriptions and, 210–214
 Negative quantifiers, 214–216
 Non-contradiction, law of, 147

O

Object-language operators, 169
 Ogawa, H., 271
 O'Hearn, K., 271
 Omniscience argument, 182, 185, 189

P

Pappas, C., 170
 Paratactic analysis
 graphical representation of, 39
 nonfinite clauses in, 54
 parsing, 52
 Paratactic parsing
 linguistic expression delineated by, 54
 Parataxis, 53
 notion of, 32, 44
 Parsimony, theoretical, 47
 Partial first-order model, 174
 Pearson, D., 271
 Pelletier, F. J., 4
Penumbral truths, 148
 Perl, T., 271
 Perry, J., 260
 Pesetsky, D., 35
 Pietroski, P., 30, 47, 57, 80, 137, 138, 195
 Pointwise entailment, 178

- Pook, P.K., 262
 Posner, M.I., 271
 Post Production systems, 256
 Potts, C., 212
 Pragmatic strategies, 86
 Pre-Fregean semantics (PFS)
 characterization of, 150
 maximalism and minimalism of, 152–156
 Presupposition obviation, 213
 Primitive visual objects
 individuating and tracking of, 265–266
 properties of, 263
 Principle of charity, 246
 Progovac, L., 212
 Propositional radical, 84
 Psycholinguistic studies
 human performance, of, 257
 Putnam, R., 137
 Pylyshyn, Z.W., 255–272
- Q**
 Quantified conditionals and compositionality
 containing either “if” or “unless,” 4
 meaning of, 3
 puzzle of, 5–7
 truth conditions of, 14–15, 29
 Quantifiers
 conditionals embedded under, 3
 semantic properties of, 47
 Quine, W.V.O., 54, 126, 127
- R**
 Radical Pragmatic Model (RPM), 84
 Radical pragmatics, 79
 debate with truth conditional semantics,
 80–84
 proponents of, 83
 Rao, R.P.N., 262
 Ray, G., 32
 Real context shifting argument (RCSA), 89, 90
 Recanati, F., 41, 80, 131–133, 138, 140,
 141, 143
 Recognition-by-components theory, 258
 Reimer, M., 99, 237–253
 Reinhart, T., 195, 201, 213
 Restricted quantifier, 198
 “Rich Rupert” stories, 95–97
 Rizzi, L., 45
 Roberts, C., 202, 204, 206–208
 Rock, I., 264
 Rooryck, J., 222, 226, 228, 232
 Rooth, M., 205–207
 Russell, B., 79, 86, 88, 193, 196–198, 214
 Rysiew, P., 127
- S**
 Safir, K., 39
 Saiki, J., 271
 Sainsbury, R.M., 198
 Salmon, N., 83
 van der Sandt, R., 206
 van der Sandt, R.A., 202
 van Santen, J.P.H., 135
 Sato, S., 264
 Schiffer, S., 126
 Scholl, B.J., 266, 267, 271
 Searle, J., 80
 Sears, C.R., 271
 Segal, G., 33, 41, 48
 Sellars, W., 201
 Semantically encoded content, 242
 Semantic maximalism, 153–154
 Semantic minimalism (SM), 80, 85, 88
 aspects of, 130
 doctrines of, 243
 interpretation of, 137
 Semantic partition, 205
 Semantics
 chameleon-like, 4, 8
 of conditionals containing “if”
 Conditional Excluded Middle, *see*
 Conditional Excluded Middle
 (CEM)
 modalized restrictive account, 12–17
 solutions, 7–8
 of conditionals containing “unless,” 17–19
 von Fintel’s account, 19–20
 modified account, 21–26
 uniqueness clauses and coordinate
 structures, 26–28
 “unless”-statements embedded under
 quantifiers, 28–29
 Dynamic, *see* Dynamic semantics
 hypothesis in, 152
 indirect discourse and propositional
 attitude, of, 31
 material conditionals, of, 5
 natural language, for, 55
 and non-semantic properties, 35
 phenomena of, 32
 properties of quantifiers, 47
 quantified “if”-statements, of, 4
 truth-conditional conception of, 187
 truth-conditionalism, 149
 unquantified “unless”-statements, for, 4
 update, truth, and disunity, 186–189
 Semantic strategies, 86
 Sethi, N., 271

- Sharvey, R., 198
 Simple "Assign and Combine" Model (SACM), 81
 Slemmer, J.A., 271
 Slifting diagnostics, 232
 Snyder, C., 271
 Soames, S., 80, 99, 101, 202
 Sober, E., 135
 Sosa, E., 64, 73
 Speas, M., 41, 222
 Speech
 assertoric, 74
 disquotational, 249
 ironic, 66
 qua contribution, 68
 Speech Act Pluralism (SPAP), 80, 85, 88, 243
 Speech reports, 65–67
 problems in context of, 67–70
 Sperber, D., 80, 238, 242
 Stainton, R., 77, 80, 102, 127
 Stainton, R.J., 32, 37, 80, 102, 131
 Stalnaker, R., 9, 126, 127
Stanford Encyclopedia of Philosophy, 4
 Stanley, J., 80, 126, 127
 Stavi, J., 198
 von Stechow, A., 205, 207
 Stokhof, M., 199, 205
 Storm, R.W., 265, 271
 Stowell, T., 55, 216
 Strawson, P.F., 201, 202, 213
 Subordinating coordinator, 54
 Suganuma, M., 271
 Superentailment, 175, 177
 Supertruth, 175
 See also Truth
 Supervaluations, 147–150
 objections to, 179–182
 Sutton, J., 172
Symmetric determiner, 25
 Syncategorematic structure, of proposition, 197
Syntactic predecessor, 47
Syntactic structure, derivation of, 51
Syntactic theory, 51
 Systematic semantic theory, 85, 88
 Szabolcsi, A., 212
 Szabó, Z.G., 4, 80, 161, 162
- T**
- Tarskian disquotational scheme, 169
 Tenny, C., 222
 Tense phrase (TP), 44
 Testimonial communication, 62
 Testimonial knowledge
 acquisition of, 64
 characteristic of, 70
 dialectical case for
 solutions for, 70–73
 problems associated with, 67–70
 what is said and, 62–65
 Thornton, I.M., 271
 Tipper, S.P., 264
 Torrego, E., 35
 Townsend, D., 140
 Transformational grammar, 256
 Travis, C., 80, 87, 88, 102, 126, 127
 Treisman, A., 263–265
 Trick, L. M., 271
 Truth
 conditional semantics of, 80–84
 disunity of, 156–158
 kinds of, 158
 ‘minimalist’ theories of, 173
 necessity as mode of, 158–162
 pointwise and setwise, 167–169
 vagueness as modality of, 174–177
- U**
- Ullman, S., 261, 262, 268
 Ungerleider, L.G., 264
 Universal quantifier, 55
 Update-to-test notion of consequence (UTC), 184
 Utterances, truth conditions of, 88
- V**
- Vagueness, disunified treatment of, 177–179
 van Marle, K., 271
 Veltman, F., 183, 184, 186
 Venn, J., 155
 Verkuyl, H., 201
 Visual Index, 255, 265
 theory of, 267–268
 Viswanathan, L., 271
- W**
- Watson, D. G., 262
 Wettstein, H., 260
 Whately, R., 155, 156
 Whitehead, A.N., 196–198
 Wiggins, D., 269
 Willett, T., 222
 Williams, M., 126
 Williamson, T., 176, 179–182
 Wilson, D., 80, 238, 242

Wilson, G., 195
Winter, Y., 201
Wolfe, J.M., 271
van der Wouden, T., 212
van Eijck, J., 150, 199, 201
Wright, C., 173, 176–178

X

Xu, F., 268

Y

Yagi, A., 271
Yantis, S., 262, 267, 271
Yokosawa, K., 271