

**Issues in the Philosophical Foundations  
of Lexical Semantics**

by

Brian Edward Ulicny

B.A. Philosophy,  
B.S. Electrical Engineering,  
University of Notre Dame (1986)

Submitted to the Department of Linguistics and Philosophy  
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Philosophy

at the

**MASSACHUSETTS INSTITUTE OF TECHNOLOGY**

May 1993

© Massachusetts Institute of Technology 1993. All rights reserved.

Author .....  
Department of Linguistics and Philosophy  
April 20, 1993

Certified by .....  
James Higginbotham  
Professor, Linguistics and Philosophy  
Thesis Supervisor

Accepted by .....  
George Boolos  
Chairman, Departmental Committee on Graduate Studies

**ARCHIVES**

MASSACHUSETTS INSTITUTE  
OF TECHNOLOGY

..... 00 1002

# Issues in the Philosophical Foundations of Lexical Semantics

by

Brian Edward Ulicny

Submitted to the Department of Linguistics and Philosophy  
on April 20, 1993, in partial fulfillment of the  
requirements for the degree of  
Doctor of Philosophy in Philosophy

## Abstract

My dissertation defends and explores the thesis that in order for a speaker to understand a natural language it is not only sufficient but necessary that the speaker tacitly know or “cognize” the truth-conditional contribution of the words and other sentential elements to the truth-conditions of the whole expression. A speaker’s semantic competence is to be explained as the employment of an internally-represented axiomatized truth theory for that speaker’s first language.

By providing a theory of truth for a language, the truth of certain sentences follows on the basis of that theory alone. In the first chapter, I develop and defend a notion of analyticity suggested by Noam Chomsky in his *Language and Problems of Knowledge* (1986) against skeptical worries due to Quine and Burge. On Chomsky’s view, analytic sentences are those sentences of natural languages true in virtue of “connections” between the semantic elements of the native linguistic endowment. I explain this idea of a semantic connection by way of a truth theory for the speaker’s language. In “Two Dogmas of Empiricism”, Quine argued that not even logical truths are analytic, however, since the seemingly fixed meanings of logical constants are empirically revisable. Quine’s worry is the seeming incompatibility of quantum theory with, e.g., the distributive law of classical logic. I argue that the acceptance of quantum theory does not bring about the revision, but rather the clarification, of the meaning of “and” and “or”. I contrast my view with those of Hilary Putnam and other writers. Natural language analyticities diverge from those of first order logics because the fixed semantic elements of natural languages and first-order logics differ. Part of the project of natural language semantics, then, is an account of logical form sufficient to expose the analyticity of a sentence. Chomskian analyticities are not coextensive with the sort of “folk analyticities” Quine targeted. “All bachelors are unmarried” isn’t analytic in this Chomskian sense.

In the next chapter, I consider Stephen Schiffer and Jerry Fodor’s arguments for the conclusion that representing and employing a truth theory is not necessary for understanding language. Schiffer’s argument consists in outlining the inner workings of a creature, Harvey, who comes to believe T-sentences appropriately but without

employing truth axioms for the elements of the sentences he hears. I argue that Harvey cannot serve as a model for our capacities if Harvey doesn't learn to understand his language. For us, learning this is most plausibly seen as learning a truth theory for the language. Fodor takes Harvey to show that an account of compositional semantic knowledge is redundant since semantic properties can be naturalized in terms of relations between the subject's brain and the world. I argue on the basis of the semantics of vague terms, tense, uninstantiated properties and other considerations that Fodor is not justified in supposing that the intentional content of thoughts and, thus, the semantic properties of expressions can be naturalized along the lines he envisions. Therefore, a theory of semantic knowledge is not superfluous to an account of understanding.

The project of producing a truth theory for a natural language frequently involves the uncovering of structure not apparent on the surface of a sentence. In the final chapter of my thesis, I explore the truth-conditional semantics of verbs in relation to the metaphysics of events. Terence Parsons' (1991) proposal goes beyond Davidson's original analysis of the logical form of action sentences in taking the logical form of, say, "Brutus stabbed Caesar" to include not only an event argument but also "thematic relations" borne by the sentence's arguments to the event. Parsons analyzes this sentence's truth-conditions as:

$$\exists e(\text{stabbing}(e) \ \& \ \text{Agent}(e, \text{Brutus}) \ \& \ \text{Patient}(e, \text{Caesar}))$$

where "Agent" and "Patient" are thematic relations borne by Brutus and Caesar to the stabbing event. I criticize Parsons' argument for such analyses, propose alternative arguments, and consider special problems arising for such accounts in treating apparent event identities.

Thesis Supervisor: James Higginbotham

Title: Professor, Linguistics and Philosophy

## Acknowledgements

I would like to express my deep thanks to my advisors, Bob Stalnaker, Noam Chomsky, and especially, my thesis supervisor, Jim Higginbotham, for their guidance in this project. It has been a great privilege to be able to discuss my work with them from whom I have learned so much both through our discussions and through their own work in this field. Moreover, they have provided me with an example I can only hope to emulate as to what it is to be a philosopher. It has been a great privilege to study at the Linguistics and Philosophy Department at MIT which seems to me to be the ideal environment in which to begin to ask questions about this deeply fascinating phenomenon of human language.

I am grateful to many other philosophers, linguists, and others at and around MIT for graciously discussing aspects of this thesis with me. These include: Richard Cartwright, Ned Block, Judith Thomson, Michael Antony, Tova Rapoport, Kate Kearns, Paul Pietroski, Trafford Welden, Tim Dilligan, Lenny Clapp, Rob Stainton, Marga Reimer, Catherine Womack, Brian Loar, Martin Davies, Jason Stanley, Katherin Koslicki, Tony Bures, Tom Green, Doug Jones, Michael Picard, Daniel Stoljar, Danny Fox, Jonathan Bobaljik, Colin Phillips, Alec Marantz, Irene Heim, Eric Lormand, Paul Bloom, Jeff Gruber, Shawn Kern, and anyone else I've inadvertently left out here. I would especially like to thank Paul Hoffman for his help and encouragement during his years at MIT, Wendy Weber for her administrative acuity, and Chris Collins for sharing his knowledge of linguistics with me so generously.

I would also like to thank Prof. Eugene Mellican at University of Massachusetts, Lowell, Prof. Robert Frederick at Bentley College, and Prof. David Layzer at Harvard University, for providing me the opportunity to teach while I was completing my thesis. Thanks to the guys in Türkoiz Köökie (Tim, Lenny, and Tom Green-the-Linguist), Col. Mahadevan ("Coach") Khandiah, and Rob Kempf, for making my stay at MIT so much fun. Special thanks, too, to Trafford Welden, for teaching me the ropes of graduate school as well as other fine points of deportment.

Sarah McConville has been an unflagging source of support, encouragement, love, and joy, over these past years. To say that I couldn't have done it without her only

begins to do the matter justice. Thanks for everything, my Sweetie!

Finally, I would like to dedicate this thesis to my earliest and best teachers, my parents, Mildred and Edward Ulicny, as well as my aunts, Srs. Cecilia and Mildred McClain, SC, for the support, love, and encouragement they've given me throughout my long years of schooling.

“Is there an Absolute?” asked His Highness [the Maharajah of Chhokrapur] suddenly. “That is what I want you to tell me. I look upon you as a kind of *weezard*; you must tell me these things. Is there an Absolute? Is there a God? Is there a future life?”

“Well,” I said, “you know the prayer of one of the Cato Street conspirators before his head was chopped off?”

“No,” said His Highness, looking at me with great expectancy. “What was the prayer?”

“He said, “O God—if there is a God—save my sould—if I have a soul.””

I smiled, and he hid his face in his sleeve and his small body shook with laughter; then looking up at me again, he said: “What did he mean?”

—from *Hindoo Holiday: An Indian Journal*

by J. R. Ackerley

# Contents

<b>1</b>	<b>Chomsky vs. Quine on Analyticity</b>	<b>11</b>
1.1	Chomsky on Analyticity . . . . .	12
1.2	Quine's Rejection of the Analytic/Synthetic Distinction . . . . .	19
1.2.1	Analyticity and The Revisability Thesis. . . . .	20
1.2.2	Analyticity and Indistinguishability . . . . .	36
1.2.3	Analyticity and the Roots of Reference . . . . .	50
1.3	Conclusions . . . . .	52
<b>2</b>	<b>Does Understanding Language Require Compositional Semantic Knowledge?</b>	<b>55</b>
2.1	Schiffer's Counterexample: Harvey . . . . .	57
2.2	Fodor on Semantics and Intentionality . . . . .	72
2.2.1	Fodor's Theory of Content . . . . .	78
2.2.2	Vagueness and Asymmetry . . . . .	81
2.2.3	The Actual History Condition . . . . .	88
2.2.4	Difficulties with the Nomic Relation Condition . . . . .	91
2.2.5	Naturalizing Logical Consequence . . . . .	97
2.3	Conclusions . . . . .	99
<b>3</b>	<b>Events and Thematic Relations</b>	<b>101</b>
3.1	Verbs and Events in Logical Form . . . . .	105
3.2	Arguing for Thematic Roles in Logical Form . . . . .	116
3.3	Parsons' Dream Machine Argument . . . . .	122
3.4	Event Identity and Thematic Roles . . . . .	135
3.5	Conclusions . . . . .	148

## Introduction

This dissertation explores the thesis that a speaker's semantic competence is explained by the speaker's internal representation and employment of a truth theory for that language. That is, the speaker's semantic knowledge determines a theorem (a "T-sentence") for each sentence *S* of the speaker's language that gives its truth conditions in some sufficiently articulated system of mental representation. This system of mental representations is to be thought of along the lines of what Jerry Fodor calls "Mentalese", the language of thought, a hypothesized calculus of brain state features over which cognitive processes are defined. On this view, understanding an uttered indicative sentence consists in the subject's coming to have the belief (or otherwise mentally representing) that what the speaker said has such-and-such truth conditions. Thus, if Alpha says to me, "Elvis swam", I come to believe of Alpha that what she said is true iff Elvis swims by employing the internally represented truth theory in some way.

Learnability constraints and constraints on a subject's computational resources entail that the truth theory for the indefinitely large number of sentences a speaker can understand must be finitely axiomatizable and recursively applicable. Thus, a speaker's semantic knowledge must consist of knowledge of the contribution of each lexical item and each elemental syntactic construction to the truth-conditions of a whole sentence. The fundamental question of lexical semantics, then, is this: "What does a speaker represent as the contribution of this particular word to the truth-conditions of expressions in which it appears?" This is to be distinguished from a fundamental question of the philosophy of mind: "In virtue of what does a mental state have the intentional content that it does?" Linguistic entities denote or are otherwise about things in the world only because mental states are, but it is not the semanticist's task to say in virtue of what a mental state has the intentional content that it does.

The thesis that a truth theory is part of a speaker's grammar is stronger than the thesis, associated with Donald Davidson's recent work, that although knowledge of a

truth theory or a language would be sufficient for understanding a language, it is not necessary. Such an account takes a truth theory to be merely a perspicuous way of systematically formulating an interpreter's strategy for determining what a particular speaker intends to convey by her utterances based on the available evidence and whatever empirical constraints there are on interpretation.<sup>1</sup> On the view I defend, understanding is not viewed as the construction of or convergence upon new truth theories (what Davidson calls "passing" theories) for each new speaker. Rather, the speaker's internally represented truth theory is taken to be a standing element of the speaker's linguistic faculty (her "I-language" in Chomsky's sense), on par with the speaker's knowledge of syntax and phonology. Just as phonological knowledge serves as the interface between language and speech, a truth theory serves as the interface between language and thought, providing truth-conditions for each structure determined by the language. As Davidson long ago pointed out, "meanings" are not required here; a truth theory is sufficient to link language and thought. It is in virtue of providing truth-conditions for each sentence of the language that this aspect of the speaker's grammar can be seen as providing "instructions" (to use Chomsky's term<sup>2</sup>) to be employed in interpreting, referring, expressing thoughts, and in other language-involving activities, just as phonological knowledge provides instructions for the articulation of speech.

This dissertation adopts the framework of Fodor's Language of Thought hypothesis as an account of mental processing. On this view, propositional attitudes and other psychological states (if not all of them) are computational relations to symbol tokens realized in some way by the subject's neural architecture. That is, we can think of the subject as having various internal buffers and boxes—one for beliefs, another for desires, and so on—individuated by their role in the subject's psychology. The contents of these inner mental spaces are Mentalese symbol tokens. Fodor has argued persuasively that mental representations must have constituent structure if they are

---

<sup>1</sup>See Davidson, "A Nice Derangement of Epitaphs". For criticism of Davidson's view, see B. C. Smith, "Understanding Language", *Proceedings of the Aristotelian Society*, 1992.

<sup>2</sup>See his "Language from an Internalist Perspective", p. 2.

to explain linguistic understanding. If we are able to understand indefinitely many sentences, and if one can understand a sentence such as “John loves Mary” only if one can understand its combinatorial variants such as “Mary loves John”, then these facts can only be explained by the employment of a finely articulated system of internal mental representations.<sup>3</sup>

The chapters that follow defend and explore this framework in relation to the question of analyticity in natural language semantics (chapter 1), the necessity of semantic knowledge (chapter 2), and the specific nature of the representation of events and event participation in the semantics of natural language (chapter 3).

---

<sup>3</sup>See J. A. Fodor, “Why There Still Has to Be a Language of Thought”, in his *Psychosemantics*, 1987, Appendix, pp. 139-154.

# Chapter 1

## Chomsky vs. Quine on Analyticity

In what follows I shall examine Noam Chomsky's objections to W. V. Quine's skeptical arguments concerning analyticity. What characterization of analyticity is Chomsky offering in response to Quine's arguments? I present Chomsky's objections and offer some criticism of Jerrold Katz's program to explicate analyticity along Chomskian lines. I proceed by examining two theses that Quine may be seen to have argued for in his "Two Dogmas of Empiricism": the Revisability Thesis—that the truth-value of any sentence is revisable in light of empirical evidence—and the Indistinguishability Thesis—that sentences true in virtue of meaning alone cannot be distinguished from sentences true in virtue of their meaning and their subject matter as well. I construct a Chomskian defense of analyticity that can be mounted against these theses.

A Chomskian defense of an analytic/synthetic distinction, as with all Chomskian arguments concerning the nature of human language, will insist upon the importance of the nature of human language acquisition in deciding the issue empirically. On what I shall offer as an explication of the Chomskian view, the class of analytic sentences in a speaker's first language can be given a clear empirical explication: a sentence is analytic in a speaker's first language just in case the right-hand side of its Davidsonian T-sentence is a logical truth with the respect to the fixed semantic

elements of “universal grammar”.

## 1.1 Chomsky on Analyticity

I will begin with two long quotations. In *Language and Problems of Knowledge*, and elsewhere,<sup>1</sup> Chomsky asserts that he is unimpressed by “the widely accepted and quite influential [conclusion] of modern Anglo-American philosophy...that there is no sharp distinction between analytic truths and statements that are true only by virtue of the facts.”

This solution seems quite erroneous. There is no fact about the world that I could discover that would convince me that you persuaded John to go to college even though he never intended or decided to go to college; nor is there any fact of experience even relevant to the judgment that you failed to persuade him if he never intended or decided to go to college. The relation between...“persuade” and...“intend” or...“decide” is one of conceptual structure, independent of experience—though experience is necessary to determine which labels a particular language uses for the concepts that enter into such relations. The philosophical debate over these matters has been misleading because it has focused on very simple examples, examples involving words that lack the relational structure of such terms as *chase* and *persuade*. Thus there is much debate over whether the statement “Cats are animals” is a truth of meaning or of fact...In such cases a decision is not easy to reach, but in others it seems quite straightforward....Furthermore, empirical inquiry can help clarify the status of a statement as a truth of meaning or of empirical fact; for example, inquiry into language acquisition and variation among languages. Thus *the distinction between truths of meaning and truths of empirical fact is an*

---

<sup>1</sup>N. Chomsky, *Language and Problems of Knowledge: The Managua Lectures*, MIT Press, 1988, hereafter (LPK). See also Chomsky, “Language and Problems of Knowledge”, ms, MIT, 1987.

*empirical issue*, not to be decided merely by reflection or, certainly, by stipulation.<sup>2</sup>

According to Chomsky, “[t]he whole matter requires extensive rethinking, and much of what has generally been assumed for the past several decades...appears to be dubious at best.”

Elsewhere Chomsky writes:

One would be hard put to find studies of language that do not assign structures and describe the meaning of *kill*, *so*, etc., in such a way that there is a qualitative distinction, determined by the language itself, between the sentences “John killed Bill, so Bill is dead” and “John killed Bill, so John is dead” Or to take another case, it would be difficult to find a study of referential dependence in natural language that does not conclude that the language itself determines that the relation holds between *Mary* and *herself* in *Mary expects to feed herself* but not when the same expression is embedded in the context “I wonder who...” yielding *I wonder who Mary expects to feed herself*. Such syntactic-semantic properties will induce cases of the analytic-synthetic distinction...But what Quine is alleged to have demonstrated goes beyond the matter of analyticity, reaching to the conclusion that there are no semantic connections that can be attributed to the language faculty itself as distinct from our general systems of belief.<sup>3</sup>

Thus, it is clear that Chomsky is emphasizing certain kinds of causative-stative entailments and sentences exploiting certain kinds of referential dependencies to be analytic, as opposed to such standard examples of the philosophical literature as “all

---

<sup>2</sup>Chomsky, 1987, pp. 33-4, italics added

<sup>3</sup>N. Chomsky, “Language and Interpretation: Philosophical Reflections and Empirical Inquiry” in J. Earman, ed., *Inference, Explanation, and other Philosophical Frustrations*, U. California Press, 1992., pp. 1-2.

bachelors are unmarried”, and so on. However, the debate over analyticity will not be settled by indicating putative analytic sentences of kinds not previously considered. What general criterion distinguishing all and only the analytic sentences of a language is being offered?

Chomsky’s response to Quine’s skepticism about analyticity is based on his conception of the language acquirer, i.e. the human subject who comes to understand a human first language. A Chomskian defense of an analytic/synthetic distinction for speakers of human languages will require appeal to the thesis that certain semantic concepts known to the speaker and employed in learning words are innate. Chomsky’s idea of “connections” in the semantic properties of sentences will then be explicated in terms of these innate concepts.

Chomsky’s views about analyticity must be understood within the context of Chomsky’s general views about the nature of human language. Chomsky views linguistics as a branch of naturalized epistemology. It is especially concerned with what he calls *Plato’s Problem*: “How comes it that human beings, whose contacts with the world are brief and personal and limited, are able to know as much as they do?” (LPK, pp. 3-4). Chomsky’s response to Plato’s Problem in the case of linguistic knowledge is well known. He holds that we have good reason to suppose that speakers of a natural language do not construct their grammatical knowledge by means of only general learning strategies since it is obvious that any of the general strategies proposed wouldn’t work, to the extent that they are detailed.<sup>4</sup> This highly intricate task is accomplished without explicit instruction and on the basis of impoverished data. “Analogical” reasoning isn’t adequate.

On Chomsky’s view, our ability to acquire a competence as intricate as that of linguistic competence as quickly as we do given only the impoverished data available to the child argues for innate structures guiding the child in acquiring the language. This is the nativist argument for innate structures guiding language acquisition on

---

<sup>4</sup>See LPK, Lecture I.

the basis of the poverty of the information contained in the learning stimulus. If linguistic competence is not explicitly taught (and it isn't), then it must be the language acquirer's native endowment that facilitates her acquisition of the set of faculties for the use and understanding of a human first language. That is, it is part of the speaker's nature that only certain language faculties may result as the mature state of the language faculty. With an innate determination of the structure of all possible languages as part of the speaker's native endowment, what remains for the human language acquirer is to be exposed to ambient linguistic data sufficient for acquiring one such possible structure. Chomsky calls this innate structure guiding the child in the task of language acquisition "Universal Grammar" (UG).

What does one acquire when one acquires a human first language? According to Chomsky, the grammar acquired determines a set of structural descriptions of all of the expressions that constitute a language. Such structural descriptions determine all of the syntactic, semantic, and phonological properties of every expression that is part of the speaker's grammar. Knowledge of this grammar is employed as providing a repertoire for such linguistic behaviors as "articulation, expression of beliefs and desires, referring, describing, and so on".<sup>5</sup> Grammatical knowledge consists in internally representing a grammar determining such a set of structural descriptions. Knowing how to employ this grammar for one's purposes is a distinct matter.

Chomsky's "poverty of the stimulus" argument for an innate parameterized space of possible syntactic principles is well-known, but how do such considerations lead to the postulation of innate aspects of lexical knowledge as part of the speaker's native endowment? And how does the postulation of innate lexical knowledge figure in a defense of the analytic/synthetic distinction?

The aspect of a speaker's competence that is crucial to the discussion of analyticity is the speaker's lexical semantic competence. This competence is perhaps of greater

---

<sup>5</sup>"Internalist Perspective", p. 2. For further details, see N. Chomsky, *Knowledge of Language*, Praeger, 1986; N. Chomsky, "A Minimalist Program for Linguistic Theory", MIT Occasional Papers in Linguistics, No. 1, 1992.

complexity in the case of some lexical elements than in others. For example, a speaker might hear the verb “to mortgage” quite a number of times in sentences in which she understands the meaning of every other expression, and yet fail to fully grasp not only all of the details of what it is to mortgage something but also fail to grasp all of the semantic and syntactic idiosyncrasies of the verb. On the other hand, as Chomsky notes, early stages of lexical learning proceed at a quick pace once the process is underway. There is a period during which the child acquires a great facility in learning words. Chomsky cites evidence that during the great swell of lexical-acquisition activity that children commonly exhibit early on, normal children are capable of becoming competent in a word’s use after only a single exposure to the word and, thus, acquire a great many words in a relatively short period of time.<sup>6</sup>

How are we to account for this speed in lexical acquisition relative to the data with which the child is presented? Although every exposure to a word’s use in ambient linguistic behavior comes with collateral information, this information is impoverished relative to the intricacies of even simple words and the concepts they express.

[T]he simplest concepts, for example, the concept of a nameable thing, . . . turns out to have remarkable intricacies, even involving the sophisticated idea of human agency, when one investigates it closely. Similarly, the concept of a person, one of the most primitive concepts available to a young child, is extremely complex and has been the subject of subtle philosophical inquiry for many centuries. Surely none of this is learned from experience. . . . The concepts that are available, independently of experience, to be associated with (or labeled by) words . . . do not constitute a mere list. Rather, like the sounds of language, they enter into systematic structures . . . [Concepts] such as action, agent of an action, goal, intent, and others, enter into the concepts of thought and language in complex ways. (LPK, p. 31)

---

<sup>6</sup>Compare this to Quine’s model of word learning as reinforcement of linguistic dispositions. It is hard to see how this process could be quick unless either the collateral information was extremely rich or the child’s dispositions are structured innately.

Chomsky here seems to be presenting a “poverty of the stimulus” argument for innate conceptual elements that the child will employ in mastering the semantic properties of words. The task of the child will be to (tacitly) determine how elements of her innate conceptual resources have their expression in the lexicon.<sup>7</sup>

As Ken Hale puts it, certain oppositions of semantic categories, e.g. nameable/unnameable, eventive/stative, causative/stative, referentially dependent/independent, are “part of the mental structures which enable human beings to acquire the semantic systems of their native languages”; any of these semantic distinctions within the lexicon would be “difficult, if not impossible, to learn on the basis of the data which a language learner would have in the normal course of language acquisition, suggesting that the semantic opposition involved is universal” i.e., an element of universal grammar. Such an opposition he concludes, is “therefore not learned, only the particular ways in which it functions in the grammar are learned”.<sup>8</sup>

According to Chomsky, words are triples of phonological, syntactic, and conceptual properties. From the postulation of semantic features of lexical knowledge one could move to a defense of analyticity along the lines of Kant’s classic notion of concept-containment: a sentence is analytic just in case the concepts of the subject contain the concepts of the predicate, and so on. Therefore, analyticity can be defined in terms of the conceptual aggregates associated with words in terms of subject-containment. Arguments for innate aspects of lexical knowledge serve to insure that some conceptual elements are shared by all speakers of a language.

One person who has tried to explicate a notion of analyticity along such Chomskian lines is Jerrold Katz.<sup>9</sup> For a number of years, Katz has defended an account of analyticity against Quinean skepticism on the basis of just such a decompositional

---

<sup>7</sup>For a discussion of the concept *nameable thing*, see Chomsky’s *Reflections on Language*. New York: Pantheon Books, 1975, pp. 46-52. There Chomsky claims that agency is involved in the concept of a nameable thing because, for example, an arrangement of leaves on a tree is not a nameable thing unless the arrangement is the product of human agency, perhaps as a kind of art form.

<sup>8</sup>K. Hale, “Notes on World View and Semantic Categories: Some Warlpiri Examples”, in P. Muyskens and H. van Riemsdijk, eds. *Features and Projections*, Foris: 1986, p. 252-3.

<sup>9</sup>Katz’s views are detailed in such works as his *Cogitations*, Oxford U. Press, 1986.

view of word meaning. On Katz's view, analyticity, like grammaticality, is a property of certain sentences and not others that is attested to by speakers of a language; therefore, a complete theory of grammar should incorporate means for making predictions as to which sentences a speaker of a particular language will judge to be analytic. This is to be accomplished by an appeal to a decompositional theory of word sense expressed in terms of tree-structured "semantic markers" and a more sophisticated notion of sense-containment employing graph theory. For example, Katz takes the semantic marker for the verb "chase" to be a tree containing such features as [physical], [activity], [movement],[purpose], [catching], etc. The analyticity of "If John chased Mary then John moved" is to be accounted for by the relation of the semantic marker of the consequent in the semantic marker of the antecedent.

This account doesn't explain analyticity, however. Rather, it is simply a means of notating it syntactically. Appeal to syntactic markers is circular in explaining analyticity. The Quinean skeptic about analyticity will ask what it is to say that word W has semantic feature [X] above and beyond saying that for the speaker or speakers about whom we are determining an analytic/synthetic distinction, appropriate sentences are judged analytic, e.g.:

(1) Everything that is W is X'.

where X' is a word whose only semantic feature is [X]? One can use semantic features to explain analyticity noncircularly only if the analytic/synthetic distinction has already been made clear. What Katz is telling us, however, is that a sentence is analytic just in case, simplifying a bit, the semantic features of the predicate are contained in the semantic features of the subject, and one is to determine the semantic features of the subject by asking which sentences the speaker finds analytic.<sup>10</sup>

---

<sup>10</sup>For similar objections, see John D. Greenwood, "Analyticity, Indeterminacy and Semantic Theory: Some Comments on 'The Domino Theory'", *Philosophical Studies*, vol. 58, 1990, pp. 41-49; N. L. Wilson, "Linguistical Butter and Philosophical Parsnips", *The Journal of Philosophy*, vol. LXIV, no. 2, February, 1967, pp. 55-67.

Katz's notion of analyticity formalizes a speaker or speakers' judgments about what sentences express definitions or criterial truths in addition to whatever else, however. The project of formalizing such "folk" analyticities as "All bachelors are unmarried" does not necessarily distinguish between truths of meaning alone (genuine analyticities) and truths which depend upon features of their subject matter as well taken to be criterial (folk analyticities). To object to Katz's theory is not, therefore, to reject the idea of semantic features. There is nothing wrong with postulating semantic features as theoretical entities if that is what the theory requires. However, Katz has provided no means for distinguishing semantic features that are implicated in genuinely analytic truths from those that figure in more "encyclopedic" criterial statements.

Further, there is no guarantee that a speaker must recognize all truths of meaning, in particular, sentences expressing complex natural language tautologies, as analytic. That is, sentences expressing complex logical truths need not be recognized as such by the speaker. As a formalization of certain (defeasible) judgments of analyticity, then, Katz's theory would not produce any deep distinction concerning the speaker's linguistic capacities that would stand up to Quine's skepticism.

I shall argue for a Chomskian analytic/synthetic distinction that relies crucially on this notion of innate semantic categories but not along Katzian lines. Such an account can be defended in light of Quine's recent retrospections on his original argument and what he now takes to be its excesses. Before making a case for the Chomskian analytic/synthetic distinction, I would like to review Quine's original skeptical arguments especially in light of his recent retrospections upon it.

## **1.2 Quine's Rejection of the Analytic/Synthetic Distinction**

In the next two sections, I want to consider two theses:

- (2) **The Revisability Thesis:** For each sentence of a theory, there could be recalcitrant experience such that the revision of that sentences' truth value would be an optimal way of accommodating that experience within the theory, and only optimal revisions must be made to the theory.
- (3) **The Indistinguishability Thesis:** Sentences true in virtue of meaning cannot be distinguished from sentences true in virtue of features of their subject-matter.

These two theses are related to ways of fleshing out the distinction between analytic truths and synthetic truths that Quine rejects in "Two Dogmas of Empiricism". Against Quine's claims, I want to examine sentences Chomsky takes to be clear instances of analyticity, such as

- (4) If John persuaded Mary to go to college, then Mary intended to go to college.

and propose a way of understanding them that would meet Quine's scruples. I will understand Chomsky's counterexamples to show that there is a determinate range of sentences whose truth follows from the speaker's knowledge of grammar. Thus, we should take the notion of analyticity to have a grammatical explication rather different from, but related to, the notion Quine was attacking. Quine's skepticism should thus be seen to lead to a deeper understanding of analyticity rather than its rejection.

### **1.2.1 Analyticity and The Revisability Thesis.**

In "Two Dogmas of Empiricism", Quine has two acknowledged targets. The first is a belief in "the fundamental cleavage between truths which are *analytic*, or grounded in meanings independently of matters of fact, and truths which are *synthetic*, or grounded in fact. Quine's second target is the dogma of *reductionism*: "the belief that each meaningful statement is equivalent to some logical construct upon terms which refer to immediate experience" (p. 20). Quine goes on to assert, late in the

paper, that these two dogmas are “at root identical” (p. 41). In what sense are these dogmas identical and on what basis does Quine reject a principled analytic/synthetic distinction?

That there is something wrong, in Quine’s view, with maintaining a distinction between analytic and synthetic truths is equivalent to saying, of some sentence *S*, that no intrinsic fact about that sentence justifies a claim that *S* is analytically (or synthetically) true. This denial of any fact of the matter of a sentence’s analyticity is initially puzzling in light of Quine’s own seemingly clear explication of analyticity:

one is tempted to suppose in general that the truth of a statement is somehow analyzable into a linguistic component and a factual component. Given this supposition, it next seems reasonable that in some statements the factual component should be null; and these are the analytic statements. (“Two Dogmas”, pp. 36-7.)

According to this explication, a sentence (Quine prefers ‘statement’) is analytically true just in case no experience could ever lead to the revision of its truth-evaluation as the optimal way of accommodating experience; otherwise it is synthetic. That is, a sentence is analytic just in case if it is part of a theory *T* held by a subject, no revision of *T* on the basis of empirical experience will fail to include it; otherwise, it is synthetic. Quine’s point is that, as such, there are no analytic sentences and no synthetic sentences. Any sentence may be rejected as part of a theory on the basis of empirical experience; further, any sentence may be maintained as part of a theory if sufficient changes are made elsewhere.<sup>11</sup>

Quine’s rejection of an analytic/synthetic distinction despite his clear explication of it depends upon his commitment to a holistic conception of truth evaluation. If there are, as Quine puts it, “logical interconnections” among *all* of a speaker’s beliefs, then experience can be reflected in the truth-evaluation of sentences in a variety of inequivalent of ways. In fact, says Quine, on a fully holistic view of language any

---

<sup>11</sup>See M. Dummett’s, “The Significance of Quine’s Indeterminacy Thesis”, 1973, p. 375.

sentence's truth-value could be revised to accommodate experience; conversely, any sentence's truth-value could be maintained by making changes in the truth-values of other sentences in order to accommodate experience. It is thus because of the "interconnection" of all sentences with sentences whose truth-values are sensitive to experience—or, as Quine more scrupulously puts it in his *Philosophy of Logic*, to neural stimulations resulting from the "cosmic distribution of microphysical states over space-time"—that Quine denies that there is an intrinsic fact determined by a sentence's truth-conditions as to whether a sentence must be held true come what may or whether a sentence's truth-value could be revised. Conversely, no intrinsic fact determines whether some sentence's truth-value depends upon the course of neural stimulation or whether it can be maintained come what may.

Quine's argument for the rejection of the analytic/synthetic distinction on the basis of the Revisability Thesis, then, involves at least this line of thought:

- (1) Every sentence (of an empirical theory) combines with others to entail a prediction about sensory stimulation in a certain context.
  - (2) If such a prediction is defeated within a theory, then it shows that the truth-value of at least one of the entailing sentences must be revised.
  - (3) "Recalcitrant experience" of this kind doesn't determine which entailing sentence's truth-value must be revised.
- Therefore, (4) the truth-value of every sentence is, in principle, revisable, even those which were taken to be semantic truths.

The upshot of this is that any decision to take one group of sentences in such a network as true or false come what may and another as true or false depending on experience, is wholly pragmatic, not grounded in any intrinsic difference in such sentences' properties. That in any history of a particular scientific world-view there is a distinction to be made between sentence's whose truth was maintained in light of experience and those that weren't is not a deep fact about the sentences so distinguished but is merely a pragmatic fact about considerations made by the scientists

whose decisions shaped the set of beliefs in question.

Quine's argument is significant for the philosophy of language if we consider a subject's idealized total set of beliefs to be the empirical theory at issue. We may, for the sake of illustration, take the sentences at issue to be the sentences inscribed in a subject's "belief box". We are to suppose that beliefs about the semantic properties of the various words and constructions of the speaker's language are included in this totality of beliefs.<sup>12</sup> Quine's argument, therefore, suggests that they are vulnerable to revision as well.

I would like to reexamine the argument against analyticity on the basis of the Revisability Thesis. The Revisability Thesis is not simply the thesis that no sentence's truth value is immune from revision because one could change the meaning of the words involved. This is obviously true but uninteresting.

Chomsky has remarked that the banality of this thesis is easily seen by considering the parallel case for phonology. It is assumed that there is a distinction between pairs of words that rhyme and those that don't. Suppose, then, that someone points out that although the word denoting cows as now pronounced rhymes with "now", we might change the pronunciation of the word that denotes cows to /dog/ while maintaining its syntactic and semantic properties. We should not conclude from this that there is no fact of the matter as to whether the word "cow" and "now" rhyme.

Let us call this sort of revisability *Saussurean Revisability* after the linguist known for emphasizing that the arbitrariness of the relation between a word's phonological and semantic properties. In cases converse to the one above, Saussurean revisability entails that one could always change the meaning of a word picked out by its phonetic or orthographic properties. This thesis is of little interest, however.

Quine's Revisability Thesis is not a thesis of mere Saussurean Revisability. Quine rather says that no sentence is analytic since every sentence's truth value might in

---

<sup>12</sup>We may think of the language at issue here either as the speaker's I-language (in Chomsky's sense; see his "Language from an Internalist Perspective") or an idealization of a speech community's beliefs, if an idealization could be coherently made.

principle be revised in light of recalcitrant experience if *so revising it would be the optimal way of accommodating that experience*. That is, experience makes our beliefs even about the semantic properties of words vulnerable to revision.

I shall argue that the argument for the denial of analyticity from empirical revisability is not a good one. On the other hand, defenses of analyticity based on a defense of a certain class of sentence's empirical irrevisability are not necessarily convincing. If the Indistinguishability Thesis is true, there might be empirically irrevisable but non-analytic sentences. Thus, Chomsky's remarks about empirical irrevisability or the irrelevance of experiential data to a speaker's disposition to assent to putatively analytic sentences do not decide the whole issue.

In his recent article "Two Dogmas in Retrospect",<sup>13</sup> Quine attempts to make his rejection of analyticity in "Two Dogmas" more perspicuous. Once again, the issue is framed in terms of the revisability of statements within a total scientific theory. Quine's empiricism entails that every statement in the formalization of science participates in some number of valid arguments of a particular form: the premises are a "cluster" of sentences of total science, and the conclusion is an *observation categorical*. "An observation categorical", Quine writes, "is a generalization of the form 'Whenever, this, that', where 'this' and 'that' are observation sentences." Thus, the argument:

- (5) If solution x is acid, then whenever litmus paper is submerged in solution x, it turns red.
- (6) Solution x is acid.
- (7) *Therefore*, whenever litmus paper is submerged in solution x, it turns red.

is an example of the sort of argument Quine believes every sentence (even, redundantly, the observation categoricals themselves) to figure in as a premise. Quine's Revisability Thesis thus comes down to at least this: logic requires that at least one

---

<sup>13</sup>Canadian Journal of Philosophy, v. 21, no. 3, (Sept. 1991), pp. 265-74.

statement in an observation-categorical entailment must come to be rejected as false if an observation-categorical is disconfirmed in experience. However, the nature of entailment doesn't privilege the maintenance of any one premise's truth over another's. Thus, the revision of any one particular sentence's truth-value is not determined by the disconfirmation of an observation categorical it jointly entails. Pragmatically speaking, some revisions may be simpler than others, and simplicity counts in accommodating experience optimally in a system of beliefs. Thus, the Revisability Thesis holds that the revision of any sentence's truth-value might be the optimal revision to make in a system of beliefs, and we must only make optimal revisions to our total theory.

This conception of the vulnerability of *every* sentence to truth-value revision in light of an observation-categorical's disconfirmation is overdrawn, however; it depends on an extreme form of holism that is indefensible as a conception of scientific method. Scientific methodology does not construe every scientific statement as up for grabs in an experimental situation, but a strictly logical interpretation of the revisability thesis entails this. Valid arguments, that is, may be "diluted"; they are not made invalid by the addition of further, irrelevant but true statements as premises. The addition of the premise

(8)  $2+2=4$

to the above entailment of the litmus-paper observation categorical is also valid, for example. However, it would introduce chaos into the scientific method Quine seeks to explicate if the truth of  $2+2=4$  were vulnerable to revision in an argument in which it plays no essential role. It would be purely gratuitous to allow any sentence whatsoever to face revision in light of a particular disconfirmation from whose entailment it can be cut. Thus, not every sentence's truth value is vulnerable to revision in light of recalcitrant experience. To hold that it is is to hold an extreme form of holism that is indefensible as a scientific methodology.

An appreciation of extreme holism's excesses, I take it, is the motivation behind

Quine's retrenchment to "moderate holism" in his recent writings. Quine no longer holds that "the unit of empirical experience is the whole of science" ("Two Dogmas", p. 42). Rather, "clusters" of sentences within a total theory or sub-theory are held to be the unit of empirical experience. Any sentence at all is vulnerable to revision only if all valid entailments of observation categoricals are treated equally; this is Quine's rejected extreme holism. On the view I take to be his moderate holism, only sentences which are essential to a valid observation categorical entailment face truth-value revision. That is, sentences that are irrelevant to an entailment are not vulnerable to revision if the observation-categorical is disconfirmed. Thus, on moderate holism, sentences from all branches of science (including mathematics) may participate in observation-categorical entailments, but only those premises without which the conclusion is not entailed face revision in light of disconfirmation. I take this to be the upshot of Quine's embrace of moderate holism and his turn from emphasizing the confirmation or disconfirmation of the whole of science to "chunks" of science.

In Quine's considered position, however, there are beliefs left invulnerable to empirical revision by the embrace of moderate holism and the uncontroversial thesis that which premise(s) one should revise is not logically determined. Every logical truth stands out as eliminable from observation-categorical entailments, and, the particular semantic properties of the logical vocabulary provides the formal structure that grounds the validity of observation-categorical entailments. Logical truths are those sentences (not sentence schemata) that remain true under substitutions of non-logical constituents, *salva congruitate*. So, any sentence which expresses the logical relation of the sentences that figure in a valid argument will be logically true (since the argument is valid). Therefore, it will be immune from revision in light of disconfirming evidence if moderate holism is adopted because it will not itself figure in any observation-categorical entailment. Thus, to illustrate: the sentence

- (9) If, if solution x is acid, then whenever litmus paper is submerged in solution x it will turn red, and solution x is acid, then whenever litmus paper is submerged in solution x, it will turn red

is a logical truth of the form *if, if p then q, and p, then q* that expresses the logical relation between the sentences of the original argument that was the basis of its validity. This sentence is not, however, vulnerable to truth-value revision directly on the basis of the disconfirmation of an observation categorical since it doesn't figure essentially in any observation-categorical entailment.

As we have seen, then, logical truths are not themselves directly vulnerable to revision in light of experience, contrary to the Revisability Thesis. However, Quine goes on to argue that science may require the reinterpretation of any lexical item over the course of scientific inquiry: even those formerly considered logical constants. Thus, Quine holds that the revision of a putative logical truth's truth-value may be warranted in the course of scientific inquiry, but not directly in the face of an observation categorical's disconfirmation. Quine subscribes to this indirect revisability because he holds that if the simplest modification of one's total science, or the revision optimal for some other pressing methodological reason, were that by which the interpretation of the logical vocabulary were changed, then the Quinean scientist would be required to change his interpretation of even the logical vocabulary. Thus, the logical truth (9) is not called into question by the denial of the observation categorical (7) alone, but by the denial of (7) along with the simultaneous affirmation of (5) and (6). The interpretation of the logical constants involved would thus be required, assuming that consistency in the speaker's beliefs must be maintained.

When would empirical experience compel us to change the interpretation of a logical constant? Let us examine the specific example Quine has cited. Quine has always maintained that the experimental results of quantum physics could require an empirical revision in the meaning of our logical constants and, thus, the rejection of previously held logical truths.<sup>14</sup> Suppose, for example, that a certain speaker believes:

---

<sup>14</sup>This is suggested in "Two Dogmas", p. 43:

[N]o statement is immune from revision. Revision even of the logical law of the excluded middle has been proposed as a means of simplifying quantum mechanics.

and maintained in his 1986 "Reply to Jules Vuillemin", p. 620:

Even a truth of logic or mathematics could be abandoned in order to hold fast some

- (10)  $[(A \ \& \ B)]$  is true iff A is true and B is true.
- (11)  $[(A \vee B)]$  is true iff A is true or B is true.
- (12)  $[A \leftrightarrow B]$  is true iff A and B are both true or both false.
- (13) P: Electron e has position within range R.
- (14) M: Electron e has momentum within range Q.
- (15) (M & P) (from (10), (13) and (14))
- (16)  $P \leftrightarrow (P_1 \vee P_2, \dots \vee P_n)$  (where the  $P_i$ s uniformly subdivide the region R).
- (17) (M &  $(P_1 \vee P_2 \dots \vee P_n)$ ) (from (12), (15) and (16))
- (18) Therefore, whenever (M & P),  $((M \ \& \ P_1) \vee (M \ \& \ P_2) \vee \dots (M \ \& \ P_n))$  (from (17), (10) and (11))

The sentence (18) follows from the other beliefs. However, if the range of subdivided positions is made small enough, each conjunction “M &  $P_i$ ” violates the Heisenberg uncertainty principle placing physical limits on the joint precision with which we can determine both the position and momentum of a particle in principle; the product of the precision with which we determine a particle’s position and the precision with which we determine a particle’s momentum must be greater than a certain constant. The more precisely we know the position of a particle, the less precisely we can *in principle* determine its momentum, and conversely. Moreover, this limitation is not due to the interference of one measurement with the other. Measurement of the one variable makes the other objectively indeterminate within a certain range. The upshot of this is that each conjunct of the observation categorical (18) must be

---

casual statement of ephemeral fact. ... Could be abandoned—very well, but *would* be? Yes, in an extremity; there are the two oft-cited examples, intuitionist logic and the deviant logics that have been proposed for quantum mechanics.

disconfirmed, and so the whole observation categorical can and must be disconfirmed by experience.<sup>15</sup>

The consequence of this is that a revision is required in the speaker's beliefs. Specifically, it seems to be the speaker's beliefs about the semantics of "&" which need to be revised. It is the interpretation of "&" which is getting one in trouble, because the speaker believed that "&" could be used to conjoin any two sentences held true and produce a truth. This is impossible with quantum-physically incompatible sentences such as those above. The conjunction of two sentences which could independently be true isn't true in quantum physics if it violates an uncertainty principle. In particular, the distribution of "&" across "∨" produces the problems here. There is no problem in principle with determining the position or the momentum of the particle to either specified precision. It is only *jointly* determining the position and momentum to a precision less than the Heisenberg constant that causes the problem.

On this basis, then, the Quinean scientist is faced with denying what was held a logical truth. The old conception of "&" and "∨" supported the law of distribution. This law holds that every instance of the following schema is logically true.

$$(19) p \ \& \ (q \ \vee \ r \ \dots \ \vee \ z) \ \text{iff} \ (p \ \& \ q) \ \vee \ (p \ \& \ r) \ \dots \ \vee \ (p \ \& \ z)$$

The quantum physicist may well assent to the left half of the biconditional while rejecting every disjunct on the right hand side. Considerations such as this, Quine says, could motivate the rejection of what had been considered a logical truth and a revision of one's logic.<sup>16</sup>

---

<sup>15</sup>For a similar presentation, see David Bostock's "Logic and Empiricism", *Mind*, Vol. XCIX, No. 396, October, 1990, pp. 571-582. On the uncertainty principle, see Roger Penrose, *The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics*. New York: Penguin Books, 1991, pp. 248-50.

<sup>16</sup>This explication of how quantum phenomena might lead one to deny the law of distribution comes from Hilary Putnam's "Is Logic Empirical?" Putnam's essay is reprinted under the title: "The Logic of Quantum Mechanics" in his *Mathematics, Matter, and Method: Philosophical Papers, vol. 1*, Cambridge U. Press, 1975. See also his "Quantum mechanics and the Observer", in *Realism and Reason: Philosophical Papers, vol. 3*, Cambridge U. Press, 1983, where Putnam writes: "Perhaps the best way to think of quantum logic is this: in quantum logic, the rule of conjunction-introduction (from,  $p$ ,  $q$ , to infer the conjunction  $p \ \& \ q$ ) is restricted to compatible propositions  $p$  and  $q$ " where the

Is it true that the Quinean must reconsider the status of *every* sentence of the form (19) on the basis of this rejection? What effect does this have on the rest of his commitments? Indeed, does the rejection of (19) in the quantum mechanical case compel *any* global change at all? No, the simplest change in theory required by the rejection of (19) is not to globally revise the truth-conditional contribution of “ $\vee$ ” or “ $\&$ ”, but rather to acknowledge the qualification against physical incompatibilities like this one within the definition of the old logical constants themselves.

Quantum logics differ from ordinary first-order logics in that no theorem logically equivalent to the law of distribution is a logical truth in that language. Typically, a quantum logic is defined in terms of operations (“meet”, “join”, and “orthocomplement”) on the lattice of the closed linear subspaces of Hilbert space that model the system quantum-mechanically.<sup>17</sup> The quantum-logical correlates of “ $\&$ ” and “ $\vee$ ” (“meet” and “join”) aren’t distributive. Quantum logic is, thus, well-suited for the quantum domain in which distributivity fails because of physical incompatibilities.

However, to accommodate quantum phenomena, we needn’t reject ordinary logic completely in favor of a non-distributive logic; that is hardly the “minimal mutilation” to our previous science that Quine advocates. Instead, the domain of our classical logical connectives should be clarified so that problematic cases are excluded. That is, when the recursive truth definitions of the parts of the language answering to classical logic are delineated, qualifications should be made for troublesome quantum-mechanical statements about electron positions and momenta (or other such statements in this or other domains that turn out to be problematic).

There is clearly no purely formal reason why the Quinean picture cannot accom-

---

“new semantical relation” incompatibility is given a definition in terms of Hilbert spaces. Putnam goes on to say, “the decision not to conjoin statements which are incompatible is a way of making the distinction in the logic itself between cases in which both of the statements we know have predictive value and cases in which only one of the statements has predictive value after [a] measurement” (pp. 266-7). I am not concerned with the correctness of Putnam’s proposal vis a vis the interpretation of quantum mechanics but only with its relevance to the question of the revisability of logical truth.

<sup>17</sup>See G. Hellman, “Quantum Logic and Meaning”, *PSA 1980*, pp. 493-511. See also R. I. G. Hughes, “Quantum Logic”, *Scientific American*, Vol. 245, October, 1981; Michael Redhead, *Incompleteness, Nonlocality, and Realism: A Prolegomenon to the Philosophy of Quantum Mechanics*, Oxford: Clarendon Press, 1987, Ch. 7, “Realism and Quantum Logic”, pp. 153-167.

moderate Tarskian definitions of several connectives with different ranges of application. Thus Michael Dummett writes that “[Q]uantum logic, as presented by Putnam, involves the proposal to introduce new connectives . . . explained operationally, alongside the classical connectives . . . explained truth-functionally” rather than the overthrow of the old logic for a new, quantum logic, as Putnam suggests in “The Logic of Quantum Mechanics”.<sup>18</sup> The quantum logician’s constants corresponding to “meet” and “join” may be introduced alongside the classical “ $\vee$ ” and “ $\&$ ” without any contradiction. We simply make the clarification that

(20)  $[(A \& B)]$  is true iff A is true and B is true and *A and B are not physically incompatible, as determined by the quantum formalism.*

For instance, one could stipulate that a conjunction is true just in case each conjunct is true and the quantum-logical “meet” of the propositions is not *bottom* or *the false*. This is true just in case the propositions expressed are quantum-physically incompatible. This characterization of “ $\&$ ” is sufficiently objective to allow logic to operate independently of our activity. The question of which sentences are quantum-physically incompatible is not arbitrary; it is formally decidable. Thus, truth-valuations of this revised scheme would be objectively decidable as well.

In the face of quantum phenomena, then, we needn’t reject *all* statements of the troubling form (here, instances of the distributive law), whatever their content. We may qualify and clarify our logic such that all instances of the distributive law are true except when quantum-physically incompatible sentences are at issue. Accepting the fact that empirical results can bear on our logical practice in this way doesn’t entail Cartesian skeptical worries to the effect that we mistook instances of the distributive law for logical truths in the past.<sup>19</sup> We haven’t found that our logical practice was mistaken relative to the sentences to which we previously applied it. Rather, we are

---

<sup>18</sup>Dummett, “Is Logic Empirical?”, in *Truth and Other Enigmas*, Harvard U. Press, 1978.

<sup>19</sup>Cf. Descartes’s First Meditation: “[S]ince I sometimes believe that others go astray in cases where they think they have the most perfect knowledge, may I not similarly go wrong every time I add two and three or count the side of a square, or in some even simpler matter, if that is imaginable?”

opting to refrain from adding certain new sentences of a form previously thought to be logically true to our stock of beliefs. This should be seen as a *clarification* of one's former logical practice, not a revision of it. We have learned, in this case, something about the limits implicit in our former logical practice regarding "&". Specifically, we have learned that our ordinary logical practice regarding "&" was limited to compatible sentences. It is not clear what other qualifications of classical logic are required by quantum phenomena.<sup>20</sup>

It may be objected that it is illegitimate to appeal to the nature of the conjoined certain sentences as a way of clarifying our logical practice while keeping out the troublemakers. The objector will respond that logic is supposed to be "topic-neutral".<sup>21</sup> This supposition, however, is false in general: some logics, at least, clearly do have subject matters. Indeed, what can it be to talk about a temporal logic unless we mean a logic that is about times? Similarly, modal logics would seem to be about possible worlds. Even standard first-order logic would seem to be about individuals as conceived in a quite particular way: all truths about individuals are taken to be compatible with one another. In virtue of this, the semantics of logical connectives can be given by means of truth tables. It is by considering alternative logics that one becomes aware of the metaphysical assumptions that standard first-order logics embed.<sup>22</sup>

In any case, the clarification of "&" proposed here seems to violate the assumption that all lexical elements of the same *syntactic type* within a language (i.e. which can be intersubstituted *salva congruitate*) are also of the same *semantic type*. That is, elements of the same syntactic type are supposed to make the same kind of contribution to the truth-value of a sentence. The objection is that no logical constant, here "&", should require special qualifications if the other logical connectives don't. Such assumptions allows one to characterize validity, entailment, and so on, syntactically

---

<sup>20</sup>Intuitionists like Dummett are faced with the converse of the situation just described. They want to make logic more accountable to experience by eliminating all verification-transcendent statements from their logical practices.

<sup>21</sup>See S. Haack, *Philosophy of Logics*, Cambridge U. Press, 1978, pp. 5-6.

<sup>22</sup>See H. Putnam, "Vagueness and Alternative Logic", in his *Realism and Reason*, *op. cit.*

without speaking of the semantics of the individual elements of a logic.

This assumption doesn't express anything necessary about formal languages, however. Nothing prevents words of the same syntactic type from being of different semantic types. Consider names. Names are all supposed to contribute to a sentence's truth-value by denoting an individual. But an element with the same syntactic distribution as a name needn't make the same semantic contribution. John Etchemendy's "Nix" is an example of this. Etchemendy defines "Nix" as an element whose syntactic distribution is identical to that of a proper name but which renders any sentence in which it appears false.<sup>23</sup> The semantics of "Nix" is given by the truth-conditional axiom:

(21) "[<sub>S</sub>...[<sub>N</sub> Nix]...]" is false.

There is no reason why such an element could not be introduced into the vocabulary of a language. The possibility of such items shows that unless appeal is made to the semantic category of an item, as well as its grammatical or syntactic category, the status of every logical truth must always be vulnerable to additions to the language's lexicon and not just to recalcitrant experience. In order to maintain any logical truth, therefore, we must make some (perhaps tacit) appeal to the idea that a sentence is logically true just in case permitted substitution of lexical items of the same syntactic and semantic type preserve truth.<sup>24</sup>

It is, therefore, not illegitimate for one to appeal to semantic properties of certain sentences in order to preserve certain logical truths in certain domains. As a logical truth, all substitution instances of (19) within the range of compatible sentences will come out true. Moreover, this envisioned logical bilingualism seems to be exactly what does take place with speakers. No one, not even quantum physicists, has taken to denying the logical truth of a sentence such as

---

<sup>23</sup>See his *The Concept of Logical Consequence*, Cambridge: Harvard, 1990. pp. 40-1.

<sup>24</sup>An closer parallel here to the quantum case would be a language containing both "Nix" and "Nox" such that "Nix" refers to Nix and "Nox" refers to Nox and any sentence of the form "...Nix...&...Nox..." is false.

(22) Sarah will drive to Mt. Monadnock and either she will go via Rte. 93 or she will go via Rte. 24 just in case either Sarah will drive to Mt. Monadnock and she will go via Rte. 93 or Sarah will drive to Mt. Monadnock and she will go via Rte. 24.

on the basis of quantum phenomena. It is only within the domain of quantum physics that they abstain from assenting to all substitution instances of (19). Quantum phenomena, or other logically recalcitrant phenomena, need never compel the wholesale replacement of previously held logical truths. At most, what phenomena like that of quantum mechanics motivate is the establishment of further logical truths in addition to those of classical logic and a demarcation of the domain of the logical connectives.

In any case, it is not clear that there are scientific phenomena which call every former logical truth into question; other than this problem for “&”, there are no other cases of empirical phenomena contravening a logical law. Further Quine has provided no general demonstration that the revision of *any* formerly held belief would be the optimal way to accommodate some recalcitrant phenomenon. Quine’s example implicates a central semantic fact as revisable, but it does not necessarily implicate all of the concepts that might figure in a truth theory for our language.

The argument of this section, then, is this: Quine held that the truth value of any sentence may be revised in light of experience. Upon examination, it was shown that, according to the moderate holism that even Quine now embraces, logical truths are not directly vulnerable to revision in light of experience. Quine, however, says that simplicity (or other methodological) concerns may mandate the repeal of logical truths on the basis of recalcitrant experience, and he cites quantum theory as an example of such a threat. On his view, instances of the law of distribution (19) may be revised due to quantum phenomena if this is the most suitable revision that can be made. I have argued that no such wholesale revision of logic or logical truths need ever be the optimal way to accommodate recalcitrant experience. The optimal revision called for in the one case Quine provides as evidence for the Revisability Thesis would seem to be the adoption of a refined and extended logic, not the

rejection of the old logic. In such a clarified practice, the classical logical apparatus applies to physically compatible sentences, and the quantum logic applies within the quantum domain or any other domain modelled by such a lattice.

This seems to me the way to view Quine's argument in "Two Dogmas". However, in a recent reappraisal of Quine's article,<sup>25</sup> Paul Boghossian argues that Quine's rejection of an analytic/synthetic distinction is far more radical than it seems: the denial of an analytic/synthetic distinction requires the wholesale abandonment of what Boghossian calls "meaning realism", where "meaning realism" is the view that expressions have semantic properties for a speaker at a time. Thus, all appeal to semantic properties in what follows in order to establish a Chomskian distinction between sentences true by virtue of a speaker's grammar, and other truths, will be for naught. Rejecting the analytic/synthetic distinction precludes any appeal to semantic properties at all, according to Boghossian.

Boghossian's argument from semantic properties to an analytic/synthetic distinction is this: if there were semantic properties, then two expressions could have identical semantic properties, and, so, a sentence which exploited this sameness of semantic properties would be true in virtue of the semantic properties of the components. (Boghossian points out that, obviously, if there were no semantic properties, then there could be no analytic/synthetic distinction since there would be nothing in virtue of which analytic sentences would be true.) That is, if there were semantic properties, then, two predicates "F" and "G" could have the same semantic properties and, thus, the sentence

(23)  $\forall x(Fx \leftrightarrow Gx)$

would be analytically true. So, since Quine denies that there are analytic truths, he must, on Boghossian's view, be denying what must be possible if expressions have semantic properties, and, thus, Boghossian considers Quine's rejection of analyticity

---

<sup>25</sup>"Analyticity", forthcoming in C. Wright and B. Hale (eds), *Blackwell's Companion to the Philosophy of Language*

to be equivalent to rejecting semantic properties of any sort.

This, however, misrepresents the point of Quine's argument in "Two Dogmas". In "Two Dogmas", as pointed out above, Quine gives a perfectly intelligible explication of analyticity as statements whose truth depends on no configuration of facts in the world; thus, it is not possible that Quine is arguing for the thesis that "is analytic" does not express a coherent property of sentences. Quine is arguing that no sentence is irrevocable, and so, not analytic.

Furthermore, Quine can't be rejecting semantic properties outright: truth itself is a semantic property. Thus, Quine is perfectly willing to talk about semantic properties in Tarski's sense, i.e. extensional properties. And it is obvious that if there are semantic properties, then there could be analytic truths, as Boghossian points out.

### 1.2.2 Analyticity and Indistinguishability

If it is false that for every belief there is an experience such that revising that belief's truth-value is the optimal way to accommodate that experience, where does this leave Quine's skepticism about an analytic/synthetic distinction? In a recent survey of twentieth-century philosophy of language, Tyler Burge writes:

[W]hat I regard as [Quine's] fundamental criticisms of analyticity have never been satisfactorily answered: No clear reasonable support has been devised for a distinction between truths that depend for their meaning alone and truths that depend for their truth on their meaning together with (perhaps necessary) features of their subject matter.<sup>26</sup>

Burge's point is this: where there is no pressure from experience to revise a sentence's truth-value, how can we distinguish truths of meaning from sentences which express substantive non-analytic truths which will never be disconfirmed? This is the problem

---

<sup>26</sup>T. Burge, "Philosophy of Language and Mind: 1950-1990", *Philosophical Review*, v. 101, No. 1 (January, 1992), pp. 3-52.

that Katz's analysis faced. What distinguishes the empirically irrevisable analytic truths from empirically irrevisable synthetic truths? In what follows I would like to show how a Quinean notion of logical truth, plus a Davidsonian truth-theory might be employed to draw a defensible analytic/synthetic distinction along Chomskian lines that meets Burge's worries.

In the chapter "Logical Truth" of his *Philosophy of Logic* (2nd ed.), Quine presents a Tarskian definition of logical truth in terms of substitution. "A sentence is logically true", he writes, "if all sentences are true that share its logical structure" (p. 49). Lest this seem unhelpfully circular, Quine goes on to add that one exhibits the logical structure of a sentence by replacing the predicates of a sentence with schematic letters. Quine gives as an example

(24)  $\neg\exists x(x \text{ floats} \ \& \ \neg(x \text{ floats}))$

which is a logical truth if all substitutions of predicates for "floats" are true as well. This is the complete story, however, only for the predicate calculus. That is, for the predicate calculi that Quine is considering, the logical structure of a sentence is simply the truth-conditional or quantificational skeleton of the sentence abstracted from the specific content of the non-logical components of the sentence.

In *The Philosophy of Logic*, Quine goes on to provide a completely general version of his definition of logical truth. "A logical truth," he writes, "is a sentence that cannot be turned false by substituting for lexicon, even under supplementation of lexical resources." That is, a sentence is a logical truth if it cannot be turned false by substitution *as suitably restricted* no matter how large we make the lexicon. However, some lexical items must be held fixed: after all, even the tautologies of propositional calculi can be turned false by means of lexical substitution among the logical connectives. Thus, the standard logical tautologies come out logically true only if one holds the standard logical vocabulary fixed.

Two issues, then, must be resolved in distinguishing a set of logical truths within a language this way: what elements are fixed, and what may be substituted for what?

Unless suitable restrictions are made, the sentence

(25) Abraham Lincoln was bearded or Abraham Lincoln was not bearded.

fails to qualify as a logical truth. That is, without suitable restrictions, the substitution of “every president” for “Abraham Lincoln” preserves grammaticality but yields:

(26) Every president was bearded or every president was not bearded.

which is false. Since it is false, then (25) would not be a logical truth by Quine’s criterion. Thus, an appeal to the fact that some NPs denote singular terms and that others are quantificational must be made. Further, even if a finer-grained conception of grammatical category allowed one to avoid these examples, one must make the assumption that elements of the same syntactic category make the same sort of semantic contribution to the sentence’s truth-value. Any language into which one could introduce Etchemendy’s “Nix”, for example, would have no logical truths similar to (25) as far as Quine’s substitution test goes. In fact, adding a lexical element like “Nix” to a language eliminates all logical truths.

As far as the fixed elements within the definition of logical truth, Etchemendy goes on to point out in *The Concept of Logical Consequence* that Quine’s generalization of the substitution criterion is both overwide and overnarrow in its scope. On the one hand, if “was president” and “was a man” are held fixed as well as the logical terms, then

(27) If John was U.S. President then John was a man.

is determined to be a logical truth, at least in this world. Nor does adding further names to the lexicon change its status. On the other hand, if only the standard constants of predicate calculi are held fixed, then

(28)  $\exists x \exists y (\neg(x=y))$

is judged to be a logical truth by this criterion, if it is true. But the fact that there are two distinct things in the universe is not usually taken to be a logical truth. Nevertheless, aside from such statements about the size of the universe, if the terms held fixed by a substitutional criterion of logical truth include only those items standardly held to be logical constants, then such embarrassments to the substitutional criterion as (27) are avoided. If the fixed set of terms is not limited to that of the standard logical apparatus, then, as we have seen, the range of logical truths determined by the substitution criterion changes from world to world.

Quine's substitution criterion, then, is not unproblematic as an account of logical truth. Depending upon what elements of the language are held fixed, certain truths in this world come out meeting the criterion for logical truth that intuitively aren't logical truths. Thus, meeting the substitution criterion is a merely necessary condition for being logically true unless the choice of fixed elements is determined independently. When we distinguish the logical truths of predicate calculi, we make appeal to such an independent criteria. Predicate calculi constitute a class of formal languages distinguished by the inclusion of specific kinds of connectives and quantifiers. We distinguish the class of logical truths substitutionally by holding fixed just these truth-functional connectives and quantifiers. We thus appeal to the type of semantic functions that these elements serve in order to distinguish a class of languages and the logical truths of languages in this class.

The connection with the Chomskian conception of analyticity can now be made. According to Chomsky, all natural language users bring a fixed set of concepts to bear upon the task of learning the lexicon. The notion of analyticity can be reduced to the notion of logical truth, then, if we suppose that the semantics of a sentence is given, in part, in terms of these fixed conceptual elements.<sup>27</sup>

---

<sup>27</sup>Chomsky has remarked (p.c.) that it is no more necessary for the semanticist to make precise the notion of analyticity than for a biologist to make precise any technical concept employed in doing biology: a science employs such conceptual distinctions as it finds useful in classifying phenomena and, if successful, discovers more about the categories as the science matures. However, there seems no reason not to clarify a concept within a science, here Chomsky's notion of 'conceptual connection', if one has the means to do so, and the notions of logical truth and innate semantic

The natural means for doing this is by means of a Davidsonian recursive truth theory provided as a theory of meaning for a natural language. On this view, to understand a sentence is to produce its truth-conditions in a meta-language (here “Mentalese”) by means of a finite recursive truth-theoretic compositional semantics. The axioms of such an internally represented theory provide necessary and sufficient conditions for an expression to be true, to be satisfied, or to refer. Recursively, they determine the truth-conditions of expressions involving these constructions or expressions as syntactic constituents. For example, axioms of a truth theory for English (or, better, a particular idiolect called for convenience “English”) might include the following:

(29) A sentence  $S$  consisting of a noun phrase and a verb phrase (i.e. [ $S$  NP VP]) is true (in English)  $\leftrightarrow$  the denotation of the NP is among the things that satisfy the VP.

(30) “Elvis” refers (in English) to  $x \leftrightarrow x = \text{Elvis}$

(31) “swims” is satisfied (in English) by  $x \leftrightarrow x$  swims

One such axiom would be required for each semantically distinguished item of a language. Given a syntactic description of a sentence as input, such axioms would generate T-sentences for each providing its truth-conditions. Knowledge of a recursive truth theory for a language would thus be employed in language understanding and production as the interface between thought and language.

Since, according to the nativist, all human beings bring a certain set of concepts to the task of language acquisition, we may suppose that the truth axioms assigned to lexical items and syntactic constructions manifests the contribution of innate conceptual items implicated in the speaker’s understanding of the word. Thus, each semantic element of UG, taken to be an innate, fixed feature of first language acquisition, will be notated as an element of Mentalese that figures in the truth axioms

---

elements of universal grammar provide such means in this case.

a speaker brings to bear in understanding sentences. The set of semantic elements of UG are then distinguished as the fixed elements of the metalanguage employed in natural language understanding. Thus, analyticity may be explicated as follows:

- (32) A sentence is *analytic* in language *L* just in case the Mentalese truth-conditions assigned to it by the speaker's T-theory meet the substitution criterion for logical truth with the semantic elements of UG taken as the fixed elements.

As such, human first languages contain analytic truths distinguishable from substantive synthetic generalizations.

Compare this with what Davidson himself says on this point:

Just as synonymy ... goes generally untreated [in a truth theory as a theory of meaning], so also goes ... analyticity. Even such sentences as *A vixen is a female fox* bear no special tag .... A truth definition does not distinguish between analytic sentences and others, except for sentences that owe their truth to the presence alone of of the constants that give the theory its grip on structure: the theory entails not only that these sentences are true but that they will remain true under all significant rewritings of their non-logical parts. A notion of logical truth thus given limited application, related notions of logical equivalence and entailment will tag along. It is hard to imagine how a theory of meaning could fail to read a logic into its object language to this degree; and to the extent that it does, our intuitions of logical truth, equivalence, and entailment may be called upon in constructing and testing the theory.<sup>28</sup>

Davidson here fails to consider the possibility of holding elements of the metalanguage fixed other than the logical constants.

Consider, for example, that one finds referential dependence as exhibited by anaphors and pronouns employed in all human languages. Thus, the notion of refer-

---

<sup>28</sup> "Truth and Meaning", p. 33.

ential dependence has a claim to being innate. Developmental evidence supporting a poverty of the stimulus argument for a child's understanding of referential dependence in language would secure such a claim. Suppose that there is such evidence. In that a speaker knows that anaphoric items can be coreferential with NPs in certain syntactic distributions, then sentences expressing this fact will be analytic. That is, their Mentalese truth-conditions will be logical truths with respect to the fixed elements of UG. Surely, knowledge of this coreferential capacity is part of a speaker's mastery of anaphoric elements. Thus, the truth-conditions of this sentence

(33) If John<sub>*i*</sub> shaved himself<sub>*i*</sub>, then John<sub>*i*</sub> was shaved by John<sub>*i*</sub>.

can be predicted to meet the substitution criterion with respect to "himself" and the index *i* on the basis of the semantics of coindexation and the syntax of anaphora. That is, the truth-conditions of any sentence structurally isomorphic with this one up to choice of words—where this includes grammatical gender, number, tense, etc.—will be true in virtue of the semantics of coindexation and the syntactic fact that anaphors such as "himself" must be bound: they must be coindexed with a c-commanding element in their domain.

The semantics of indexation, that is, is trivial:

(34) [<sub>*x*</sub> α<sub>*i*</sub>] denotes *x* with respect to sequence  $\sigma \leftrightarrow x$  is the *i*-th element of  $\sigma$

This axiom guarantees, however, that coindexed elements in a sentence denote the same element. Such an axiom would make the truth-conditions for (33) a logical truth.<sup>29</sup> Thus, the analyticity of sentences such as (33) is predictable. What this example shows is that if speakers bring the notion of referential dependence to the acquisition of language, then it follows that there will be sentences that can be predicted to be analytic on the basis of their syntactic form.<sup>30</sup>

---

<sup>29</sup>I will not detail the semantics of the passive construction, the *by*-phrase, and thematic relations here. See the third chapter, "Events and Thematic Relations".

<sup>30</sup>Furthermore, it is not clear how one could explain the analyticity of anaphoric relations in terms of Katzian semantic markers and features. Is the relevant semantic feature for "himself"

Consider also the example Chomsky often uses: the relation between “persuade” in “x persuades  $y_i$  [PRO<sub>*i*</sub> to  $\phi$ ]” and “intend” in “ $y_i$  intends [PRO<sub>*i*</sub> to  $\phi$ ]” (where “PRO to  $\phi$ ” stands for an infinitival clause with the coindexed element interpreted as the subject of the clause). In these contexts, these words encode a relation of cause to effect. That is, there is a “semantic connection” between “persuade” and “intend” such that

(35) John persuaded Mary to go to college.

entails, but is not entailed by,

(36) Mary intended to go to college.

because a speaker understands “John persuaded Mary to go to college” as being true just in case John caused Mary to intend to go to college (in the relevant way). Thus, the sentence

(37) If John persuaded Mary to go to college, then Mary intended to go to college

is analytically true, and in general, such relations between a causative expression and its correlated stative are analytic.

In order to make a Chomskian case for the analyticity of (37) it must be established that there is a conceptual opposition innate to speakers that is brought to bear in learning such pairs of words as *persuade* and *intend*, *cure* and *recover*, *teach* and *know*, *kill* and *die*. That is, in order to show the analyticity of (37), we must be able to appeal to a general fact about the design of language known to a speaker and exploited by the language learner in learning words. What general fact, then, is at work in the “persuade/intend” case? The fact at work here known to speakers

---

[+ anaphor]? If so, where is its counterpart in the marker for “If John shaved himself then John was shaved by John”? Are indices themselves semantic features? If so, are all sentences involving subjects and objects with coindexed NPs analytic? Surely, “John<sub>*i*</sub> loved his<sub>*i*</sub> mother” is not analytic.

concerning the design of the lexicon is that, roughly, if there is a verb  $V$  with  $n$  arguments expressing being in a state or participating in a certain kind of event, then there may be correlated causative verbs, let's symbolize them "CAUSE+ $V$ ", with  $n+1$  arguments, that express something causing that state or that event in some way perhaps idiosyncratic to that lexical item. The extra argument encodes the thing responsible for causing the other arguments to be in a relation of  $V$ -ing.

At least two verbs are related to "intend" in this construction in this way: "inspire", as in "John inspired Mary to go to college", and, of course, "persuade". To intend to do something is to be in a certain (mental) state. Thus, the speaker is disposed to incorporate into her lexicon verbs expressing something's causing someone to intend something. The notion of inspiring is such that the source of the inspiration need not act to cause the intention. The subject of "inspire" is not an *agent* in some sense encoded by the language, while the subject of "persuade" is. A persuader must *act* to cause the relevant state, not simply initiate a chain of causation leading to the state. The notion of persuasion does require that the subject intend to cause the intention and brings it about in some range of relevant ways.

One would make a case for the analyticity of (37) by appeal to a truth-conditional axiom of the following sort as internally represented and employed in understanding the causative "persuade".<sup>31</sup>

(39)  $\langle x,y,\phi,e,s \rangle$  satisfies "persuade"  $\leftrightarrow$  If [CP,VP] is infinitival, then (Agent( $e,x$ ) & Causes( $e,s$ ) & Intending( $s,y,\phi(y)$ ) &  $e$  is done in an appropriate manner) or if [CP,VP] is noninfinitival, (Agent( $e,x$ ) & Causes( $e,s$ ) & Believing( $s,y,\phi$ ) &  $e$  is done in an appropriate manner)

---

<sup>31</sup>A simple causative like "kill" would have a truth-axiom formalizable in Mentalese corresponding to:

(38)  $\langle x,y,e,e' \rangle$  satisfies "kill"  $\leftrightarrow$  (Agent( $e,x$ ) & Causes( $e,e'$ ) & Dying( $e'$ ) & Patient( $e',y$ )) &  $e$  is done in an appropriate way.

Thus, *Brutus killed Caesar* entails, but is not entailed by, *Brutus caused Caesar to die*.

That is, [CP,VP] is the complement phrase that is an argument of “persuade”: In the first sentence, [CP,VP] is infinitival; in the second, it is noninfinitival.

(40) [<sub>S</sub> John [<sub>VP</sub> persuaded Mary [<sub>CP</sub> [PRO to go to college]]]

(41) [<sub>S</sub> John [<sub>VP</sub> persuaded Mary<sub>i</sub> [<sub>CP</sub> that she<sub>i</sub> should go to college.]]]

In that the Mentalese truth-conditions determined for (37) by a T-theory along these lines would contain elements such as *Causes*, the truth-conditions would be a logical truth in Mentalese with such elements considered fixed. That is, the Mentalese truth-conditions determined for (37) by a theory making explicit the concepts involved would be (simplifying somewhat):

(42) (37) is true iff ( $\exists e \exists s (\text{Agent}(e, \text{John}) \ \& \ \text{Causes}(e, s) \ \& \ \text{Intending}(s, \text{Mary}_i, [\text{PRO}_i \text{ to go to college}] \ \& \ e \text{ is done in an appropriate manner}) \rightarrow \exists s' (\text{Intending}(s', \text{Mary}_i, [\text{PRO}_i \text{ to go to college}])))$ )

The truth-conditions of (37) are logically true with the innate semantic elements of universal grammar, such as *Cause* here, held fixed. On the other hand, the truth-conditions of “If John urged Mary to go to college, then Mary intended to go to college” would not be logically true, since “urge” is not a causative verb associated with the notion of intention in the requisite way.<sup>32</sup>

These considerations allow us to make sense of Chomsky’s comment that the concept of causation partially explains the child’s facility in learning the meaning of words. To say that the concept of causation underlies the child’s acquisition of words

---

<sup>32</sup>Of course, we must either hold fixed the logical constants of the metalanguage as well or make fine-grained distinctions, perhaps distinguishing a grammatical category of implication relations, within which substitutions can take place innocuously. This is to say that the child brings the concept of conditionality, and other logical concepts of the metalanguage, to the task of acquiring language. If, on the contrary, substitutions were allowed among the logical constants of the metalanguage, falsehoods could result from truths. Trivially, for example, if we are allowed to substitute “only if” for “if” above, as well as the other licensed substitutions, we will produce a falsehood in a case in which  $\alpha$  intended to go to college but wasn’t persuaded to do so by  $\beta$ . Cf. Gareth Evans’ discussion of “structurally valid inferences” in his paper, “Semantic Structure and Logical Form”, 1976.

is to say at least that a causal structure is reflected in the structure of the lexicon. That is, the child knows that if there is a word that expresses the notion of being in a certain state, then there can be words that express with their arguments the notion of something else's causing that thing to be in that state (in, perhaps, idiosyncratic ways). The child does not need to learn the meaning of the causative from scratch independently of learning the meaning of the stative. Nor does the child need to learn the meaning of the stative independently of learning the causative, if the order of learning goes this way. All that the child needs to learn is the phonological form associated with each. Applying this to the case at hand, if, as empiricists like David Hume have pointed out, it would be difficult to acquire the notion of causation on the basis of the available data, we have reason to believe it figures in a universal subclassification of the human lexicon.<sup>33</sup>

It may be objected, however, that positing such lexical subclassifications simply begs the question against Quinean skepticism. After all, is the analyticity of (37) empirically determinable consistent with Quinean scruples? Its meeting the substitution criterion is not enough to establish its analyticity without ruling in unwanted analyticities as well. Imagine a Quinean scenario such as this: suppose there was a world (or, at least, a world insofar as a speaker has experienced it) in which (37) meets the substitution criterion, but so also does

---

<sup>33</sup>Chomsky thinks that this specification of analyticity in terms of logical truth is unnecessary. Chomsky writes that it must be possible to establish empirically that the relation between "persuade" and "intend" is a conceptual connection by elaborating "the structure of the concepts, their primitive elements, the principles by which they are integrated and related to other cognitive systems, and so on". Then, one must "seek to show that other properties of language and other aspects of the acquisition and use of language can be explained in terms of the very same assumptions about the innate structure of the language faculty, in the same language and others, and that the same concepts play a role in other aspects of thought and understanding" ("Language and Interpretation", p. 14.) This talk about conceptual structure of the terms seems to beg the question, however. Many concepts are related to "persuade" and "intend", but, as Quine would ask, how are we to decide which conceptual liaisons are matters of meaning and which are matters of correlated belief? On the other hand, in suggesting that the child's task of acquiring the lexicon is explained in part in terms of causative/stative correlations, the analysis I suggest does explain the analyticity of (37) in terms of assumptions about the innate structure of the language faculty based on developmental and cross-linguistic evidence. Cross-linguistic and developmental evidence is taken to suggest that children are disposed to expect correlated causative verbs for stative verbs modulo certain restrictions.

(43) If John urged Mary to go to college, then John caused Mary to intend to go to college.

On the basis of this speaker's experience alone, one would have just as much reason to posit a connection in meaning between "urge" or "exhort" and "intend" in this structure as "persuade" and "intend".

However, even if the particular circumstances don't distinguish urgings from persuadings, this doesn't mean that the truth-axioms for the verbs aren't distinguished. What the linguist needs to establish the analyticity of (37) non-circularly is that the speaker knows facts about language on the basis of which one can predict a connection in meaning between "persuade" and "intend", and not between "urge" and "intend". Facts about experienced persuadings will not establish such a connection, however. Therefore, if such a connection is to be established, the case must be made on the basis of other grammatical knowledge. "Persuade" can be established as a causative verb on the basis of a speaker's knowledge that "persuade" is subclassified into a class of verbs to which "cure", among others, also belong; and these  $n$ -place verbs are related to certain  $n-1$ -place verbs as a causative to its correlated non-causative. That is, even in a world in which urgings are always followed by intendings, a case for a grammatical difference between "urge" and "persuade" can be established on the basis of a speaker's disposition to classify "persuade" differently than "urge". Speakers would acknowledge that "persuade" is to "intend" as "cure" is to "recover" but not as "urge" is to "intend" or "treat" is to "recover".<sup>34</sup> Such an appeal does not depend on the speaker's ability to imagine appropriate possible worlds or other introspective techniques legitimately called into question by such arguments as Putnam's against the analyticity of "all cats are animals".

The grammatical subclassification of pairs of words can be used to distinguish verbs classified in a speaker's grammatical knowledge as causatives from verbs which

---

<sup>34</sup>In an extremely recalcitrant world, such pairs could be made up and introduced by the linguist in order to elicit the data.

apply to merely contingent antecedent events.<sup>35</sup> Similar tests can be used to distinguish “intend” as the lexically-determined effect of “persuade” in certain constructions from a merely contingent consequent. That is, in a speaker’s experience, all persuadings to go to college might have been followed by attendings of college, but comparisons with a suitable other minimal pair can be used to distinguish “intend” as the lexically-determined consequent of “persuade”.

To generalize, then, intralingual grammatical subclassifications of lexical items into such categories as anaphora, causatives, etc., allow for analytic connections between certain words and constructions even when facts about use of these words in experience fails to distinguish analytic from non-analytic truths. That is, even if a speaker’s experience fails to distinguish persuadings from urgings in the expected way, a grammatical difference between “persuade” and “urge” can be established by a speaker’s assimilation of “persuade” to another causative for which the expected differences do hold.

This Chomskian criterion meets Burge’s challenge of distinguishing truths due to meaning rather than substantively true generalizations, because the relation of the relevant elements is not established by how things are in the world, but by grammatical evidence that the speaker classifies these words according to an innate semantic

---

<sup>35</sup>For a discussion of the differences in syntactic distributions between causative and noncausative verbs, see David Pesetsky’s *Zero Syntax*, ms. MIT, September 1992. Pesetsky cites, for example, Chomsky’s observation that if a verb appears in an inchoative construction, e.g. *Tomatoes grow*, then although there corresponds a causative verb “grow” as in *John grew the tomatoes*, the related derived nominal lacks a causative sense: \**John’s growth of tomatoes*. On the other hand, derived nominals with a causative sense, such as in *Bill’s cultivation of the tomatoes* (=Bill’s causation of the tomatoes’ cultivation) lack a corresponding inchoative: \**Tomatoes cultivate*. Such verbs do allow the matrix causative, however: *Bill cultivated the tomatoes*. Pesetsky suggests an explanation for this based on the hypothesis that causative verbs have an unpronounced causative affix. The fact that we don’t find derived nominals with a causative sense for inchoatives is to be explained by a further hypothesis concerning restrictions on affixation: words with unpronounced morphemes do not permit the affixation of further derivational morphemes. This is known as “Myers’ Generalization”. See Pesetsky, 1992, §3.4.5. In that the derived causative nominal “growth” would have both an unpronounced causative affix and a “nominalizer” affix, it would be impermissible according to Myers’s generalization. Further such evidence provides the syntactic side of evidence for the grammatical reality of the syntactico-semantic category “causative verb” as a well-defined subclass of the (non-intransitive) verbs. “Persuade” patterns just as “cultivate” in these examples: there is no inchoative form, and the derived nominal has a causative sense.

opposition. This evidence would come from cross-linguistic data and developmental studies. Further, this Chomskian criterion has nothing to do with empirical irrevisability.

The scope of this Chomskian notion of analyticity, then, is confined to expressions that encode universal semantic relations between expressions and constructions. It follows from the Chomskian view that no analytic connection exists between two expressions of lexical types that do not belong to universal semantic categories. Thus, such sentences as “all bachelors are unmarried men” fails to be analytic if the concept of marital status is not one of the innate semantic oppositions that the child brings to bear upon the task of acquiring the lexicon. There might have been such creatures, but it seems safe to say that this is not true of us.

In order to understand a word such as “bachelor”, all that one needs to know is how it contributes to the truth-conditions of sentences in which it appears. One may have any other set of beliefs about bachelors at all. It is neither necessary nor sufficient for understanding sentences including “bachelor” that one believe that all bachelors are unmarried men. One is free to believe that bachelors are Martian automata while still understanding the word in that one understands its truth-conditional contribution.<sup>36</sup>

This Chomskian notion of analyticity delimits a set of analytic truths based upon the semantic categories the language acquirer brings to the task of language acquisition in the initial, or universal, stage of her grammar. As such, Chomskian analyticity involves a relativization to the speaker’s species. Imagine a Martian who came to speak a language extensionally equivalent to your idiolect but who brings a completely different set of concepts to the task of language acquisition. Sentences analytic for you will not be analytic for him (her? it?), since the same elements of

---

<sup>36</sup>Cf. Burge on the dubitability of necessary truths by means of “non-standard theorizing” in his “Intellectual Norms and the Foundations of Mind”, *op. cit.* Burge argues, correctly I think, that one can understand a word without believing any criterial or necessary truths about it. Thus, even if I don’t believe that sofas are pieces of furniture made for sitting, perhaps hypothesizing that they are really liturgical objects, it does not follow that I don’t understand the word “sofa”. Perhaps some notion of “communal meaning” can be explicated with respect to which my understanding is partial. However, such conceptions of language communities have yet to be made clear and are, further, unnecessary to linguistic inquiry, as Chomsky has often emphasized.

the Martian's truth-conditions would not be held fixed as for the truth-conditions of sentences of your language.

### 1.2.3 Analyticity and the Roots of Reference

In *The Roots of Reference* (1974),<sup>37</sup> Quine gives a different characterization of analyticity than the one he gives in "Two Dogmas". In §21 ("Analyticity") of *Roots*, Quine writes,

A sentence is analytic if *everybody* learns that it is true by learning its words. Analyticity, like observationality, hinges on social uniformity. . . . The formula wants some refining. We should limit the people to those who learn the language as mother tongue. Also we should allow for chains of proof; we would want a recondite sentence to count still as analytic if obtainable by a chain of inference each of which individually is assured by the learning of the words. Perhaps this version of analyticity succeeds in drawing a rough line between sentences like "No bachelor is married" . . . and sentences that are not. At any rate it would seem that we all learned 'bachelor' uniformly, by learning that our elders are disposed to assent to it in just the circumstances where they will assent to 'unmarried man'.

Later, in his "Reply to Herbert G. Bohnert" (1986),<sup>38</sup> Quine writes:

There are sentences that we learn to recognize as true in the very process of learning one or another of the component words. 'No bachelor is married' is a paradigm case. Anyone who learned English as his first language, rather than through translation, will have learned 'bachelor'

---

<sup>37</sup>La Salle, IL: Open Court Press, 1974.

<sup>38</sup>In *The Philosophy of W. V. Quine*. La Salle, IL: Open Court Press, 1986, pp. 93-95. See also §22, "Domestic Meaning" in *Quine's Pursuit of Truth* (Revised Edition). Cambridge, MA: Harvard University Press, 1992.

through a paraphrase in which manhood and the exclusion of marriage are explicit.... Here ... is the germ of an intelligible and reasonable notion of analyticity. However, I see little use for it in the epistemology or methodology of science.... [S]ome truths are learned by learning words. Which ones are thus learned, however, and by whom, is scarcely to be determined, apart from a few clear cases of the sorts just noted. What is more, we need not care; for beliefs, however acquired, are fitted into a system without regard to individual pedigree. My old point about legislative definition and postulation, that it is episodic only and confers no enduring distinctive trait, applies equally to analyticity in the psychogenetic sense just now proposed.

In this revised view, Quine maintains the Revisability Thesis, while allowing that if everyone in a speech community learns that a sentence is true in learning a new word, then that sentence is *analytic* while still revisable. Quine suggests that there is only a germ of a criterion for analyticity here. The core of this suggestion is that we are to imagine that a child learns words by forming hypotheses concerning the contexts in which others of their community will assent to their use.

The appeal to community-wide similarities in learning seems to have the purpose of ruling out idiosyncratic learning patterns that don't correspond to analyticities: if one learns that if one's own elders are disposed to assent to "Uncle Arthur" then they will assent to "bachelor" and these circumstances exhaust their dispositions to assent to these queries, then "Uncle Arthur is a bachelor" would come out analytic for that individual.

So, too, "All bachelors are over one foot tall" could be learning-analytic for the general community. In general, circumstances inducing such analyticities of learning are possible, if perhaps less likely, at the community level as well, so it is unlikely that every sentence meeting Quine's learning criterion would be intuitively analytic. Imagine, for example, an isolated community in which the word for sin was invariably introduced via the story of Adam and Eve. There the sentence equivalent in that

community's language to "Adam and Eve sinned" would be learning-analytic for Quine, but surely not analytic in the intuitive sense.

Further, not every sentence that is intuitively analytic meets Quine's learning criterion. For example, Chomsky's examples of anaphoric coreference are surely intuitively analytic. However, no one need learn the meaning of "John", "shave", "himself", and so on, by learning "If John shaved himself then John was shaved by John". Even if one did learn one such truth by learning the words, there would be indefinitely many others just as analytic that one didn't learn in that way. The examples of analytic sentences Chomsky mentions need not be learned to be true in the course of learning the meanings of their constituents in order to be analytic.

### 1.3 Conclusions

Both Quine in his later pronouncements and Chomsky emphasize learning in their accounts of analyticity, but they employ this notion very differently. For Quine, a sentence must be learned to be true in learning its words if it is analytic. For Chomsky, however, it is the conceptual structures that the language learner brings to learning the meanings of words that makes the sentence analytic, independently of whether the speaker has ever even formulated the sentence to herself. On the Chomskian view presented here, a sentence is analytically true just in case the formalization of the truth conditions derivable for it by the speaker's internally represented T-theory meets the substitution criterion for logical truth with the semantic elements of universal grammar held fixed.

Quine still maintains the Revisability Thesis even though Quine has produced no argument for the conclusion that for every one of a speaker's beliefs, its truth-value might require revision as the optimal way of accommodating recalcitrant experience. Quine's repeated references to the threat that quantum phenomena pose to the meaning of the logical connectives are mistaken. Quantum phenomena such as failures of

distributivity do not require that we revise our opinion about formerly held logical truths, such as non-quantum mechanical instances of the distributive law, as the optimal way of accommodating that experience. Rather, the meanings of the classical logical connectives may be clarified and new connectives may be accommodated alongside of the previous logical constants.

In any case, the Revisability Thesis is only worrying to the extent that one accepts Quine's "web of belief" model of subjects along with its talk of "core" and "peripheral" beliefs. To the extent that the Chomskian need not accept Quine's picture of the ordinary subject as surveying and optimizing her web of beliefs, the Revisability Thesis is of no great concern in linguistic inquiry. Further, to the extent that the Chomskian is free to accept a certain degree of modularity for the internally represented grammar including a T-theory, the Revisability Theory need not be a concern for linguistic inquiry at all.

Finally, if the speaker's semantic competence consists of an internally represented T-theory, the Indistinguishability Thesis can be answered by the conception of analytic truths as sentences whose truth-conditions, as derived by the internally represented T-theory, are logical truths with the semantic elements of universal grammar taken to be the fixed elements. These fixed elements would be assumed constant across the whole species. Together with the T-theory encoding a speaker's semantic competence, they would determine a set of analytic truths for that language distinguishable from substantive truths and mere folk analyticities.

Quine and Chomsky do agree, however, that analyticity is an empirical matter. We must *discover* which sentences are analytic in a language. For neither writer does analyticity play the central epistemological role analyticity played for Carnap or other early analytic philosophers: analytic statements do not form the irrevivable core of a linguistic framework upon which we settle for some reason or other to do science. Further, for neither Quine nor Chomsky do analytic sentences express vacuous truths<sup>39</sup> or merely encode conventions.

---

<sup>39</sup>Analytic truths, that is, are not those for which an axiom of the speaker's truth theory simply

---

**asserts the sentences's truth: "S" is true". Analytic sentences have truth-conditions, and their analytic status depends upon them.**

## Chapter 2

# Does Understanding Language Require Compositional Semantic Knowledge?

Must speakers know a theory of meaning for the languages they understand? More specifically, must a speaker internally represent and employ the axioms of a truth theory for that speaker's (first) language? Following Donald Davidson's recent work, some philosophers hold that although knowing the axioms of a truth theory would be sufficient for a speaker to understand the sentences of a natural language, such knowledge, whether conscious or tacit, is not necessary. Davidson writes that "[t]o say that an explicit theory for interpreting a speaker is a model of the interpreter's linguistic competence is not to suggest that the speaker knows any such theory" ("A Nice Derangement of Epitaphs", p. 438; in LePore, 1986.). In what follows, I will examine an influential<sup>1</sup> recent attempt to demonstrate that speakers needn't internally represent the axioms of a truth theory for the natural languages they understand. This demonstration, due to Stephen Schiffer, proceeds by describing a creature, Har-

---

<sup>1</sup>See Hornstein, 1988 and Johnston, 1988.

vey, who comes to have the relevant T-sentential beliefs for indefinitely many natural language sentences by finite means and without representing or employing the axioms of a truth theory. I shall argue that Schiffer's counterexample does not show that internally representing the axioms of a truth theory isn't necessary for us based on considerations of our acquisition of the capacity to understand natural language. Thus, I shall argue that Schiffer has not refuted the early Davidsonian thesis that the best explanation of our *learning* to understand indefinitely many sentences via finite means is that we internally represent and employ a truth theory for that language in understanding.

Jerry Fodor is especially enthusiastic in supporting Schiffer's conclusion that a speaker need not internally represent and employ truth axioms in order to understand a natural language. Fodor supposes that for a speaker to internally represent the axioms of a truth theory would be *redundant* in that the intentional content of mental representations, and, derivatively, the semantic values of natural language expressions, is determined by naturalistic head-world relations, not by what the speaker knows. For Fodor, a theory of meaning for a natural language is derived from a theory of intentional content for the speaker's internal representations, and a theory of the intentional content of internal representations is, in turn, a theory of how the intentional content of mental symbols is determined by naturalistic head-world relations. Therefore, a theory of meaning is a theory about naturalistic relations obtaining between representations of natural language expressions in a speaker's head and the world; it is not something that must be internally represented and employed by a speaker in understanding. I shall argue, however, that Fodor has yet to provide us with good reasons to suppose that his program for naturalizing semantics and intentionality in terms of naturalistic head-world relations can be accomplished. In particular, I shall argue on the basis of vague predicates that Fodor's theory hasn't yet provided a solution to the crucial problem of misrepresentation. Other problems abound with the theory of content he has recently sketched. Thus, there is no reason to suppose that semantic theory is a theory about the relation of speakers' heads to the world and, therefore, no reason to suppose that internally represented semantic

theories are redundant to an account of speakers' understanding.

## 2.1 Schiffer's Counterexample: Harvey

In his book *Remnants of Meaning*<sup>2</sup>, Stephen Schiffer attacks the view that an appeal to referential properties is necessary to account for natural language understanding and linguistic behavior. The centerpiece of this is Schiffer's outline and defense of a counterexample to what he supposes is the widely-accepted thesis (U):

- (1) (U): It would not be possible to account for a human's ability to understand utterances of indefinitely many novel sentences of a language without the assumption that that language had a finitely storable recursive truth-theoretic compositional semantics.

More precisely, Schiffer's actual target in this chapter is the thesis that I will label (U').

- (2) (U'): It would not be possible to account for a human's ability to understand utterances of indefinitely many novel sentences of a language without the assumption that that speaker *internally represents* and *employs* a finitely storable recursive truth-theoretic compositional semantics in understanding those sentences.

By attacking the thesis (U'), Schiffer thinks he will have undercut the reason most often cited in favor of thesis (U).

The "axioms of a finitely storable recursive truth-theoretical semantics"<sup>3</sup> that are at issue are of the familiar Tarskian sort, giving necessary and sufficient conditions for

---

<sup>2</sup>Cambridge, MA: MIT Press, 1987. See, especially, Chapter 7, "Compositional Semantics and Language Understanding", and the first two sections of Chapter 8, "Compositional Semantics, Meaning Theory, and Ontology", pp. 179-219.

<sup>3</sup>I will use the phrase "axioms of a truth theory" (for a language) for short.

an expression to be true, to be satisfied, or to refer. These axioms recursively determine the truth-conditions of expressions involving these constructions or expressions as syntactic constituents. For example, axioms of a truth theory for English<sup>4</sup> might include the following:

- (3) A sentence S consisting of a noun phrase and a verb phrase (i.e. [<sub>S</sub> NP VP]) is true (in English) iff the denotation of the NP is among the things that satisfy the VP.
- (4) “Abe Lincoln” refers (in English) to x iff x=Abe Lincoln
- (5) “was President” is satisfied (in English) by x iff x was President

On the view that Schiffer is attacking, such axioms as these are internally represented by speakers of a natural language, and these axioms are deployed to compute or derive the truth-conditions of heard sentences in understanding. Schiffer’s strategy is to produce a counterexample to (U’) by describing a conceptually possible creature that has the capacity to generate appropriate beliefs about the truth-conditions of indefinitely many sentences but that doesn’t represent and employ truth axioms in generating them.

Schiffer proceeds to describe just such a creature, “Harvey”. Harvey’s inner processing operates according to formal algorithms such that whenever Harvey hears a sentence of English<sup>5</sup> uttered, he comes to believe that what the speaker said is true if and only if the relevant conditions obtain. For the purposes of his example, Schiffer takes coming to have a belief to be tokening a sentence in “Mentalese”, the hypothesized calculus of brain state features, in a buffer of brain space distinguished by its propositional attitude function, i.e. a “belief box”, or a “fear box”, etc. Furthermore, Schiffer asserts that he can say enough about Harvey so that (i) it will be undeniable that he understands natural language, and (ii) it will be undeniable that

---

<sup>4</sup>Better: an idiolect called “English” for convenience.

<sup>5</sup>Actually, English minus indexical and ambiguous lexical items. Schiffer proposes ways to handle these, too, within an extension of his counterexample, but I will not need to advert to this refinement.

his inner processing does not involve the axioms of a compositional semantic theory *either consciously or tacitly*.<sup>6</sup> In so describing Harvey, Schiffer asserts that he will have presented a counterexample to (U') showing "that we might be so constituted that our language comprehension abilities proceed without access to a compositional semantics for the language we comprehend" (p. 192). I shall argue, however, that we must internally represent and employ the axioms of a truth theory since our capacity to understand natural language is learned.

Harvey's mental processes are operations defined over sentences in Mentalese. As mentioned above, these sentences are the contents of various propositional attitude "boxes": one box contains the Mentalese formulae that constitute Harvey's beliefs, a second contains the Mentalese formulae that are his desires, and so on. The contents of the boxes and their interrelation serve to determine Harvey's behavior according to whatever sort of computational psychology turns out to be true.

According to Schiffer, the ability to understand sentences of natural language consists of nothing more than the ability to come to have the belief that what was said has the appropriate truth conditions whenever a speaker utters a sequence of sounds belonging to that language, *ceteris paribus*. Schiffer believes that this transition from the impingement of uttered sentences to the belief that what was said has such-and-such truth-conditions can be explained purely in terms of the "conceptual role" of three Mentalese expressions: "UTTERED", "SAID THAT" and "IS TRUE IFF" (p. 207). (Mentalese formulae will be notated by means of all caps.) The conceptual role of "UTTERED" consists in the fact that if the speaker denoted by Mentalese  $\alpha$  produces an appropriate acoustic signal, then, *ceteris paribus*, Harvey tokens a belief (produces a Mentalese formula within his belief box) of the form " $\alpha$  UTTERED  $\delta$ ". The conceptual role of "SAID THAT" is entirely given by:

---

<sup>6</sup>In his paper "Is Semantics Necessary?" (Proceedings of the Aristotelian Society 88 (1987)), James Higginbotham supposes that Schiffer's counterexample allows 'disinterpreted' knowledge of perhaps sophisticated semantic axioms in the form of permanent elements of a speaker's beliefs (p. 230). This, however, runs counter to Schiffer's explicit assertion that his counterexample make no appeal to any tacit semantic knowledge, stored in any form (*Remnants*, p. 195). Tacit knowledge of a complete truth theory for Mentalese in conjunction with knowledge of a translation scheme from English to Mentalese would seem to be equivalent to tacit knowledge of a truth theory for English.

- (6) If the formula  $[\alpha \text{ UTTERED } \delta]$  is in Harvey's belief box, then, *ceteris paribus*, so is the formula  $[\alpha \text{ SAID THAT } \mu]$  (p. 197).

The formula  $\mu$  is the output of a procedure  $f$  to be defined solely in terms of the formal, non-semantic properties of the input  $\delta$  that produces a Mentalese sentence with a specifiable intentional content given representations of utterances as input. Crucially, the procedure  $f$  is such that

- (7) if the referent of  $\delta$  can be used to say that  $p$ , then  $f(\delta)$  ( $=\mu$ ) would token the belief that  $p$  (p. 196).

Finally, the conceptual role of "IS TRUE IFF" is as follows

- (8) If the formula  $[\alpha \text{ SAID THAT } \sigma]$  is in Harvey's belief box, then, *ceteris paribus*, so is the formula  $[\text{WHAT } \alpha \text{ SAID IS TRUE IFF } \sigma]$  (p. 199).

Schiffer takes it that since these conceptual roles are sufficient for understanding and clearly do not involve the axioms of a truth-theoretic compositional semantics for English, then Harvey is a counterexample to (U').

The following example illustrates Harvey's capacity to generate T-sentences in his belief box that manifest his understanding without knowledge of the axioms of a truth theory for English. Suppose that Carmen makes the following assertion<sup>7</sup> to Harvey:

- (9) "Salsa is the ketchup of the 90s."

Schiffer supposes that the acoustical signal Carmen produces triggers a formal computational process that produces the following belief (a Mentalese sentence in his belief box) about who did the uttering and the phonological form of what was uttered.

---

<sup>7</sup>Only sentences in the assertive mode will be considered here for simplicity.

(10) CARMEN UTTERED [<sub>PF</sub> SALSA IZ THA KECH'UP OV THA NAYN'TEEZ]

Having this sentence in his belief-box triggers a procedure that takes the representation of the phonological form of the utterance and produces a belief about what was said in making that utterance.

(11) CARMEN SAID THAT SALSA IS THE KETCHUP OF THE 90s.

Finally, having this sentence in his belief box, results in the production of the T-sentence:

(12) WHAT CARMEN SAID IS TRUE IFF SALSA IS THE KETCHUP OF THE 90s.

The addition of such Mentalese sentences to Harvey's belief box constitutes understanding the utterance, according to Schiffer, just in case tokening the right-hand side of the biconditional alone in his belief-box would constitute Harvey's believing that salsa was the ketchup of the 90s. That is, the right-hand side of all the T-sentences so produced by Harvey must be *belief-realizing* if Harvey is to be truly said to understand the sentences he hears. In our example, that is, producing these Mentalese formulae would not be sufficient for understanding unless the formula "SALSA IS THE KETCHUP OF THE 90s" were such that if Harvey were to have it as a complete item in his belief box and as part of no other formula, then he would thereby believe that salsa is the ketchup of the 90s. Further, the tokening of such a formula must potentially figure in a complete explanation of Harvey's psychology.

It must be emphasized that Schiffer is not suggesting that if I came to acquire comprehensive phonological knowledge of, say, Urdu, then I would understand Urdu sentences if I were to token in my belief box a T-sentence with the mental representation of the disquotation of the heard Urdu sentence on the right-hand side. That is, if Carmen asserts the Urdu sentence "Mumbo jumbo", it wouldn't be sufficient for understanding the sentence that I came to have the following belief.

(13) WHAT CARMEN SAID IS TRUE IFF MUMBO JUMBO

In that I don't, by hypothesis, understand Urdu, the right-hand side of the produced T-sentence would not be belief-realizing for me. The right-hand side of the biconditional has to figure in a subject's psychology in the right way. Since "MUMBO JUMBO" isn't belief-realizing for me, it would not be produced by any of the usual processes of belief-fixation, nor interact with desires in the way that beliefs are said to in computational accounts, and so on. The represented formula would be psychologically inert.

Given this proviso that the right-hand side of the biconditional must be belief-realizing, the strategy of Schiffer's counterexample is easily seen. Schiffer thinks that the capacity to understand sentences of a natural language can be had simply by being such that one's head correlates the lexical items of English with the lexical items of Mentalese via their phonological properties. Harvey's inner processing merely maps the phonological form of the utterance to the corresponding Mentalese formula word by word. Here Harvey's inner processing must employ tacit phonological knowledge, but this does not violate Schiffer's strictures. Harvey's understanding capacity consists in his ability to produce T-sentences whose right-hand side is a belief-realizing Mentalese formulae by mapping phonological representations of heard English words with their Mentalese counterparts. Since *f* doesn't correlate phonological representations of the words of, say, Urdu with any Mentalese terms, Harvey will produce no T-sentences in his belief box upon being exposed to utterances in Urdu or any other language. Furthermore, we can assume that symbols that aren't elements of propositional attitude-realizing Mentalese formulae are psychologically inert. No cognitive process is defined over them.

Has Schiffer produced a satisfactory account of how human natural language understanding might work? Could understanding English really require nothing more than knowing the phonological properties of the words correlated with one's Mentalese formulae? If it were, then Harvey's capacity to understand wouldn't differentiate between interpretable and uninterpretable strings of English words. Harvey would be a

“(syntactic) garbage in, (semantic) garbage out” system, since Harvey would assign truth conditions to any string of uttered sounds that consisted of English words. For example, upon hearing Carmen assert

(14) “Some every favor good boy deserves”

Harvey would eventually token

(15) WHAT CARMEN SAID IS TRUE IFF SOME EVERY FAVOR GOOD BOY  
DESERVES

This counts against Harvey’s having the capacity to understand English. One doesn’t understand a language if one doesn’t distinguish nonsensical truth-conditions from those that are coherent. Harvey needs to be able to filter utterances that may be assigned truth-conditions from those that can’t.

A more sophisticated version of Schiffer’s Harvey model can easily be produced to remedy this. We may supplement Harvey’s inner processing with a function that recursively assigns a syntactic structure to every string of words recognized as part of a speakers’ utterance, i.e. a parser. We can then modify Schiffer’s procedure  $f$  so that it correlates all of the terminal nodes of a parse tree at the relevant syntactic level with a Mentalese lexical item just in case the string can be assigned at least one grammatical structure.

We may further specify that this parsed syntactic representation indicate the scope relations and so on that determine the interpretation of the asserted sentence. Syntactically ambiguous assertions could have more than one syntactic form assigned. Harvey’s parser, that is, will have to produce a syntactic representation at something like the level of syntactic representation called LF within Government-Binding syntax. At this level, the scope of quantifiers and operators and so on is disambiguated syntactically. The mapping from syntactic representation to Mentalese truth-conditions would then be trivial. The procedure would simply map each terminal symbol of an

LF parse tree to a specified Mentalese term left to right just in case the whole structure was assigned an acceptable LF parse. Our model of Harvey would now require both syntactic *and* phonological knowledge but no knowledge of the axioms of a truth theory for that language.<sup>8</sup>

Under this revision, if Carmen made the following assertion:

(16) “Everybody loves somebody”

this would produce a belief encoding the phonological properties of the uttered words and their order in Harvey’s belief box. This in turn would result in beliefs that the string can be assigned these two (acceptable) LF structures:

(17) [<sub>S</sub> [<sub>NP</sub> somebody<sub>i</sub>] [<sub>S</sub> [<sub>NP</sub> everybody<sub>j</sub>] [<sub>S</sub> t<sub>j</sub> loves t<sub>i</sub>]]]

(18) [<sub>S</sub> [<sub>NP</sub> everybody<sub>j</sub>] [<sub>S</sub> [<sub>NP</sub> somebody<sub>i</sub>] [<sub>S</sub> t<sub>j</sub> loves t<sub>i</sub>]]]

The belief that the sentence can be assigned the first parse, in turn, would result in the following beliefs.

(19) CARMEN SAID THAT SOMEBODY<sub>i</sub>; EVERYBODY<sub>j</sub>; t<sub>j</sub> LOVES t<sub>i</sub>

(20) WHAT CARMEN SAID IS TRUE IFF SOMEBODY<sub>i</sub>; EVERYBODY<sub>j</sub>; t<sub>j</sub> LOVES t<sub>i</sub>

That is, we assume that Harvey’s procedure *f* mapping well-formed LF representations into Mentalese T-sentences involves pairing the LF terminal symbol “somebody<sub>i</sub>” with the Mentalese element “SOMEBODY<sub>i</sub>”, the LF terminal symbol “loves” with “LOVES”, and so on. By hypothesis, tokening the right hand side of the biconditional (20) is sufficient for Harvey’s believing that somebody loves everybody (on the reading in which “somebody” has wide scope).

---

<sup>8</sup>Schiffer has objected to semantic knowledge on the grounds that it is “beyond the ken of plain folk” (p. ?), but so, too, is syntactic and phonological knowledge that is necessary in order for Harvey to understand.

Once Harvey knows that an uttered string can be assigned an acceptable LF structural representation, he can simply employ the trivial left-to-right mapping of LF terminal items into Mentalese terms in order to produce the relevant Mentalese T-sentence. All of the intellectual work on Harvey's part comes in recovering the string of words from the acoustical signal and in determining whether that string can be assigned an acceptable LF structure. Harvey need not employ the axioms of a truth theory to produce the T-sentences.

With these revisions, Harvey is such that he produces appropriate T-sentences in his belief box ("T-beliefs") for all and only the assertions that can be assigned grammatical LF representations without employing truth axioms.<sup>9</sup> Harvey is thus input/output equivalent to a speaker who produces all and only the same T-beliefs given the same assertions as input but who employs truth axioms to do so.

Does the conceptual possibility of Harvey entail anything about whether we speakers must represent and employ truth axioms? The issue of language acquisition is crucial here. Schiffer says nothing about how Harvey came to have this capacity to generate T-beliefs. For all that Schiffer says, Harvey may simply spring into being with his capacity to understand in place. In this, Harvey differs from us. Presumably, we speakers of natural languages are born without any capacity to generate T-beliefs. Furthermore, the capacity we do acquire depends upon the ambient linguistic environment. Since Schiffer's model builds in no dependence of Harvey's understanding capacities to his environment, it is perfectly consistent with Schiffer's model that, for example, Harvey be such that he maps English terminal nodes to Mentalese terms in a nonstandard way, mapping English "white" to Mentalese "BLACK", and so on. Or, Harvey could be such that he possesses only an Urdu understanding capacity in the midst of the monolingual English-speaking community with which he has had exclusive contact since birth. Harvey's capacity to understand language bears no relation

---

<sup>9</sup>Of course, even these revisions are not quite enough since we have no trouble assigning interpretations to nearly grammatical sentences such as "the child seems sleeping", non-assertoric utterances, and, perhaps, to "minor" expressions such as "your shoe", said to draw attention to one's conversant's smoldering footwear. (Rob Stainton has drawn my attention to the difficulty minors pose.) I will leave these further refinements aside.

to what Harvey's compatriots speak. If Harvey's brain were set up one way, he would understand certain sentences in a certain way no matter what others around understood by them; if his brain were arranged in a different way, he would understand sentences in other ways, again independently of the understanding of those around him. Such independence of understanding capacities to linguistic environment is not true of us speakers, and so, in this, Harvey is not a model of our capacities.

For Harvey to be a model of us, then, we must further revise him in order to incorporate a mechanism of language acquisition that makes Harvey's capacities dependent upon the ambient linguistic environment without requiring Harvey to employ axioms of a truth theory. Only a triggering mechanism meets these requirements. Harvey, that is, must be such that it is a brute fact that he comes to be disposed to map, e.g., "mango" to "MANGO", when he has been exposed to (a certain number of) well-formed utterances containing "mango", and he must also be such that he comes to be disposed to map "mangue" (French) to "MANGO" just in case he has been exposed to (a certain number of) well-formed utterances containing "mangue". The same must apply to every word of every possible natural language: it must be a brute fact about Harvey that for each word he comes to be disposed to map it to the appropriate Mentalese term in constructing a T-belief when this disposition is triggered by appropriate uses of the word in the linguistic environment. Otherwise, Harvey's capacity to understand wouldn't be correctly correlated with his linguistic experience. Harvey must not be supposed to induce truth axioms as facts about English in order to come to understand it; his capacity to understand an expression is a disposition triggered by his exposure to uses of that expression.

On the other hand, we might acquire the capacity to generate T-beliefs expected on the basis of our environment by learning facts about the ambient language. Here, learning semantic facts is understood as the selection and testing of hypotheses about the contribution of an expression or construction to the truth-conditions of whole sentences on the basis of the available data. Nativism concerning language learning assumes that somehow the space of hypotheses available to the learner is smaller

than the space of total possible hypotheses; the learner therefore has a head start in determining what facts capture the available data since a number of mistaken hypotheses need not be tested. Nativism does not entail that language acquisition is mere triggering. A nativist need not, that is, suppose that mere exposure to certain acoustical signals triggers a disposition to assign utterances of an entire language particular structures or particular truth conditions. Unlike semantic triggering mechanisms, then, semantic learning mechanisms would require the internal representation and employment of truth axioms for the expressions and constructions of a language.

Harvey can serve as a counterexample to claims about *our* linguistic capacities—the linguistic capacities of us human native speakers of natural languages—only if we could acquire language by means of triggering. If, however, we must learn to understand language, then it must be assumed that what we learn is the contribution of expressions and constructions to the truth-conditions of whole sentences. If learning is understood as a process of hypothesis selection and testing, then it follows that truth axioms must be represented and employed by human speakers. Thus, the Harvey counterexample shows nothing about us.

Which of these acquisition mechanisms, learning (theory construction) or passive triggering, is true of us? This is an open empirical question. However, the fact that children rapidly, but not flawlessly, acquire a mature semantic competence strongly favors the learning hypothesis over the pure triggering model of language acquisition. For example, the empirical fact (Bowerman, 1982; Pinker, 1990, discusses this example) that children are likely, at a certain point, to accept a sentence such as

(21) Daddy filled water into the glass

as well-formed and believe it to be true just in case Daddy poured water into the glass, suggests that exposure to utterances of “fill” didn’t trigger the disposition to have all and only the same beliefs that mature speakers have upon hearing “...fill...”. This fact suggests, rather, that the child supposes that the contribution of “fill” to the truth-conditions of sentences in which it appears is as follows.

(22)  $\text{Val}(\langle x,y,e \rangle, \text{fill}) \leftrightarrow (\text{filling}(e) \ \& \ \text{Agent}(e,x) \ \& \ \text{Theme}(e, y))$

The idiolect of the child who accepts (21) comes to more closely resemble our own when the child comes to believe that “fill” must have a container as its direct object, or Theme. That is, the child must come to know that not only is  $y$  a Theme, it must also be a container. “Pour”, on the other hand, requires a material as its direct (active) object, as in (21).

There are two possible explanations for the child’s revision of her understanding of a word such as “fill”. Under the learning hypothesis, the child is assumed to have revised her internally represented truth axiom for “fill”. Alternatively, under the triggering hypothesis, the triggering of the disposition to map “fill” to “POUR” is assumed to have been overridden by the triggering of a second disposition to map “fill” to “FILL”.

This postulation of a series of triggerings paralleling learning stages seems much more unlikely than the child’s revision of an internal hypothesis about the truth-conditional contribution of “fill”. Furthermore, as we saw above, the triggering hypothesis requires an infinity of triggering mechanisms within the finite brains of speakers, since every actual or possible word must have a mechanism dedicated to triggering the relevant disposition upon exposure to the word. Thus, these considerations suggest that we acquire our capacity to produce indefinitely many T-beliefs according to the linguistic environment by learning and not by triggering.

Why should it be an axiomatized truth theory that the child learns? Again, this is largely an empirical question, but empirical constraints would seem to favor seeing the child as learning a truth theory. For example, whatever compositional theory the child learns must be evaluable by the child during learning. Truth theories seem particularly well-suited for this; the child has the ability to check her theory against the conditions in which assertions are made. On the other hand, the constraint of evaluability rules out pairing utterances with world-time pairs or other abstract items pressed into service by a theory of meaning. Further, learning general conditions of

speakers' use of the words as opposed to the contributions of words to truth conditions seems a far too open-ended task for the child; the conditions under which speakers would use a word, its assertibility conditions, are much broader than the conditions under which the word is used truly. Thus, if the child must learn what words mean, as I think the child must, then it seems particularly plausible that what the child learns is the axioms of a truth theory.

Could the child, however, learn a procedure for producing T-sentences appropriately without learning a truth theory? We can imagine that the child has an inner homunculus whose job it is to figure out a finite procedure for producing T-sentences for heard sentences without learning a truth theory.<sup>10</sup> Might not the inner homunculus simply learn the procedure by which it maps the terminal nodes of the syntactic representation to appropriate positions in the Mentalese formula that represents its truth-conditions? No. In that the homunculus is behaving rationally, it seems that we can only explain why he comes to map "snow" with "SNOW" by invoking the notions of reference, satisfaction, and so on, to explain his behavior. The question here is why would the homunculus choose the intended procedure, mapping "[*S*...[*N* snow]...]" to "...SNOW...", rather than some other unless the homunculus were employing the notions of reference, satisfaction, and truth?

With no referential concepts guiding the homunculus, the problem of identifying the right mapping procedure would be just as hard as the Quinean picture of language learning would seem to be.<sup>11</sup> With no principles guiding the homunculus, tentative mappings would have to be tested against huge numbers of cases to get things right. Thus, learning would be quite hard. On the other hand, if the homunculus were disposed to assign truth-axioms to elements on the basis of general principles already at his disposal, e.g. that nouns, verbs, determiners, and so on have a characteristic way of contributing to the truth-conditions of sentences in which they appear, then

---

<sup>10</sup>Jim Higginbotham suggested this image of the inner homunculus as an device for thinking about unconscious learning.

<sup>11</sup>Quine: "[T]here is no justification for collating linguistic meanings, unless in terms of men's dispositions to respond overtly to socially observable situations." (*Word and Object*, MIT Press, 1960, p. ix.)

coming to acquire the ability to produce T-sentences appropriately becomes easy, but obviously amounts to acquiring a truth theory for that language. Thus, the ease with which children acquire semantic competence seems much more plausibly explained by the supposition that the homunculus accomplishes his task by bringing to bear the elements of a truth-theory which he then seeks to apply to the syntactic data rather than by simply casting about for ways in which syntactic elements can be paired with Mentalese features in the absence of any guiding principles except that one wants both sentences to apply to the same observable situations.

Of course, once a truth theory is acquired, one might employ the truth theory in producing T-sentences without using some sort of proof procedure. That is, once the truth theory is acquired, one might derive the T-sentences in some way other than by means of a proof procedure involving the truth axioms. This would still amount to employing the truth theory as a truth theory, however, since the procedure employed would depend upon the truth theory internally represented.

To conclude this section, Harvey *is* a counterexample to the thesis (U') considered as a thesis about conceptually possible creatures and not about human beings as they are. The best explanation of how human beings come to acquire language is by learning, and learning a grammar requires the inner representation of hypotheses as to what structural and semantic properties can be assigned to strings in that language, perhaps guided by innate principles. Schiffer rightly refuted this claim:

(23) (U'): It would not be possible to account for a creature's (cf. human's) ability to understand utterances of indefinitely many novel sentences of a language without the assumption that that speaker *internally represents* and *employs* a finitely storable recursive truth-theoretic compositional semantics in understanding those sentences.

If, however, language acquisition *for human beings* within the observed period requires learning or constructing a truth theory rather than acquiring a mapping procedure by passive triggering or in some other way, then the conceivability of a creature like

Harvey does not count against (U') as a thesis about us. Schiffer's discussion, then, forces us to formulate the thesis at issue more perspicuously as (U'').

(24) (U''): It would not be possible to account for a human's ability to *learn* to understand utterances of indefinitely many novel sentences of a language without the assumption that the speaker comes to internally represent and employ a finitely storable, recursive, truth-theoretic compositional semantics in understanding those sentences.

This thesis is essentially that embraced by Davidson himself in his early paper "Theories of Meaning and Learnable Languages" (1965). There, Davidson writes:

When we can regard the meaning of each sentence as a function of a finite number of features of the sentence, we have an insight not only into what there is to be learned; we also understand how an infinite aptitude can be encompassed by finite accomplishments....This argument depends, of course, on a number of empirical assumptions: for example, that we do not at some point suddenly acquire an ability to intuit the meanings of sentences on no rules at all....

In that paper, Davidson was concerned more with the idea of the learnability of a theory rather than with defending the thesis that such theories were in fact learned. That the semantics of a natural language must be learned by human beings and not merely triggered or "intuited" is an empirical assumption. If, however, we make this empirical assumption, then Schiffer's Harvey says nothing about us. Of course it is a *conceptual* possibility that there might be a creature who, as a matter of brute fact, is disposed to token T-beliefs upon hearing sentences uttered à la Harvey. For that matter, it is a conceptual possibility that there be a creature who is disposed to map indefinitely many complete well-formed utterances into the appropriate T-beliefs without employing any of the finite, recursive means of Harvey. This shows nothing about the necessity of internal truth axioms for creatures like ourselves who, it seems,

must come to be able to understand language by learning and for whom learning must be guided by innate principles.<sup>12</sup>

## 2.2 Fodor on Semantics and Intentionality

In his review of Stephen Schiffer's book *Remnants of Meaning*,<sup>13</sup> Fodor says that he thinks that the conclusion Schiffer derives from his description of Harvey is "exactly right": "To understand a sentence of English on this account just is to compute its M[entalese]-translation" (p. 186). Moreover, he says,

it's far from obvious that you have to know the *semantics* of an English expression to determine its M-translation; on the contrary, the translation algorithm might well consist of operations that deliver Mentalese expressions under syntactic description as output given to English expressions under syntactic description as input *with no semantics coming in anywhere* except, of course, that if it's a good translation, then semantic properties will be preserved. That purely syntactic operations can be devised to preserve semantic properties is the philosophical moral of proof theory...Syntax is about what's in your head, but semantics is about how your head is connected to the world.(pp. 186-7)

Fodor thus endorses Schiffer's view that truth axioms need not be internally represented and employed by subjects in understanding natural language. According to Fodor, semantic theories do not explain what speakers themselves know about meaning; the purpose of a semantic theory is for theorists to explain how natural language relates to the world via mental representations. Semantic theories are, therefore,

---

<sup>12</sup>See R. J. Matthews, "The Plausibility of Rationalism", in Matthews and Demopolous (eds.) 1989: *Learnability and Linguistic Theory*. Dordrecht: Kluwer.

<sup>13</sup>J. Fodor, "Stephen Schiffer's Dark Night of the Soul: a Review of *Remnants of Meaning*", Chapter 7, *A Theory of Content and Other Essays*, MIT Press, 1990.

supposed not to be internally represented nor are they supposed to explain how the speaker is able to understand indefinitely many sentences.

In the previous section, I argued that speakers must internally represent truth axioms if language is learned. In this section, I shall argue that we have no reason to suppose that there will ever be a naturalistic theory about how speakers' heads are connected to the world adequate to accounting for natural language understanding. The theory of intentional content that Fodor has outlined doesn't, in fact, contain even a solution to what Fodor feels is the most pressing problem for naturalistic theories of intentionality: the disjunction problem. (I will describe this problem and Fodor's treatment below.) There is, thus, no reason to suppose that internally represented truth axioms are redundant to an account of natural language understanding given purely in terms of head-world relations.

Fodor accepts Schiffer's conclusion that Harvey shows how we might employ nothing but syntactic and phonological knowledge in order to generate indefinitely many T-beliefs, but Fodor insists that the formal account of Harvey's understanding must be supplemented by a naturalistic theory of the intentional content of Harvey's T-beliefs. This is part of Fodor's main recent philosophical project: what he calls "the problem of meaning in the philosophy of mind". The problem is to produce an explanation of how it is that internal mental states (as individuated by their non-semantic properties, or "syntactically", as Fodor sometimes says) have a determinate intentional content within the context of a physicalist picture of the world. Fodor has repeatedly argued that a non-intentional, merely formal or syntactic characterization of Harvey's propositional attitudes isn't sufficient for psychological explanation. That is, it is not an adequate psychological explanation of Harvey's behavior to say, "He had formula B in his belief box and formula D in his desire box, so he invested in a salsa-making firm." A commitment to physicalism and intentional realism, as Fodor sees it, requires an account of intentional properties in non-intentional terms. Here Fodor's "physicalism" consists in the claim that the instantiation of any property of a "special" science (i.e. any non-physical science) is to be explained in terms of the

properties of ultimate physical theory, i.e. physics when it is completed.<sup>14</sup> In what follows, I examine the physicalist account of intentionality Fodor develops in his recent essay “A Theory of Content (Parts I and II)”<sup>15</sup> as complementary to Schiffer’s model of a language user.

Fodor considers that there are only three sorts of facts that could determine the propositional content of Mentalese sentences, including T-beliefs: either their intentional content is a matter of irreducible brute fact; or it is determined by the “saying potential” of the sentences to which they correspond; or it is determined by a naturalistic relation between the symbol tokens and features of the world. These possibilities are taken to be exhaustive, and he argues that only one of them is plausible. The open possibility, a naturalistic account of intentionality, is the subject of his essay “A Theory of Content I & II”.

The first possibility, that the propositional content of a Mentalese sentence is simply a brute fact, is rejected by Fodor out of hand. Fodor writes, “that a symbol means what it does *can’t* be a brute fact; it’s not the right *kind* of fact to be brute” (Review of Schiffer, p. 189). For Fodor, no complete naturalistic account of the world can include irreducible intentional properties and relations. Physicalism, the doctrine that the instantiation of all scientific properties must be cashed out in terms of the vocabulary of ultimate physics, precludes this. As he states in his *Psychosemantics*

I suppose that sooner or later the physicists will complete the catalog they’ve been compiling of the ultimate and irreducible properties of things. When they do, the likes of *spin*, *charm*, and *charge* will perhaps appear upon their list. But *aboutness* [i.e. intentionality] surely won’t; intentionality simply doesn’t go that deep. It’s hard to see, in face of this consideration, how one can be a Realist about intentionality without also

---

<sup>14</sup>Cf. Loewer and Rey, “Editors’ Introduction”, *Meaning in Mind*, 1991. For a critical discussion of this thesis and Fodor’s commitment to it, see T. Crane and D. H. Mellor, “There is no Question of Physicalism”, *Mind*, vol. XCIX, no. 394, April, 1990, pp. 157-206. Crane and Mellor argue, convincingly I think, that all of the features of intentionality that are supposed to be metaphysically problematic are found within acceptably physicalistic sciences like physics itself.

<sup>15</sup>In J.A. Fodor, *A Theory of Content and Other Essays*, MIT Press, 1990.

being, to some extent or other, a Reductionist. If the semantic and the intentional are real properties of things, it must be in virtue of their identity with (or maybe of their supervenience on?) properties that are themselves *neither intentional nor semantic*.<sup>16</sup>

As an argument, Fodor's line of thought here is completely question-begging; Fodor seems only to be stating his metaphysical intuitions. In any case, however, Fodor thinks he has compelling reason to suppose that the intentional content of Harvey's beliefs must have some reductive explanation.

Fodor next considers Schiffer's favored explanation: that the content of a Mentalese expression is given by the "saying potential" of its correlate in natural language. (See Sections 8.2-3 of *Remnants*.) Schiffer denies that a finite compositional theory of content for Mentalese expressions can be given. He therefore supposes that the only way to assign content to Mentalese expressions is via natural language; the most that can be done is to map Mentalese sentences to their natural language counterparts. Schiffer, that is, holds that the intentional contents of Mentalese formulae are determined by facts of the following form.

(25) The intentional content of Mentalese formula M is the proposition *that S* just in case M would be correlated with the sentence "S" canonically used to say that S as "S"'s truth-conditions if the subject were to hear "S" uttered, *ceteris paribus*.

Fodor rejects "saying potentials" as determinants of propositional content on grounds of circularity: such an account attempts to say both that sentences of a natural language "have their semantical properties in virtue of their relation to sentences of Mentalese", and that the sentences of Mentalese have their intentional properties in virtue of their relation to sentences of natural language. No such circular account will provide an adequate reduction of the content of a Mentalese formulae.

---

<sup>16</sup>Op. cit., Ch. 4, "Meaning and the World Order", p. 97

It is not entirely clear, however, that Fodor correctly imputes a circularity to Schiffer here. If the saying potential of a sentence is determined independently of facts about the Mentalese formula that gives its truth conditions, then there is no circularity involved in explicating the content of the Mentalese in terms of the saying potential of the sentence with which it is correlated. On the other hand, Schiffer doesn't provide us with any sort of recursive function from Mentalese formulae into their saying potentials (in a population? if so, how is the relevant population determined?). Nor does he seem to think that there can be such a function. Schiffer himself seems to think that the best that can be done is to produce a list of Mentalese formulae and their correlated saying potentials. However, the saying potentials of individual Mentalese expressions can't be given; Schiffer, therefore, holds that there can be no finite, recursive theory of the intentional content of Mentalese formulae. Our theory of the intentional content of Mentalese formulae can at most be partial if there are indefinitely many distinct Mentalese formulae.

A consequence that should give one pause regarding Schiffer's attempt to explicate the intentional content of Mentalese formulae in terms of the saying potentials of natural language expressions not taken up by Schiffer is that it rules out all forms of psychological explanation involving content not expressible in the (current) vocabulary of the natural language a subject speaks. Thus, Schiffer's account would be unable to give content to beliefs figuring (tacitly) in, to give two examples, explanations of early vision processing à la Marr or in phonological theory, that weren't statable in the subject's natural language. This consequence of Schiffer's thesis would seem to require an independent argument he does not provide.

Having dismissed the possibilities that (i) the intentional content of a Mentalese symbol is a brute fact, and (ii) that the intentional content of a Mentalese symbol is determined by the "saying potential" of the English expressions with which it is correlated, Fodor supposes that the only plausible option is to conclude that there is a compositional, naturalistic account of the propositional content of Mentalese sentences. It must be compositional in order to explain the *productivity* of the

propositional attitudes. That is, compositionality is required to explain our capacity as finite creatures to have indefinitely many syntactically and semantically distinct beliefs. Secondly, it must be naturalistic because Fodor's physicalism about scientific discourse requires that intentionality be ultimately reducible to the (nonintentional) vocabulary of ultimate physics. Thus, Fodor concludes that intentional realism requires us to have an account of the content of propositional attitudes, and this must take the form of a compositional theory of the content of the Mentalese sentences in terms compatible with the vocabulary of ultimate physics.<sup>17</sup>

The structure of Fodor's account of natural language understanding, then, is this. The disposition to produce indefinitely many appropriate T-beliefs upon hearing sentences uttered suffices for understanding only if the intentional content of these formulae is determined compositionally by a satisfactory physicalistic theory of intentionality. That is, the ability to produce a symbol of the form

(26) WHAT  $\alpha$  SAID IS TRUE IF AND ONLY IF  $\gamma$

is part of a complete physicalistic account of natural language understanding only if a compositional account of the intentional content of that formula can be given for each of its elements: "WHAT  $\alpha$  SAID", "IS TRUE", "IF AND ONLY IF", and  $\gamma$  (where " $\gamma$ " is the Mentalese formulae representing the truth-conditions of the uttered natural language sentence S). Intentional content must be determined for every part if the intentional content of the whole is to be determinate. For example, saying that a subject has a belief about Caesar and the kissing relation doesn't go far enough in determining the propositional content of that subject's belief. An acceptable theory must provide a complete proposition as the intentional content of every Mentalese sentence.

---

<sup>17</sup>Cf. B. Loewer and G. Rey: "Stating the whole of Fodor's theory succinctly, we might put it this way: propositional attitudes are computational relations to symbols encoded in the brain, whose broad content is determined by the properties onto which they lock and whose narrow content consists in a disposition to so lock. Psychology consists in stating laws about such dispositions, laws that are true by virtue of underlying, ultimately physical mechanisms that implement the computations performed on the symbols to which those dispositions are attached. ("Editors' Introduction", *Meaning in Mind*, 1991, p. xxx.)

In his recent work, Fodor has sought to naturalize propositional attitude content attribution by formulating a satisfactory set of *sufficient* conditions, all of which are expressed in non-semantic, physicalistic terms, for some Mentalese formula to be attributed a determinate intentional content. This is inadequate, however. The fact that Jupiter is larger than Mars is a naturalistic *sufficient* condition for a formula in my head to denote the proposition that snow is white, since I do believe that snow is white and, therefore, on Fodor's view, must have a formula in my head with this intentional content.<sup>18</sup> What an adequate theory of content must provide are *necessary and sufficient conditions* such that a formula in my head has a certain intentional content just in case certain physicalistically specifiable conditions obtain. Two questions then present themselves: (1) is Fodor's theory *adequate*, i.e. can it completely determine the intentional content of every Mentalese T-sentence? and (2) is it *correct*, i.e. are its assignments of intentional contents right?

### 2.2.1 Fodor's Theory of Content

Fodor's attempt to provide a physicalistic account of the meaning of Mentalese predicates is based on the causes of tokenings of Mentalese words ("concepts") in a subject. In his "Theory of Content" (*TC*), Fodor outlines physicalistic conditions for a predicate of Mentalese to denote a property instantiated by the cause of that token. Properties, for Fodor, carve up the world according to nomological relations. Thus, *being a horse* is, in the first instance, a property because the nomological relations of all horses are the same, but *being a horse other than Ol' Paint* is not a property in the Fodorian sense since the nomological relations of horses other than Ol' Paint and Ol' Paint don't, by hypothesis, differ.<sup>19</sup>

Fodor's theory is as follows:

---

<sup>18</sup>I owe this point to Ned Block.

<sup>19</sup>See *TC*, pp. 102-3. It is not clear on what basis Fodor rules out Quinean indeterminacies: is the property at issue *being an F*, *being a time-slice of an F* or *being an undetached part of an F*? Fodor does not confront this issue.

(27) The Mentalese predicate F means the property of *being F-ish* if [I will add: *and only if*]

- (i) '*F-ish* things cause "F" symbol tokens' is a law.
- (ii) Some "F" tokens are actually caused by *F-ish* things.
- (iii) For all *G-ish* things that are not *F-ish*, if *G-ish* things qua *G-ish* things actually cause "F" tokens, then *G-ish* things causing "F" tokens is *asymmetrically dependent* on *F-ish* things causing "F"s.

(Adapted from TC, p. 121)

Three factors, then, are important: the first is the (*ceteris paribus*) nomic connection between *F-ish* things and "F" tokenings.<sup>20</sup> It is on the basis of such a nomic causal relation that Fodor supposes a symbol to *carry information* about its nomic causes. The second requirement is the actual occurrence of "F" tokenings caused by *F-ish* things in a subject's causal history. The actual history condition is supposed to rule out, for example, twin-Earth XYZ from the extension of an earthbound subject's Mentalese "WATER" predicate.<sup>21</sup> Thirdly, the asymmetric dependence of non-*F-ish*-caused "F" tokenings upon *F-ish*-caused F-tokenings is the device by means of which Fodor prevents a symbol from incorrectly denoting a disjunction of properties: a disjunction of the property about which a subject carries information and the property of being *G*, where *G* denotes whatever other kind of thing that happens to cause an "F" tokening (by mistake). *G*-caused "F" tokenings are asymmetrically dependent upon *F*-caused "F" tokenings just in case *G*s wouldn't cause "F" tokenings but that *F*s did, but *F*s would cause "F" tokenings even if *G*s didn't.

It is not unfair to Fodor to turn his sufficient conditions into necessary and sufficient conditions since, first, Fodor calls for a complete physicalistic theory of in-

---

<sup>20</sup>For Fodor, *ceteris paribus* qualifications stands for the fact that not all interfering conditions need be spelled out within the theory of a special science.

<sup>21</sup>Thomas Kuhn has argued that it is physically impossible for anything other than H<sub>2</sub>O to have all of the testable properties of water. This, however, is beside the point; all that is required for the twin-earth cases are substances indistinguishable according to the subjects' present capacities to make distinctions. As such, even jadeite and nephrite, two minerals that both count as jade but which are indistinguishable by most ordinary means.

tentionality in order to explain natural language understanding, and necessary and sufficient conditions are required to make such a theory complete. Secondly, Fodor seems to have employed in his sufficient conditions the only resources general enough for the entire range of his project. Fodor's disregard for such special-scientific analyses of intentionality as given in what he calls "pop Darwinist" theories of intentionality<sup>22</sup> indicates a bias towards such basic, pan-scientific notions as the notions of causality, actuality, and counterfactual asymmetry that he employs. It is not apparent what other notions there are for Fodor to employ in such a theory. In any case, it is on the basis of the theory he has worked out in terms of these relations that Fodor bases his faith in a future physicalistic theory of content; so, it is not unfair to ask if this faith is well-founded.

Consider, then, this example of how Fodor's account is supposed to work. Mentalese "COW" denotes the property of *being a cow* if and only if, firstly, cows nomically cause "COW"-tokenings in a subject, *ceteris paribus*, i.e. just in case cows occasion mental states that include the Mentalese predicate "COW";<sup>23</sup> secondly, it is necessary and jointly sufficient that some cows have actually caused such tokenings; and, finally, non-cows wouldn't cause "COW"-tokenings but that cows did, but not conversely. Thus even if a horse on a dark night happens to have occasioned the tokening of a "COW"-belief in a subject, the asymmetric dependence condition prevents Mentalese "COW" from denoting the property of *being-a-cow-or-a-horse-on-a-dark-night* because horses on dark nights wouldn't cause "COW"-tokenings at all but for the fact that cows did. On the other hand, cows would presumably cause Mentalese "COW" tokens even if horses on dark nights didn't (presumably, that is, cows would even if *nothing* else did). Thus, for Fodor, these conditions provide physicalistic necessary and sufficient conditions for Mentalese "COW" to denote the property of *being a cow*.

---

<sup>22</sup>E.g. As in Ruth Millikan's *Language, Thought, and Other Biological Categories: New Foundations for Realism*, Cambridge, MA: MIT Press, 1984.

<sup>23</sup>Isn't it true that only cows in the vicinity cause "COW" tokenings? Why then doesn't "COW" denote the property of *being a local cow*? Presumably, Fodor would reply that local cows and non-local cows don't differ in the laws in which they figure, so "COW" doesn't simply mean local cow.

Let us begin to evaluate the adequacy of Fodor's account by considering some of the types of words for whose Mentalese correlates physicalistic content attributions would have to be provided.

## 2.2.2 Vagueness and Asymmetry

I shall argue that Fodor's theory of intentional content can't adequately account for the semantics of vague natural language predicates. Fodor's asymmetric dependence condition was supposed to solve the central problem for physicalistic theories of intentionality: the disjunction problem or the problem of misrepresentation. That is, the asymmetric dependence condition is what was to prevent "COW" from denoting the property of *being-a-cow-or-a-horse-on-a-dark-night* despite the fact that perceiving horses on dark nights sometimes caused "COW" beliefs. I shall argue that this condition necessarily gets the semantics of vague predicates wrong.

Vagueness is an ineliminable feature of the lexicon of a natural language, and thus must be dealt with by Fodor's theory of understanding. A predicate is vague just in case there are no necessary and sufficient conditions for its satisfaction statable in terms of the underlying facts. Grasp of a vague predicate  $\psi$  involves nothing more than grasping the applicability of  $\psi$  to a range of clear cases; grasping its inapplicability to certain other clear cases of non- $\psi$  things; and grasping that  $\psi$ -ness doesn't go over into non- $\psi$ -ness by a small change.<sup>24</sup> There is no reason to assume that competence

---

<sup>24</sup>See C. Wright, "Language-Mastery and the Sorites Paradox", in G. Evans, J. McDowell (eds), *Truth and Meaning: Essays in Semantics*, Oxford: Oxford U. Press, 1976. Wright develops ideas in Michael Dummett's paper "Wang's Paradox", *Synthese* 30, 1975. Wright writes that natural languages contain predicates

whose application is a matter of rough and ready judgment. We should have no use for a precisely demarcated analogue in contexts in which the word is typically used. It would, for example, be ridiculous to force the question of obedience to the command, "pour out a heap of sand here" to turn on a count of the grains. Our conception of the conditions which justify calling something a heap is such that the the appropriateness of the description will be unaffected by any change which cannot be detected by casual observation. (pp. 230-1.)

with such predicates can't be learned.

Suppose, then, that what it is to understand an English sentence "...heap..." is for a subject to token the T-belief

(28) WHAT THE SPEAKER SAID IS TRUE IFF...HEAP...

Any Fodorian relation  $R$  determining the intentional content of "HEAP" will determine necessary and sufficient conditions for satisfying that predicate since the nomological condition presupposes that all possible objects either would or would not cause a token of "HEAP" in the subject's head. Thus, whatever  $R$  relation Fodorian inquiry converges upon, there can't be any vague Mentalese predicates in Wright's sense since the  $R$  relation involves causation, and things either nomically cause "HEAP" tokens in a subject's head, *ceteris paribus*, or they don't.<sup>25</sup>

Moreover, Fodor's solution to the disjunction problem itself results in further problems for the semantic physicalist in dealing with the semantics of vague predicates. Consider the consequences of Fodor's asymmetric dependence conditions in the case of a vague natural language predicate such as 'is bald'. Suppose that Carmen asserts to Harvey, "Carlos is bald". Harvey, in that he understands English, then comes to token the Mentalese T-belief:

(29) WHAT CARMEN SAID IS TRUE IFF CARLOS IS BALD.

Fodor's views requires that there be a physicalistic determination of the property denoted by "BALD", but because of his asymmetric dependence condition his present theory must get the denotation of "BALD" wrong.

---

<sup>25</sup>Cf. Roy Sorensen's discussion of vague Mentalese predicates in "Vagueness within the Language of Thought", *Philosophical Quarterly*, Vol. 41, No. 165. (October, 1991), pp. 389-413. There Sorensen argues that Mentalese must contain vague lexical items, and the explanation of this vagueness is epistemic: we simply don't know where the cut-off for a vague predicate is. Ascriptions of vague predicates to borderline cases are simply unknowably true or unknowably false. Fodor can't accept Sorensen's conclusion because it doesn't involve naturalism about intentional content.

Here is where Fodor's asymmetric dependence condition runs into trouble. Suppose that one judges a man with  $N$  hairs on his head<sup>26</sup> to be bald. Vague predicates license the use of comparatives; that is, grasping the concept of baldness entails knowing that one man can be balder than another. However, one would not think of a certain man as bald (and, thus, token the Mentalese predicate "BALD") but for the fact that one would be disposed to judge that he was bald if he had somewhat fewer hairs, say 10 per cent fewer, on his head. Similarly, of course, one would not be disposed to think of men with 10 per cent fewer hairs as bald but for the fact that one would be disposed to think of men with 20 per cent fewer hairs as bald, and with 30 per cent fewer hairs, and so on. That is, in order to be thinking that a person falls under the predicate "bald", and not some other predicate, it is necessary that one be disposed to judge persons with even fewer hairs to be bald as well. Vague predicates, then, have an inherent asymmetry built into their use: in that a predicate is vague, one would not judge non-core members of a vague predicate's extension to fall under that concept but that one would judge core cases to fall under that concept. That is, one would not judge someone with severely receding hair, like John Malkovich, to be bald but that one would judge someone completely hairless (e.g. Yul Brynner) to be bald.

Furthermore, one would judge core cases to be bald even if one were not disposed to judge non-core cases to be bald. Judgments of baldness are sensitive to context. Thus, there might be circumstances in which I stopped judging John Malkovich to be bald, even while being aware of cases of men with far more hair, in that I am currently surrounded by Yul Brynner-like men whose hairs number orders of magnitude smaller than Malkovich's (let alone Tiny Tim's). Core cases cause "BALD" tokenings even if non-core cases wouldn't.

Consider, then, applying Fodor's theory to an ordinary user of "BALD" whose practice is described as follows: (1) Core bald persons cause "BALD" tokenings.

---

<sup>26</sup>Or ratio  $N$  of hairs to surface area. Note, too, that what I shall say about "bald" is true for the phenomenal predicate "seems bald", as well.

(2) Some “BALD”-tokenings are actually caused by core bald persons. (3) “BALD” tokenings caused by non-core cases of baldness are asymmetrically dependent upon those caused by core cases of baldness. That is, non-core bald men wouldn’t cause “BALD” tokenings but that core bald men would (as we said above); however, core bald men would cause “BALD” tokenings even if (in certain circumstances) non-core bald men wouldn’t. We can see then that the asymmetric dependence condition of Fodor’s theory has the effect that “BALD” denotes only the core cases of baldness, those persons that the subject always and everywhere would judge to be absolutely clear cases of baldness. This, however, gets things wrong. It is not true that “bald” is truly said of someone only if they belonged to this set of core cases. The concept of baldness has a context-relativity that allows it to embrace different extensions in different circumstances as long as the general pattern for vague concepts is followed.

Similar considerations hold for other vague concepts, such as “young”, “tall”, “red”. In addition, there are asymmetries for predicates that exhibit prototypicality effects; for example, penguins are less prototypically birds than robins. If so, it seems that penguins wouldn’t cause “BIRD” tokenings but that robins would. In that there is an asymmetry built into the extension of vague and prototypicality-exhibiting kinds, the counterfactual asymmetry by which judging peripheral members of an extension to fall under that predicate depends upon judging core members of that extension to fall under it determines the intentional content of the predicate as the most central cases. This effect doesn’t come up when considering natural kinds such as cows. One thinks of one cow as being as much of a cow as any other, so asymmetric dependence merely rules out non-cows from the extension of the predicate: non-cows wouldn’t cause “COW” tokenings but that cows do.

There are, then, two problems with naturalizing the T-sentences containing the Mentalese correlates of vague English predicates in Fodor’s way. Firstly, it gives predicates such as “BALD” a precise extension. This, however, misrepresents what it is to grasp the concept of baldness; if an account like Wright’s is correct, as it seems to be, then grasp of the concept of vagueness involves nothing more than grasping

the applicability of baldness to a range of core cases; grasping its inapplicability to certain other core cases of non-baldness; and grasping that baldness doesn't go over into non-baldness by some small change. Secondly, Fodor's asymmetric dependence condition determines makes only core cases of baldness fall within the denotation of the Mentalese predicate "BALD". This obviously gets the intentional content of Mentalese "BALD" wrong: it is not true that the sentence "Carlos is bald" is true just in case Carlos is completely or very nearly completely hairless.

Perhaps, then, Harvey's understanding algorithm is set up to flag vague English predicates and assign a Mentalese truth-condition that avoids these problems. Such a development need not violate Fodor's abstention from incorporating semantic knowledge into the workings of Harvey. Refined truth conditions would then be supplied for the sentence. Could Fodor employ one of the theories employed to make logical sense of vague predicates, i.e. a supervaluationist or degree of truth account, to make Harvey's understanding of sentences with vague predicates turn out correctly?

Supervaluationist accounts take sentences with vague predicates to be true just in case the sentence would be true on every "sharpening" of some vague predicate  $\psi$ , where a sharpened predicate is a predicate that specifies a precise cut-off for those instances now considered to fall between the definitely  $\psi$  and the definitely not- $\psi$ .<sup>27</sup> Could such an account help Fodor?

Suppose that Harvey's inner processing distinguishes vague from non-vague predicates of English and that he tokens the following T-sentence upon hearing Carmen utter "Carlos is bald":

(30) WHAT CARMEN SAID IS TRUE IFF CARLOS IS BALD<sub>0</sub> &  
CARLOS IS BALD<sub>1</sub> & CARLOS IS BALD<sub>2</sub> & ...

Here there are as many predicates "BALD<sub>n</sub>" as there are ways of sharpening the English predicate "is bald". This T-belief is true just in case Carlos falls within the

---

<sup>27</sup>For discussion, see R. M. Sainsbury, *Paradoxes*, New York: Cambridge U. Press, 1987; Section 2.2, pp. 31-40.

extension of every sharpened predicate "BALD<sub>n</sub>".

Such an account will not help Fodor because, first, each predicate "BALD<sub>n</sub>" suffers the same fate as the original predicate "BALD": each such predicate's denotation is asymmetrically dependent upon the core cases of baldness. Therefore, each conjunct in the supervaluationist reconstruction of the T-sentence says precisely the same thing: it says that "Carlos is bald" is true just in case Carlos is clearly bald. In addition, of course, the supervaluationist account swells the Mentalese lexicon enormously, perhaps infinitely. Consider how much computational space would be necessary for the supervaluationist T-sentence for "The integers are numerous" in that "is numerous" is vague.

The supervaluationist account thus seems to be of no help to Fodor. Perhaps, then, the "degree of truth" approach, as suggested by George Lakoff and others could help.<sup>28</sup> On this view, Harvey's inner processing produces a Mentalese T-sentence of the following form upon hearing Carmen's utterance:

(31) WHAT CARMEN SAID IS TRUE IFF THERE IS SOME  $n$  SUCH THAT  
CARLOS IS DEGREE(Bald,  $n$ ) &  $n > m$

Here, one introduces the Mentalese relation "DEGREE(BALD,  $n$ )" which is intended to denote being bald to a certain degree  $n$ . On this scheme, clearly bald men are taken to have the property of being-bald-to-degree-1; clearly unbald men are taken to have the property of being bald-to-degree-0. Everyone in between has the property of being bald to some degree between 0 and 1. On this scheme, something is bald just in case it is bald to at least some (mysteriously unspecified) degree.

Does the degree of truth account help Fodor? Consider the following Mentalese formula.

(32) CARLOS IS DEGREE(BALD, .8)

---

<sup>28</sup>See William G. Lycan, *Logical Form in Natural Language*, Cambridge, MA: MIT Press; 1984, pp. 62-70, for discussion.

This sentence would be tokened if the subject believed that Carlos was bald-to-degree-.8. The previous difficulty with asymmetric dependence will not arise here. It is not the case that things that have the property of being-bald-to-degree-.8 wouldn't cause tokenings of "DEGREE(BALD, .8)" unless things that had the property of being-bald-to-degree-1 (i.e., completely bald things) did; completely bald things presumably don't cause tokenings of "DEGREE(BALD, .8)" at all.

The problem with the degree of truth approach, then, isn't that it succumbs to the asymmetric dependence problem; rather, the difficulty with this account is in explaining why it is we are so much more precise and articulate in Mentalese than in our first languages. How can we suppose that our heads are nomicly attuned to the property of being-bald-to-degree-.8, yet not explain why we can't say when something falls under precisely this concept? We cannot say when a person is .8-degrees-bald, nor by how much one would have to change in order to go from being .8-degrees-bald to .81-degrees-bald or .79-degrees-bald. Nor can we say what the lowest degree of baldness is that still counts for being bald. Why not, if such concepts appear in our belief boxes and provide the truth-conditions of vague predicates? The degree of truth approach seems to require more conceptual resources than are plausibly attributed to speakers, and so, it is of no use in explaining away the asymmetric dependence problem for vague Mentalese terms.

In that the only way around the asymmetric dependence problem for vague predicates is the degree of truth method that inexplicably makes us more articulate in thought than in language, there does not seem to be any way around the asymmetric dependence problem for Mentalese terms correlated with vague English predicates. There does not seem to be any plausible account of the Mentalese truth-conditions for sentences involving vague predicates for which Fodor's theory would determine the correct intentional content.

### 2.2.3 The Actual History Condition

Fodor's actual history condition is adopted in order to rule out the property of *being water-or-twin-water* as the intentional content of an Earthling's Mentalese "WATER" predicate, since, by hypothesis, if he were to encounter some twin-water from Twin-Earth, he would be unable, modulo his present capacities, to distinguish it from water. Therefore, in that twin-water would nomically cause "WATER" tokenings in an Earthling just like earth-water, something must be done in order to prevent "WATER" in an Earthling's head from denoting the disjunction of both properties. Fodor's solution would seem to require that a Mentalese symbol denotes a property only if instances of that property have caused tokenings of that symbol.

An obvious problem with Fodor's actual history condition as a necessary condition for Mentalese content determination is this: presumably, no instances of such properties as the property of *being a unicorn* have occasioned the tokening of the Mentalese predicate "UNICORN" in a subject. Nevertheless, it can be correct to say of Mary that she believes that unicorns roam Albany and, thus, that she tokens a Mentalese formula in her belief box with this intentional content. Fodor is prepared to bite the bullet on this, however. He accepts that there is no simple Mentalese symbol that denotes the property of *being a unicorn* or any other uninstantiated property (TC, pp. 100-1). However, even if there is no Mentalese formula syntactically isomorphic to "unicorn", there can still be (composite) Mentalese formulae semantically or extensionally isomorphic to "unicorn" (e.g., descriptions or metalinguistic glosses such as "satisfiers of 'is a unicorn'") that can figure in belief attribution in the way that "UNICORN" was supposed to figure.<sup>29</sup>

This modification is necessary if Fodor is to avoid a serious problem. Suppose that

---

<sup>29</sup>This solution seems to bring a host of problems in its wake. For example, is there a difference between believing that unicorns roam Albany and believing that horse-like one-horned creatures roam Albany? Presumably, that is the sort of solution Fodor has in mind. However, if I believe that horse-like one-horned (created) *automata* roam Albany, do I thereby count as believing that unicorns roam Albany? Is there a difference between believing that unicorns roam Albany and believing that things that satisfy "is a unicorn" (in English? in a particular idiolect?) roam Albany? On the general problem of attributing mental contents with linguistic forms, see R. Stalnaker, 1990.

you and I talk about a person named Rob Kempf with whom I am acquainted and about whom everything you are destined to believe concerning him (*qua* Rob Kempf) comes from my talk about him alone. You, that is, never become acquainted with him. Moreover, there is no transitive causal relation from him to my utterances and from my utterances to your belief-tokenings. I do not speak of him, and thus cause you to token an appropriate belief, only if his presence causes me to have the relevant beliefs; my speaking of him is not nomically controlled by any of his properties. Suppose, then, that we assumed that understanding my assertion, "Rob Kempf is F", involves your tokening the following belief.

(33) WHAT THE SPEAKER SAYS IS TRUE IFF ROB KEMPF IS F

According to Fodor's theory, then, you would thereby have the belief that what I say is true iff "Rob Kempf" is F (i.e. iff his *name* is F). This unfortunate consequence follows from Fodor's theory because: it is the utterance of his name alone which nomically causes you to token "ROB KEMPF" (nomic relation condition); some utterances actually cause this tokening (actual history condition); and these tokenings are asymmetrically dependent upon nothing else (asymmetric dependence condition). In particular, these tokenings do not asymmetrically depend upon Rob-caused tokenings. That is, my utterances would cause you to token such beliefs even if Rob himself wouldn't. Although you might, for example, bump into Rob all the time without realizing that this is the person of whom I speak, the beliefs you would form about the person you bumped into wouldn't be of the form "...ROB KEMPF...". You token beliefs of the form "...ROB KEMPF...", by hypothesis, only as T-beliefs correlated with my utterances. Thus, if understanding my utterance produces a "ROB KEMPF" T-belief in you, then the intentional content determined by Fodor's theory is metalinguistic. This is true even if we interpret Fodor's theory as only providing sufficient, and not necessary, conditions.

Obviously, it isn't correct for Fodor's theory to attribute beliefs with such metalinguistic intentional content to you. A way to circumvent this is to suppose that

one's natural language understanding mechanism is set up to flag those words denoting things to which our heads are not nomically attuned, either because such things are not instantiated or because we do not have the capacity to recognize them as instances of that word. That is, if one doesn't have the capacity to recognize Rob as the referent of "Rob Kempf", then one doesn't token "ROB KEMPF" in the T-beliefs occasioned by talk about Rob. Rather, one's head is set up to token a belief with something like one of the following forms.

- (34) WHAT THE SPEAKER SAYS IS TRUE IFF THE REFERENT OF "Rob Kempf" IN L IS F.
- (35) WHAT THE SPEAKER SAYS IS TRUE IFF THE PERSON OF THE SPEAKER'S ACQUAINTANCE THAT UNIQUELY HAS THE PROPERTIES A & B & C...IS F (where "A, B, C" are attributes of Rob gleaned from my talk.)

The intentional content of the constituents of these Mentalese formulae is to be determined in the usual way. Thus, the intentional content of such beliefs would not be metalinguistic.

All of this makes it impossible, however, to view Harvey as employing only syntactic and phonological knowledge in producing T-beliefs. On Fodor's view, the Mentalese T-beliefs occasioned by, on the one hand, the English sentence "the unicorn is a mythical beast" and, on the other hand, by the English sentence "the stenographer is a mythical beast" differ depending upon whether or not one has encountered a stenographer or a unicorn during the course of one's experience. Harvey therefore needs to know more than what was said in order to generate a T-belief. He must also keep a running inventory of the world in order to generate only T-beliefs with the right content. While not strictly inconsistent with Fodor's endorsement of the Harvey model, this violates what Fodor says about the *modularity* of linguistic understanding.

Fodor's view, which he outlines in his monograph *The Modularity of Mind*, is that language processing is *modular*: i.e. language processing is not part of general intelligence but, rather, is accomplished by special mental mechanisms for computing the

linguistic properties of utterances *without accessing or employing general background knowledge*. On this view, the determination of what was said in an utterance is, to put it crudely, a kind of “reflex”.<sup>30</sup> The idea that modules are processors not affected by the content of one’s general beliefs is supposed, for example, to explain why we can’t help but perceive an optical illusion in an illusory way even when our beliefs about what we are seeing contradicts our perceptions. As such, the suggestion that the production of T-beliefs depends upon access to an inventory of the world seems to violate this conception of language processing as informationally encapsulated. If the form of the T-belief assigned to an utterance varies depending upon whether one believes one has encountered instances of the relevant properties or whether one believes that one can recognize instances of the property, this violates the informational encapsulation of language processing Fodor elsewhere endorses.

#### 2.2.4 Difficulties with the Nomic Relation Condition

Finally, the semantics of tense and modality raises serious problems for Fodor’s nomic relations condition. The nomic relation condition requires that if a symbol denoted the property of being *F*, then *F*-ish things cause tokens of Mentalese “*F*”. I shall argue that Fodor’s account is inadequate for ascribing the intentional content of tensed beliefs and modal beliefs correctly.

If Harvey understands Carmen’s utterance of the sentence “Carlos swam”, then, on a fairly simple account, he must come to have a T-belief of the form:

(36) WHAT CARMEN SAID IS TRUE IFF  $(\exists e)(\text{PAST}(e) \ \& \ \text{Swimming}(e, \text{Carlos}))$

That is, on a simple view of things, if Harvey has a tensed belief, he tokens a belief of the form “...PAST...” or “...PRESENT...” or “...FUTURE...”. That is, he tokens a T-belief with a Mentalese tense predicate. In order for Fodor’s theory to apply here,

---

<sup>30</sup>In the dedication of his monograph, Fodor reports Merrill Garrett as having made just such a remark.

Mentalese “PAST” must be sensitive to the property of *being past*. The extension of this property changes every instant; thus something is past iff it is earlier than  $t_1$  at  $t_1$ , and earlier than  $t_2$  at  $t_2$ , and so on. For the nomic correlation condition to apply here, past things must cause the subject to token “PAST”; present things must cause the subject to token “PRESENT”; future things must cause the subject to token “FUTURE”. None of these is strikingly plausible. Future things *qua* future things cause nothing in the present. Similarly, in that every cause takes time to propagate through space, present things will never cause tokenings of “PRESENT”. Further, it seems odd to suppose that occurrent experiences, presumably a link in all psychophysical causal chains, would cause tokenings of “PAST”. If I am currently presented with the visual sensations corresponding to the scene of a bird singing, let’s say, this will not cause me to believe that the event of the bird’s singing is past. If we are to have a naturalistic account of beliefs with present, past, and future tense intentional content, then, it doesn’t seem that explaining this in terms of nomic causal relations between instances of temporal properties and symbol tokenings will get one very far.

Suppose, then, that we represent the truth-conditions associated with tensed sentences without using Mentalese tense predicates but simply in terms of times and their temporal relations. We get a T-belief such as:

(37) WHAT CARMEN SAID IS TRUE IFF  $(\exists e)(\exists t')(\exists t)(\text{EARLIER THAN}(t',t) \ \& \ t=\text{TIME OF UTTERANCE} \ \& \ \text{Swimming}(e, \text{Carlos}) \ \text{IS TRUE AT } t')$ .

Here, too, we can see that a host of difficulties arise for Fodor’s physicalistic program. To dwell on one, it is not at all clear how to extend the physicalistic account by which Fodor explains that Mentalese “COW” denotes the property of *being a cow* in order to establish the necessary nomic causal relation between tokenings of “EARLIER-THAN” and the things that instantiate the earlier-than relation physically. There must be some way to do this if Fodor’s project of intentional naturalism is to go through.

One problem is: what instantiates the earlier-than relation? Ordered pairs of times? This will not do since times are acausal. I will suppose, then, that time-slices of a system (e.g. the universe) are sufficiently causal to serve as the physical basis of the instantiation of the earlier-than relation. In passing, it should be noted that the use of time-slices here raises the question of why “COW” should denote the property of being a cow and not a time-slice of cow.

In order for Fodor’s project to be completed, then, there would have to be a nomic causal relation obtaining between tokenings of “EARLIER-THAN” and ordered pairs of time-slices of the universe. That is, a physicalistic basis to the direction of time requires that there be some physical property P such that any later time-slice of the universe is more P-ish than any earlier stage. It is a matter of controversy that there is such a physically specifiable basis for the direction to time, but many believe that increases in entropy or the occurrence of certain quantum phenomena are candidates.<sup>31</sup> Even if there is such a physical basis for the direction of time, it is hardly plausible that ordinary speakers must be attuned to it on such a grand scale in that they understand tensed sentences. Moreover, it is hard to understand how one could establish a nomic causal relation between physical instances of the earlier-than relation and tokenings in a subject’s head. One can’t “rub the subject’s nose” (*Psychosemantics*, p. 115) in instances of the earlier-than relation (in, that is, ordered pairs of time-slices of the universe) in order to see what lights up in his head and thus establish a nomic relation. It is not conceptually possible for pairs of time-slices to be the sort of thing that a subject can interact with at a particular time. If it is impossible, then there is no way to apply Fodor’s nomic relation condition to determine an intentional content for Mentalese tense predicates.

Furthermore, the time-slices instantiating the earlier-than relation can’t be much more human-scaled than time-slices of the universe as a whole if they are to perform

---

<sup>31</sup>On whether there is a physical basis to the direction of time, see Paul Horwich, *Asymmetries in Time*, MIT Press; See David Layzer’s, *Cosmogogenesis*, Oxford U. Press, for a discussion of contemporary cosmology’s conviction that the universe is infinitely large and contains infinitely much matter; it is not even conceptually possible for a finite creature to be attuned to the instantiations of temporal relations by time slices of an infinitely large universe.

the function required. The earlier-than relation required as the intentional content of belief must be universal (relative to an inertial frame) if differently placed observers are to be able to have beliefs with the same intentional content; this entails that the earlier-than relation specified as the intentional content of Mentalese “EARLIER-THAN” must order all time-slices smaller than those of the whole universe according to the order of the universal time-slice to which they belong. However, if one attempted to use pairs of time-slices significantly smaller than universal time-slices as the causal correlate of earlier-than tokenings, there is a danger that the ordering of these smaller time-slices will differ from the ordering (according to the physical basis of time) of the universal time-slices. Locally, there may be a decrease in P from one local (non-universal) time-slice to the next even though, on a universal scale, there is an increase in P during that same interval. The ordering relation of local time-slices would thus not be necessarily isomorphic to the ordering of the universal time-slices, nor to orderings of time-slices of another locale. It follows, then, that two observers both attuned to local increases in property P as the physical basis of the temporal earlier-than relation, couldn’t have beliefs with the same intentional content if their local time-slices weren’t always isomorphically ordered according to increasing P. If there were ever a pair of times such that A’s local time-slices decreased in P while B’s increased, then the intentional content of their “EARLIER-THAN” symbol must differ. However, this defeats a reasonable ambition for a naturalized semantics of tense that any two subjects may be ascribed thoughts with the same temporal intentional content no matter where they are in the universe (relative to an inertial frame). Time marches along just the same for my cousin on Alpha Centauri and for me, so we should be able to have thoughts with the same temporal intentional content. This desideratum is defeated on a too local view. However, grounding temporal intentional content in too large a time-slice makes it implausible for a speaker’s head to be attuned to its P-ishness.

Thus, Fodor has yet to convince us that he has the theoretical resources to provide an adequate naturalized account of tensed or temporal intentional content. Instances of tensed properties or temporal relations *qua* instances of those properties aren’t the

sorts of things that have nomic causal relations with subjects' heads, and this was to be the basis of Fodor's theory.

A moment's reflection shows that a further example of an area of discourse that poses difficulties for intentional naturalization à la Fodor is modal discourse.

A well-known episode in the history of philosophy indicates that Fodor should not expect smooth sailing. Consider Descartes's example in the Second Meditation of his perception of a piece of wax, by which he criticized the Aristotelian account of our grasp of essential properties. Descartes argued that sensory encounters were an inadequate basis for grasping the indefinitely many ways that a piece of wax could deploy its quantity of extension. Further, the application of the imaginative faculty to sensory information was also an inadequate basis for grasping this; we can only imagine so many transformations of the wax, while grasping that indefinitely many such transformations are possible. So, Descartes concluded that information provided by the wax itself in sensory interaction with the subject couldn't by itself provide an adequate basis for our grasp of this modal property. Grasp of such a modal property required special powers of the mind.

This Cartesian episode illustrates the basic problem with naturalizing the modal within an information-based framework. Natural signs are signs of what actually cause them, but the semantics of modality requires a realm of the merely possible. The merely possible *qua* merely possible cannot cause anything; only the actual has causal properties. It is not at all clear how anything could bear information concerning what merely *could* be actual, but isn't. At most, information about the actual could be pressed into double duty, telling us about what is the case and about what could be the case as well. Like the limited powers of the Cartesian imagination to determine how many ways the perceived wax could be rearranged, this would not provide exhaustive information about the merely possible, however. As Salvador Dali remarked, so little of what could happen actually does. Thus, in that the realm of the possible far outstrips the realm of the actual, it seems implausible that information concerning what is actual could also encode the complete story of what is

possible. The moral of this is that if Mentalese T-beliefs contain the predicates “ACTUAL” and “POSSIBLE”, it is not clear how their extensions could be differentiated naturalistically.<sup>32</sup>

Since approaching modal intentional content naturalistically via Mentalese modal operators seems doomed to failure, there would seem to be no option left to Fodor but to take the English modal operators “possibly”, “necessarily”, “actually”, to be correlated with Mentalese logical operators naturalizable via their possible-world truth conditions. For example, if Carmen remarks to Harvey, “it is possible that p”, Harvey might token the T-belief:

(38) WHAT CARMEN SAID IS TRUE IFF  $(\exists W)(\text{POSSIBLE-WORLD}(W) \ \& \ P \text{ IS TRUE AT } W)$

The Fodorian naturalist is then faced with the task of naturalizing the intentional content of the predicate “POSSIBLE-WORLD” in terms of nomic causal relations. Again, a nomic causal relation between possible worlds and “POSSIBLE-WORLD” tokenings would seem hard to establish; how can one’s nose be rubbed in a possible world in order to see what lights up in one’s head? This possible world (i.e. the actual world) is too much with us. Furthermore, there would be no difference in the extensions of “ACTUAL-WORLD” and “POSSIBLE-WORLD” since all and only the same worlds would be causally correlated with these predicates. The actual world is a possible world, so it must be nomically related to tokenings of both predicates. On the other hand, any possible world would cause “POSSIBLE” tokenings only if it were actual, and, thus, it would cause “ACTUAL” tokenings as well. Thus, believing that it was possible that p would be the same thing as believing that it was actually the case that p.

As Descartes pointed out long ago, then, information concerning what is actual provides an insufficient basis for grasp of what is possible. Fodor’s intentional natu-

---

<sup>32</sup>I will be concerned only with the intentionality of possibility here. The semantics of “it is necessary that p” could presumably be given in terms of its dual: “it is not possible that it is not the case that p”.

realism requires that the causal be a sufficient basis for naturalizing modal intentional content, but what is merely possible and not actual has no causal powers *qua* merely possible.

### 2.2.5 Naturalizing Logical Consequence

Fodor supposes that the intentional content of non-denoting terms, such as the logical connectives, is given by their “inferential” or “conceptual role”.<sup>33</sup> Thus, the intentional content of Mentalese “AND” is determined by introduction and elimination rules that uniquely determine this propositional function. The conceptual role of “AND” is given as follows:

- (39) For any two Mentalese sentences  $\alpha$ ,  $\beta$ , if  $\alpha$  and  $\beta$  appear in the belief box, then so does  $[\alpha \text{ AND } \beta]$ .
- (40) If any Mentalese sentences of the form  $[\alpha \text{ AND } \beta]$  appears in the belief box, then so does  $\alpha$  and so does  $\beta$ .

Fodor supposes that adequate conceptual roles for the rest of logical connectives can be given as well.

This strategy for naturalizing the intentional content of non-denoting terms obviously presupposes that the introduction and elimination rules applying to a Mentalese connective are determinate. It is infeasible to instantiate every applicable instance of such rules, however; for example, from just two atomic Mentalese sentence, infinitely many applications of the “AND”-introduction rule above are possible. We cannot, therefore, expect finite creatures to apply the rule in every instance. If not, then less than total application of the rules must make the rules determinate, but this runs afoul of Kripke’s Wittgensteinian skepticism regarding the determinability of rules on

---

<sup>33</sup>TC, pp. 110-1

the basis of partial application.<sup>34</sup>

Consider, then, the Mentalese expression “IS TRUE IF AND ONLY IF” that figures so crucially in T-beliefs. In his presentation of Harvey, Schiffer says that “the conceptual role” of “IS TRUE IF AND ONLY IF” can be stipulated thus:

(41) If the formula  $[\alpha \text{ SAID THAT } \sigma]$  is in Harvey’s belief box, then, *ceteris paribus*, so is the formula  $[\text{WHAT } \alpha \text{ SAID IS TRUE IFF } \sigma]$

Obviously, however, this does not suffice in determining the intentional content of “IS TRUE IF AND ONLY IF”. This correlation between contents of a subject’s belief box is consistent with “IS TRUE IF AND ONLY IF” having the intended intentional content of “is true just in case the Red Sox win the pennant or” (under appropriate circumstances) or “is true just in case  $2+2=4$  or” and so on. Schiffer’s conceptual role for “IS TRUE IF AND ONLY IF” completely underdetermines the intended intentional content.

Can Fodor’s theory do better? It seems clear that Fodor should not want the intentional content of “IS TRUE IF AND ONLY IF” to be determined by some nomic head-world relation. This would require the countenancing of a physicalistically specifiable correspondence relation between sentences and their truth-conditions, e.g. between tokens of the formula “BRUTUS STABBED CAESAR” and the stabbing of Caesar by Brutus. Furthermore, it would require that the subject’s head be somehow attuned to the presence of this relation, again, in something like the way in which it is attuned to instances of *is a cow*.

On the other hand, if the intentional content of “IS TRUE IF AND ONLY IF” can’t be determined by its conceptual role as specified by introduction and elimination rules<sup>35</sup> “IS TRUE” and “IF AND ONLY IF”? The determination of the intentional content of the biconditional “IF AND ONLY IF” by means of introduction and

---

<sup>34</sup>S. Kripke, *Wittgenstein on Rules and Private Language*, Harvard U Press: 1982; also, P. Boghossian, “The Rule-Following Considerations”, *Mind*, October, 1989.

<sup>35</sup>Schiffer doesn’t discuss the obvious elimination rules for “IS TRUE IFF”.

elimination rules runs afoul of Kripkensteinian skepticism.

Furthermore, the intentional content of “IS TRUE” is also indeterminate. The following introduction and elimination rules

(42) If [A] is in one’s belief box, then so is [“A” IS TRUE]

(43) If [“A” IS TRUE] is in one’s belief box, then so is [A]

do not serve to determine whether the intentional content of “IS TRUE” is “is true” or “is true or grass is green” and so on. Thus, Fodor has not provided an account of what it is that a speaker believes in producing a T-sentence along naturalistic lines.

## 2.3 Conclusions

As we have seen, the truth-conditions of ordinary talk involves vague predicates, predicates with no actual instances, and properties for which the perception-based model of information-carrying is inapplicable. Further, if understanding sentences involves believing that they have such-and-such truth-conditions, then understanding requires the speaker to have at least the concept of truth. As such, to the extent that Fodor’s physicalistic theory of content determination doesn’t yet have theoretical resources capable of covering these cases adequately, there is no reason yet to suppose that the theory of intentionality is a naturalistic theory of the relation of speakers’ heads to the world. Further, there is no reason to suppose that semantic theory, which Fodor supposes to be derivative of the theory of intentionality, is a naturalistically specifiable theory of the relation of speakers’ heads to the world, either. Specifically, Fodor needs to produce a more sophisticated naturalistic theory of “carrying information” that gets beyond the basic case in which a subject’s cow perceptions carry information about the property of *being a cow*.

If the project of naturalizing content is merely to answer Brentano by providing *sufficient* conditions for a certain symbol to denote a property, then this project shows

nothing as to whether semantic knowledge is redundant to an account of linguistic competence. If, however, the project of naturalizing content is that of providing necessary and sufficient conditions for Mentalese T-sentences, then this project seems destined to fail without further theoretical resources. The elements of a theory of intentionality adequate to a complete account of linguistic understanding just can't be specified on the basis of the resources Fodor and other theorists provide. Moreover, the doubts raised here are intended to show that naturalistic theorists of intentionality should make some attempt to show that their theories could be extended to meet the hard cases outlined above that figure in accounts of linguistic understanding. Success with naturalizing the semantics of words like "cow" doesn't show very much concerning the general project of explaining linguistic competence.

Furthermore, as was argued in the first section, the representation and employment of truth axioms is necessary to explain how a human being, as opposed to a merely possible creature, can acquire the capacity to assign truth conditions to indefinitely many sentences of a natural language by learning, *pace* Schiffer.

## Chapter 3

# Events and Thematic Relations

In this paper, I shall be concerned with the relationship of the semantics of verbs and the metaphysics of events and event-participation. Davidson's (1967) paper inaugurated the contemporary practice in semantics of taking verbs (or some, at least) to incorporate an implicit event argument in their truth axioms in addition to their other arguments. Nevertheless, this proposal has been resisted by some on the basis of the unresolved metaphysical question of event identity. When do two ways of picking out an event pick out the same event? Or, closer to Davidson's project, when do the truth-conditions of two sentences that quantify over an event quantify over the same event? "Neo-Davidsonian" accounts of logical form take Davidson's proposal one step further and make explicit the relationship of an argument to the quantified-over event (or state) in the truth-conditions as well. I shall examine neo-Davidsonian accounts of logical form both for what they imply about the relationship of a verb's syntax to its semantics and about the relationship of the semantics of eventive verbs to the metaphysics of events.

In the first two parts of the paper, I provide arguments for incorporating events and thematic relations into a truth-conditional semantics. In the next section, I examine and criticize an argument of Terence Parsons' for the same conclusion. Along the

way, I try to clarify the distinction between arguments and non-argument expressions with special attention to the passive construction. Since the subject of an active sentence seem to appear only optionally in the corresponding passive construction, is the subject of an active sentence an argument of the verb? Finally, I examine some puzzling applications of the neo-Davidsonian theory of logical form. There are cases of pairs of sentences in which it seems clear that the pair can quantify over the same event, but the neo-Davidsonian view seems to prevent identifying the events. I argue that the conflict is only apparent once we abandon the assumption that the neo-Davidsonian truth-conditions of a verb is determined by its apparent syntactic projection.

Within Davidson's program, to understand a verb, as with any other word of a language, is to understand its contribution to the truth-conditions of expressions of that language. So, to ask the question, "what do I know when I understand a verb?" is to ask for the truth-conditional contribution of the verb to the constructions of that language in which it appears. In the case of verbs, it is to ask what sequences are such that all and only those sequences satisfy that verb. This suggests that all there is to know about the meaning of a verb is its adicity, the number of elements in the n-tuples that constitute its extension. Evidence for any further structure to the truth axiom of a verb is motivated by attempts to explain inference patterns.

What is the relation between explaining meaning and explaining inference in the semantics of a natural language? One sometimes hears the idea rejected that it is the business of semantics to account for inferences. This is perhaps true in the same sense that it is not the business of the particle physicist to account for the collision of atomic particles. Nevertheless, it is in providing a theory of what happens when atoms collide that one proceeds to form an account of subatomic structure. The motivation for positing additional internal structure to a particle comes from observing the interaction of physical particles. Similarly, the primary business of the semanticist according to the present view is to account for the truth-conditional contribution of each element of sentential structure. However, these contributions may

produce relations of entailments between (indefinitely many) sentences. Any formal theory of natural language semantics must, therefore, delimit a class of inferences judged informally to be valid on the basis of meaning alone (rather than pragmatic factors as well) and explain formally what makes them valid. Intuitions about the validity of inferences is, thus, a way of uncovering more structure for which semantic theory is responsible.

Let us reconsider the project of excavating hidden features of logical form. Suppose we accept this inference as informally valid and involving no pragmatic factors:

(1) Sport is a dog. Therefore, Sport is a mammal.

A goal of Davidsonian semantic theory is to account for this inference formally. How do we explain the validity of (1) within a formal theory of logical consequence? We have two options. On the one hand, we may suppose that (1) is an enthymeme with the premise “Every dog is a mammal” implicit. As such, (1) is logically valid only with its implicit premise made explicit.<sup>1</sup> On the other hand, we could suppose that the truth-axiom for “is a dog” is something like

(2)  $x$  satisfies “is a dog” iff  $x$  is a canine mammal.

In that case, (1) is logically valid as it stands. How do we adjudicate between these two options? Both are adequate formal explanations of the validity of the entailment, and so, either will do as a formal account of what speakers know (perhaps considered as permanent elements of their belief boxes à la Schiffer). The mere fact that adding to the truth conditional contribution of “is a dog” would explain the inference does not in itself tip the scales in its favor. In order to show that the inference is valid as it stands and not an enthymeme, one would have to show that minimal semantic

---

<sup>1</sup>I am thinking of logical or formal validity along Quinean lines such that an argument is logically or formally valid just in case all instances of the schema it embodies are valid. Further, I am thinking of analyticity in natural language as follows: a sentence is analytically true in  $L$  just in case the statement of its truth-conditions is a logical truth of the metalanguage. If (1) is considered an enthymeme, its implicit premise “all dogs are mammals” is not analytic.

competence regarding “dog” required knowing that, for every  $x$ ,  $x$  satisfies “is a dog” only if  $x$  is a mammal. This, I think, is incorrect. What it is to understand “dog” is to know that it denotes a particular kind of thing; one needn’t know how to recognize such things nor know the superordinate (or subordinate) kinds to which it is related. Knowing the simple “homophonic” truth axiom “ $x$  satisfies “is a dog” iff  $x$  is a dog” encodes this knowledge.<sup>2</sup>

At one point, at least, Davidson himself held the view that a theory of the semantics of natural language need not always explain entailments formally. Davidson writes:

It is not part of my programme to make all entailments matters of logical form. ‘ $x$  [is greater than]  $y$ ’ entails ‘ $y$  [is less than]  $x$ ’, but not as a matter of form. ‘ $x$  is a grandfather’ entails ‘ $x$  is a father’, but not as a matter of form.

(“Reply to Castañeda on Agent and Patient”, in *Essays on Actions and Events*, p. 125.) On the view I advocate, however, this task is required of a formal theory of natural language semantics. Thus, the inferences Davidson mentions must either be explained formally by the semantics of the expressions and constructions of just those sentences or they must be considered enthymematic for an entailment that is

---

<sup>2</sup>What notion of semantic competence is at issue here? In general, I take homophonic truth conditions to suffice for minimal competence in that they encode what a speaker knows about the contribution of an element to a sentence’s truth-conditions on the basis of, perhaps, nothing more than the syntactic category of that expression alone. By “homophonic” I mean that there is not more than minimal structure on the right hand side of the expression’s truth axiom. When homophonic truth axioms are not sufficient for semantic competence, the only supplementary concepts invoked are those that must be supposed to be part of a speaker’s native endowment for one reason or another. There is no good reason to thus ascribe the concept “mammal” to the speaker’s native endowment. Therefore, the informally valid entailment above must be an enthymeme. Cf. T. Burge, “Intellectual Norms and the Foundations of Mind”, *Journal of Philosophy*, vol. LXXXIII, no. 12, December, 1986, pp. 697-720. Burge there distinguishes the “conventional meaning” of a term, the fully articulated normative use of a term for a single speaker or community, from its “cognitive value”, the “minimum competence necessary for ratiocination specifiable with the term” (p. 717). I take it that knowledge of homophonic or nearly homophonic truth axiom qualifies as a word’s cognitive value since, by hypothesis, such an item of knowledge correlates an expression with an element of thought. “Responsible ratiocination” is therefore possible.

demonstrably logically valid. Such inferences are not a matter of meaning if neither course is possible.

### 3.1 Verbs and Events in Logical Form

Davidson's original proposal was to treat the truth-conditions of a sentence such as "Brutus stabbed Caesar" as involving a single three-place relation between the referent of the active sentence's subject, object, and an event. Accordingly, a truth theory for English must aim to prove theorems of the form:

(3) "Brutus stabbed Caesar" is true iff  $\exists e(\text{stabbing}(e, \text{Brutus}, \text{Caesar}))$ . (Davidsonian)

Under what I shall call the neo-Davidsonian analysis of event sentences<sup>3</sup>, a truth theory for English must prove theorems such as (ignoring tense):

(4) "Brutus stabbed Caesar" is true iff  $\exists e(\text{stabbing}(e) \ \& \ \text{Agent}(e, \text{Brutus}) \ \& \ \text{Patient}(e, \text{Caesar}))$ . (Neo-Davidsonian)

The neo-Davidsonian analysis differs from Davidson's original analysis only in that the relationship of the various arguments to the event are specified. Such a relationship is called the "thematic role" or "thematic relation" of that argument.

Let us turn, then, to arguments for a hidden event position for some verbs. The simplest proposal is to suppose that the truth-conditional contribution of "stab" is nothing more than a 2-place relation. That is, "stab" is satisfied by an ordered pair  $\langle x, y \rangle$  iff  $x$  bears the stabbing relation to  $y$ . Or, employing the relation "Val( $x, y$ )" (read " $x$  is the semantic value of expression  $y$ "). See Higginbotham, 1985.)

---

<sup>3</sup>Historically, the late Hector-Neri Castañeda was an early proponent of this sort of modification to Davidson's original view despite Davidson's reservations. See his "Comments" on Davidson in *The Logic of Decision and Action*, ed. N. Rescher, Pittsburgh: U. of Pittsburgh Press, 1967.

$$(5) \text{ Val}(\langle x,y \rangle, [{}_V \text{stab}]) \leftrightarrow \langle x,y \rangle \in \text{Stab}.$$

Most introductory logic textbooks would take such a view.

Moreover, the view that “stab” contributes a two-place relation to the truth-conditions of sentences in which it appears seems, on the face of it, to be all that is needed for a Davidsonian account of the truth-conditions, e.g. “Brutus stabbed Caesar”. A standard Davidsonian semantics proceeds recursively from some representation of the sentence’s syntactic structure to an account of the sentence’s truth-conditions. So, for example, the syntactic structure of the sentence “Brutus stabbed Caesar” might be represented:

$$(6) \text{ } [{}_S [{}_{NP} \text{Brutus}] [{}_{VP} [{}_V \text{stabbed}] [{}_{NP} \text{Caesar}]]]$$

(This structural description of the sentence is undoubtedly not in line with state-of-the-art syntactic theory since, for starters, tense is incorrectly represented. This representation is good enough for my purposes, however.) Truth-conditional axioms sufficient for deriving the truth-conditions of the sentence as a whole are then invoked for each labeled bracketing. For example, the structure  $[{}_S \text{ NP VP}]$  might be attributed the truth-conditional axiom

$$(7) \text{ Val}(\text{True}, [{}_S \text{ NP VP}]) \leftrightarrow \exists x(\text{Val}(x, \text{NP}) \ \& \ \text{Val}(x, \text{VP}))$$

This axiom states that every sentence with the structure  $[{}_S \text{ NP VP}]$  is true just in case the semantic value of the NP (its denotation) is a semantic value of the VP (a satisfier). Further axioms required for this structure are:

$$(8) \text{ Val}(x, [{}_{NP} \text{Brutus}]) \leftrightarrow x = \text{Brutus}$$

$$(9) \text{ Val}(x, [{}_{VP} \text{ V NP}]) \leftrightarrow \exists y(\text{Val}(\langle x,y \rangle, \text{V}) \ \& \ \text{Val}(y, \text{NP}))$$

$$(10) \text{ Val}(\langle x,y \rangle, [{}_V \text{stabbed}]) \leftrightarrow \text{stabbed}(x, y)$$

$$(11) \text{ Val}(x, [{}_{NP} \text{Caesar}]) \leftrightarrow x = \text{Caesar}$$

These axioms are nicely compositional in that each axiom gives the truth conditions of a labelled bracket in terms of the immediate constituents of that bracket. The compositional structure of the axioms is determined by current linguistic views concerning phrase structure. Since phrase structure is taken to involve at most binary branching, these axioms involve at most two immediate constituents for a node. From these axioms one can prove that

$$(12) \text{Val}(T, [_S [_{NP} Brutus] [_{VP} [_V stabbed] [_{NP} Caesar]]]]) \leftrightarrow \exists x \exists y (x = \text{Brutus} \ \& \ \text{stabbed}(x,y) \ \& \ y = \text{Caesar})$$

That is, “Brutus stabbed Caesar” is true just in case for some  $x$  and for some  $y$ ,  $x$  is Brutus and  $y$  is Caesar and  $x$  stabbed  $y$ . This is a materially adequate truth-condition for the sentence.

Davidson and his followers have argued, however, that there is more to the satisfaction conditions of a verb such as “stab” than that it is satisfied by elements of a certain set of ordered pairs. Davidson’s (1967) suggestion is that “stab” is satisfied by an ordered triple of an event and two individuals. What motivates this deviation from the mere provability of extensionally equivalent T-sentences as the goal of semantic theory?

Davidson’s original line of reasoning for an event argument proceeds from the observation that the learnability of language requires, on the one hand, that the truth-conditions of all sentences be given a finite basis and, on the other hand, that modification by adverbs and adverbial prepositional phrases seems to require an infinity of primitive elements in the metalanguage. That is, since indefinitely many modifiers may be appended to “John buttered the toast” (e.g. “in the bathroom”, “at midnight”,...), each modification would correspond to a primitive metalanguage relation (“buttered”, “buttered-in”, “buttered-in-at”) requiring its own truth axiom. Thus, the truth theory for that language would be unlearnable. This led Davidson to the realization that worries about an infinite number of axioms in the language’s truth-definition can be allayed by both (i) taking “action” verbs (like “butter”) to

quantify over an implicit event-position and (ii) taking adverbs and adverbial prepositional phrases to be predicates of events. Adverb-dropping entailments are thereby explained by the metalanguage rule of inference Conjunction Elimination (from “A and B” infer “A”).

Davidson’s observations do not count as a valid argument for event positions, however, without a premise added to the effect that there is no other way to meet these constraints than with the event argument. Such alternatives do exist, however. In “Events and Reification” (1985), for example, Quine mentions the possibility of an analysis along lines suggested by Ajdukiewicz whereby adverbs and adverbial prepositional phrases are treated as functors operating upon verbs to form an  $n+1$ -adic relation from an  $n$ -adic predicate or relation. Something similar could be done within a truth-conditional framework as follows.

$$(13) ((\text{Val}(X, VP) \leftrightarrow \Phi(X)) \rightarrow (\text{Val}(\langle X, y \rangle, [_{VP^*} VP \dots [_{PP} \text{Prep NP}]]) \leftrightarrow \Psi(\Phi(X), y))$$

(where  $X = x_0, x_1 \dots x_i$ , and where  $VP^*$  differs from  $VP$  only in including the prepositional phrase  $[_{PP} \text{Prep NP}]$ ). That is, if the ordered pair  $\langle \text{Brutus}, \text{Caesar} \rangle$  satisfies “stab” then the tuple  $\langle \langle \text{Brutus}, \text{Caesar} \rangle, \text{Rome} \rangle$  satisfies “stab...in”. This metalanguage is not first-order, but so what? There are independent reasons already for thinking that English and other natural languages aren’t “firstorderizable”.<sup>4</sup> Similar considerations apply to arguments for event positions from the semantics of causal statements.

Davidson, that is, argued that we must posit event positions in logical form because only treating adverbs as predicates of events allows us to handle them finitely in a first-order framework, as he supposes we must. However, with independent arguments to show that the English is not formalizable in a first-order language, Davidson’s argument doesn’t go through. Thus, a new argument for event positions

---

<sup>4</sup>See George Boolos “For Every A There is a B”, *Linguistic Inquiry*, vol. 12, no. 3, 1981, pp. 465-7. Boolos there shows that the truth-conditions of such ordinary English sentences as “For every drop of rain that falls, a flower grows” can’t be given in a first-order language (even with identity).

is required. I will argue that a certain class of verbs can be determined non-circularly that all support inferences to sentences explicitly quantifying over an event. This generalization is counterfactual supporting. I therefore suppose that the grammar simply represents verbs falling within this class as having a hidden event position.

Consider this data. There are verbs which do and verbs which don't allow the progressive form:<sup>5</sup>

(14) Harry is running home.

The dog will be attacking the mailman.

Fred was driving to the store when the rain started.

\* Bill is knowing French. (cf. Bill knows French.)

\* My car is weighing a ton. (cf. My car weighs a ton.)

\* We are having a new house. (cf. We have a new house.)

Moreover, roughly, for verbs that take the progressive form, it is valid to infer "something happened". Thus:

(15) If Harry ran home, then something happened.

If the dog attacked the mailman, then something happened.

If Fred drove to the store, then something happened.

Notice, too, that the *happen*-entailments don't go through with the non-progressivizable verbs above.

(16) If Bill knew French/my car weighs a ton/we have a house then something happened.

That is, in possible worlds in which my car has always had the same material constituents, in which Bill has always known French, and in which we have always had the house we now have, the premise is true but nothing happens. These facts

---

<sup>5</sup>Examples from E. Frawley, *Linguistic Semantics*, Lawrence Earlbaum, 1992; p. 149.

suggest that there is a correspondence represented in an English speaker's grammar between a verb's ability to take the progressive form and its semantic quantification over an event. Events unfold in time, and it is this sense of unfolding in time that the progressive expresses. It is reasonable, then, to suppose that a verb's truth-conditions involve quantification over an event only if it is progressivizable, at least to a first approximation.

I take it that the truth conditions of "something happened" involve nothing more than existential quantification over an event (ignoring tense):

(17) "Something happened" is true  $\leftrightarrow \exists e(e \text{ is an event})$

I will assume, then, that something is an event just in case it is true to say that it happens. Further, I will take the truth-conditions of a sentence to involve quantification over an event just in case it licenses a *happen*-entailment. The truth conditions associated with verbs that lack this property may involve quantification over states but not events. States do not happen; only events do.

If true, the correspondence between event-quantification and progressivizability would be counterfactual supporting. If so, we would have a compelling reason to include an event position within the truth-axiom of all progressivizable verbs. *Happen*-entailments would thus be analytic and not enthymemes lacking the premise that, for certain verbs V, if a sentence with matrix verb V is true, then there is a happening. Consider, for example, the made-up English verb "to berv". We could not consider this to be a valid inference on the basis of grammatical knowledge alone:

(18) John bervs (+s=Present) Mary. Therefore, something happened.

"To berv" might mean something along the lines of "have the same hair color as". However, if it were true that progressivizability entailed event-quantification, then if a speaker knew that "berv" took the progressive, the inference would be licensed on the basis of grammatical knowledge alone. For example:

(19) John is berving Mary. Therefore, something is happening.

We would accept this inference on the basis of nothing other than our knowledge of the semantics of progressivizable verbs and not our knowledge of the things to which they apply.

Is progressivizability both necessary and sufficient for event quantification? To establish this we would need to establish, first, that there were no verbs that took the progressive but that didn't support *happen*-entailments.<sup>6</sup> Secondly, we would need to establish that there were no verbs that supported *happen*-entailments but that weren't progressivizable. Apparent counterexamples to both theses can be produced. Nevertheless, a slightly modified version of the progressivizability hypothesis can be defended on principled grounds.

There are some verbs that allow the progressive but don't support *happen*-entailments. Consider this data (from Kearns, 1991):

(20) Your slip is showing.

An old hunting horn was hanging on the wall.

The stars were shining brightly.

A book was lying on the table. (Dowty, 1979)

None of these sentences licenses a *happen*-entailment yet they all involve progressive matrix verbs. Notice, however, that these examples fall into a certain class: as Kearns points out, in each case the sentences with the progressive are logically equivalent to sentences with the related simple non-progressive verb (modulo tense). Thus:

(21) Your slip is showing ↔ Your slip shows

An old hunting horn was hanging on the wall ↔ An old hunting horn hung on the wall

---

<sup>6</sup>Rob Stainton informs me that in Indian English, such uses of the progressive to express non-events is common. I will not pursue this here.

The stars were shining brightly ↔ The stars shone brightly

A book was lying on the table ↔ A book lay on the table. (from Dowty, 1979)

Cf: Mary is jogging ↔ Mary jogs (false)

Mary lacks enthusiasm ↔ Mary is lacking enthusiasm. (false)

Thus, we must revise our hypothesis as follows: a verb involves quantification over an event only if it takes the progressive and the progressive form is not logically equivalent to the simple form. Notice, too, that only intransitive “unaccusative” verbs exhibit this symmetry. Unaccusative verbs are those whose only argument is semantically passive rather than active; “unergative” intransitive verbs are those that have an agentive subject. For example, the semantics of “to wilt” requires that its subject passively undergo the wilting and so is unaccusative; the semantics of “to cough”, however, requires that its subject actively do the coughing and so is “unergative”.<sup>7</sup>

A further apparent class of counterexamples not ruled out by the above considerations are certain “psychological” verbs. Psychological verbs often resist the progressive:

(22)\* John is knowing the answer/believing that p.

At least, if the progressive for such verbs is possible, it often requires such special contexts as “more and more” constructions, and these don’t license *happen*-inferences:

(23) John is knowing the right answers more and more these days.

John is believing Mary/in ghosts more and more these days.

The first of these expresses that John has repeatedly known the answer to various questions recently. (Compare:

(24)\* ? John is knowing the Pythagorean theorem more and more these days.

---

<sup>7</sup>See Malka Rappaport and Beth Levin, “What to Do with  $\theta$ -Roles”, in *Syntax and Semantics 21: Thematic Relations*, 1986, and references cited there.

where the reading involving repetition is harder to get.)

However, some psychological verbs both take the progressive and fail to exhibit the progressive/simple logical equivalence pointed out above. Witness (data adapted from Kearns, 1991):

(25) I am enjoying the party  $\leftrightarrow$  \*I enjoy the party.

I was enjoying the party  $\leftrightarrow$  I enjoyed the party. (false)

I'm hearing voices  $\leftrightarrow$  I hear voices. (false on dispositional sense of "hear")

I remember my eighth birthday  $\leftrightarrow$  I am remembering my eighth birthday (false)

I regretted that I wore a tie  $\leftrightarrow$  I was regretting that I wore a tie. (false)

Cf: Dad is hoping to retire early  $\leftrightarrow$  Dad hopes to retire early (true)

Moreover, the licensing of *happen*-entailments with such psychological verbs is irregular.

(26) John believed that p. Therefore, something happened. (invalid)

John remembered his eighth birthday. Therefore, something happened. (invalid)

John was remembering his eighth birthday. Therefore, something was happening. (valid)

I am forgetting something. Therefore, something is happening. (invalid)

John perceived/saw that the butler wanted to speak. Therefore, something happened. (valid)

I regretted wearing a tie. Therefore, something happened. (invalid)

Darkness frightened the children. Therefore, something happened. (invalid)

Rimbaud's poetry preoccupied Joe. Therefore, something happened. (invalid)

This data suggests the explanation that verbs involving ascriptions of *processing* by inner psychological mechanisms, as opposed to the mere occurrence of psychological states, licenses *happen*-entailments. That is, we think of remembering (in the sense of intentionally recollecting the particulars of something), seeing, perceiving, and

so on, in terms of internal psychological processes, whereas we don't associate any particular form of inner processing with regretting, enjoying, believing, or forgetting something. According to folk psychology, the states associated with these verbs simply are instantiated; they do not happen.<sup>8</sup>

Are there verbs that license *happen*-entailments but that aren't progressivizable? According to the Vendler classification, "achievement" verbs are those that describe punctual or momentary events, e.g., "notice", "realize", "find", "touch". In that the progressive is taken to imply non-punctuality, achievement verbs would, by hypothesis, not take the progressive but still license a *happen*-entailment. Thus:

(27) Holmes noticed the bell-cord. Therefore, something happened. (valid)

A noticing is punctual. Therefore, it should not take the progressive.

(28)\* Holmes is noticing the bell-cord.

Thus, it "notice" seems to be a non-progressivizable verb that entails a *happen*-entailment.

As Kearns notes, however, punctuality is a matter of context. Kearns uses the example of commenting on a videotape in slow motion to illustrate this. Holmes' noticing of the bell-cord may be punctual in ordinary time, but in reviewing a videotape of a Sherlock Holmes movie in slow motion to prove some point about its plot, one might say "See! Here Holmes is noticing the bell-cord" or "See! Here he's touching the glass" of events that would be punctual in ordinary time. Achievement verbs or verbs that denote punctual events in ordinary time are therefore not counterexamples to the thesis.

Our thesis about event quantification and progressivizability must be modified in light of this data. The correct generalization is that, for nonpsychological verbs,

---

<sup>8</sup>A further minor category of exceptions here is the "prospective" use of "have" as in "We are having a baby" and "I am having a party". I will ignore these.

excepting those unaccusative verbs for which the progressive is logically equivalent to the non-progressive, the truth-conditions associated with a verb involve quantification over an event just in case the verb is progressivizable; for psychological verbs, progressivizability entails event-quantification if the verb is thought of as implying inner psychological mechanisms. This is counterfactual supporting. Consider this entailment with the non-specified verb V in the progressive (“V-ing”):

(29) Bill was V-ing the couscous with a spoon. Therefore, something happened.

(30) Alex saw Bill V-ing the couscous. Therefore, Alex saw something happening.

This verb can't be a psychological verb since psychological verbs don't involve the instrumental use of spoons. One can't regret something with a spoon, for example. Nor are psychological processes visible in the required sense. Moreover, unlike the unaccusative verbs whose progressive and non-progressive forms were logically equivalent, this verb is transitive. This shows that if a verb were known to fall within the required class, one would know that its truth-conditions involved event quantification even without the conditions under which such sentences were true.

Strange goes on! Davidson's original argument for event quantification didn't involve inferences to statements that explicitly quantified over events, even though the very first sentence of Davidson's paper is such an explicit quantification. That is, what can “strange goings on” express except that there are goings-on, and they are strange? To say that something happened or occurred or went on is to quantify over an event. Therefore, to determine which verbs do and don't support such inferences seems the most direct way of arguing for an implicit event argument. Arguing for event quantification on the basis of adverb-dropping inferences or the semantics of singular causal statements is overly indirect. Further, once the requirement of a first-order metalanguage is given up, these arguments are no longer compelling.

## 3.2 Arguing for Thematic Roles in Logical Form

The argument for thematic relations (e.g. “Agent” and “Patient” in (4)) in a truth-conditional semantics is only partially based on the argument for event positions. One argues for thematic relations as part of the grammar by pointing out that there is more to one’s grammatical knowledge about a verb than knowledge of its adicity. That is, one knows more about “stab” than that it is satisfied by certain ordered pairs. One knows at least that “Brutus stabbed Caesar” is true iff Brutus was the agent and Caesar the patient of the stabbing. The argument positions, syntactically specified, have a definite significance. “Brutus stabbed Caesar” is not true of a situation in which Caesar did the stabbing and Brutus was stabbed. Positing

(31)  $\text{Val}(\langle x,y \rangle, \text{stab}) \leftrightarrow x \text{ stabs } y$

as the truth-axiom for “stab” is unilluminating because the metalanguage here is nothing but English itself. Despite the virtues Davidson sees in disquotational or homophonic accounts of semantics, if one is interested in exposing hidden elements of semantic structure, a disquotational account of semantics will be inadequate for the job. The hidden structure will never come to light.

Arguing for thematic roles is independent of arguing for event positions because non-eventive verbs have thematic roles, too. Consider, for example, the stative verb “to know”. “Know” doesn’t meet the criteria for event quantification given above. However, its arguments have a definite significance: “to know” requires a knowing subject and a known object. The subject and object of this relation play similar roles to the subject and the object of “believe”. Therefore, one might cross-classify the arguments of these verbs as, perhaps, the Intentional-Subject and Intentional-Object of the verbs rather than merely as “knower”, “thing known”, “believer” and “thing believed”

The thematic role of an argument is nothing other than the semantic significance of an argument position considered independently of the particular kind of event or

state that the verb denotes. The licensing of certain kinds of non-specific entailments by a variety of verbs is evidence for such semantic roles. For example, consider these *do*-entailments:

- (32) John kissed/betrayed/promoted Mary. Therefore, John did something.  
John put the groceries into the car. Therefore, John did something.  
John sold the car to the man for a nickel. Therefore, John did something.  
John walked/sang/studied/passed out. Therefore, John did something.  
Mary was fired/promoted/hired by John. Therefore, John did something.  
The beets stained the shirt. Therefore, the beets did something.

The verbs in the premises denote quite different kinds of events, but in all such cases, only one argument is licensed as the subject of the entailed *do*-sentence. The examples do not entail “Mary/the groceries/the car/the shirt did something.”

Linguistic data further suggests that the number and nature of thematic roles considered as the significance of an argument position independently of the event or state denoted is essentially invariant across languages. The thematic roles of arguments do not vary as idiosyncratically as the kinds of events or states picked out by verbs. A small number of thematic roles should be sufficient for partitioning all of the arguments of all verbs. Consider this representative catalogue of basic thematic roles compiled by David Dowty (Dowty, 1989 and 1991 contain extensive surveys of thematic roles):

- (33) (i) *Agent*: a participant which the meaning of the verb specifies as doing or causing something, possibly intentionally.  
(ii) *Patient*: a participant which the verb characterizes as having something happen to it, and being affected by what happens to it.  
(iii) *Experiencer*: a participant who is characterized as aware of something.  
(Perhaps better: a participant characterized by their intentional state.)

- (iv) *Theme*: a participant which is characterized as changing its position or condition, or as being in a state or condition.
- (v) *Source*: object from which motion proceeds.
- (vi) *Goal*: object to which motion proceeds.

This catalogue doesn't partition all arguments into one of these categories, but it is an aim of descriptive semantics to do so. It might take a while to reach reflective equilibrium in formulating a catalogue to partition not only a single language's arguments but those of all natural language. Note, for example, that in the above catalogue it is hard to see what difference there is in being a Patient and being a Theme.

According to this catalogue of thematic roles, the *do*-entailments above can be explained formally if one supposes that the truth-conditions of all of these premises includes an Agent thematic relation: e.g.  $\exists e(\dots \text{Agent}(e, \text{John})\dots)$ , and also that the truth-conditions of "x did something" are given by

$$(34) \text{ "x did something" is true} \leftrightarrow \exists x \exists e (\text{Agent}(x, e))$$

Conjunction elimination is then all that is needed to explain the entailment formally. Of course, one might motivate a finer-grained catalogue of thematic roles such that the subject of, e.g., "stain" is considered an "Actor" rather than an "Agent" (like the subject of "walk") in that being an "Agent" differs from being an "Actor" in requiring intentional behavior. In such a case, the truth-conditions of "x did something" might be given as:

$$(35) \text{ "x did something" is true} \leftrightarrow \exists x \exists e (\text{Agent}(x, e) \vee \text{Actor}(x, e))$$

Either approach would be consistent with the neo-Davidsonian approach. However, if one followed a purely Davidsonian approach (as in (3)), then *do*-entailments would have to be considered enthymemes.

Why should the semantic significance of argument positions be constrained to fall within a certain range? Why couldn't there be, for example, a verb such that

only entailments not expressible in terms of thematic roles are licensed for one of its arguments? For example, why couldn't there be a verb V such that only this entailment is licensed

(36) ...argument<sub>i</sub> V-ed.... Therefore, argument<sub>i</sub> is F

where F isn't expressible in terms of the thematic roles that partition the rest of the language's verbal arguments? Imagine, for example, that there were a 3-place verb "kisk" such that:

(37) For every x,y,z, "x kisked y z" is true iff x kissed y and z is allergic to peaches.

Would this verb fall outside of the neo-Davidsonian paradigm? The only entailment licensed for the third argument would be:

(38) John<sub>i</sub> kisked Mary Bill. Therefore, Bill is allergic to peaches.

In not participating in the kissing, "Bill" would have no thematic role associated with event participation. Nevertheless, Bill could be considered the Theme of the state of being allergic to peaches. Thus, it follows that any catalogue of thematic roles will be small relative to the size of the verbal lexicon because the truth-conditions of verbs are given in terms of events and states in which their arguments figure as participants or otherwise realize the event or state more or less directly. There are, for example, no verbs such that we infer "x Verbs y" is true iff the second cousin of x's daughter does something to y. There is only a small number of thematic roles because there are only a few ways in which something can directly participate in or otherwise realize an event or state independently of the particular kind of event or state denoted. In any case, we employ limited conceptual resources in thinking about how event and state participation or realization, and this is encoded in the verbal lexicon.

There is empirical evidence for this from learning studies. Steven Pinker,<sup>9</sup> for example, emphasizes that expectations about the syntactic positions of arguments

---

<sup>9</sup>See Pinker, 1989, 1990 and Gropen *et al*, 1991.

according to their thematic roles facilitates the learning of verbs. There is evidence, for example, that if children know that a verb involves an Agent and a Patient, then they expect the Agent to be in the active subject position. We have not lost this expectation; thus, we find this inference valid:

(39) Glooping involves doing something to something. Mary glooped John. Therefore, Mary did something (Agent), and John underwent something (Patient).

Logically, this needn't be so; the reverse is also logically possible. The agent argument might logically be realized in the object position and the patient argument might be realized in the subject position. However, the grammar makes only one of these logical possibilities available. This suggests that expectations about where thematic roles will appear syntactically does facilitate learning the verbal lexicon. Empirical work on the expected syntactic hierarchy of thematic roles suggests that significant cross-linguistic generalizations can be made (see Jackendoff, 1972).

Might not the inferences that are supposed to be formalized by thematic roles be enthymemes lacking an additional known premise about the role played by an argument? Why should we not suppose, that is, that in addition to knowing that, e.g., "stab" is dyadic, speakers also believe

(40) If "x stab y" is true, then x is an agent/does something/acts intentionally

Such a view suggests that to have a minimal understanding of "stab" it is enough to know that it is satisfied by a certain set of ordered pairs; knowledge about thematic roles is supplementary to this. However, if such a view were right, one would suppose that a speaker would have no prior expectations about the thematic role distribution of arguments in a verb, and, as we have seen this is false.

Similar considerations explain knowledge of alternations such as these: If one knows that a verb involves an agent, a patient, and a goal, then one expects the verb to pattern like give:

(41) John gave the book to Mary.

John gave Mary the book (“Mary” is Goal)

Verbs with agents, patients, and instruments pattern like “cut”.

(42) John cut the bread with a knife.

The knife cut the bread.

The bread was cut (by John) (with a knife). (“the knife” is Instrument)

Verbs with agents, patients, and benefactives pattern like “made”.

(43) John made Mary a sandwich.

John made a sandwich for Mary (“Mary” is Benefactive).

Similarly, speakers know that if a dyadic verb has an agent and a patient or theme, then the verb passivizes. That is, the subject of a passive construction must be affected in some way. This “affectedness” criterion is widely attested. Witness:

(44) John kissed/greeted/fired Mary.

Mary was kissed/greeted/fired (by John).

(45) The coat cost fifteen dollars.

\*Fifteen dollars was cost (by the coat).

(46) John resembled Bill.

\*Bill was resembled (by John).

In (44) Mary is affected in various ways, but the fifteen dollars of (45) and Bill in (46) are not. The ungrammaticality of the passives in (45) and (46) is evidence that speakers know more about verbs than their adicity.<sup>10</sup>

---

<sup>10</sup>I will show below, however, that the thesis that a verb passivizes only if the referent of its direct object bears the thematic relation “Theme” or “Patient” cannot be maintained. Surfacing as a Patient or Theme is sufficient but not necessary for passivization.

The standard neo-Davidsonian view involves two claims, then: (i) truth-conditions for sentences involving a certain class of verbs involves quantification over an event; and (ii) speakers associate with each argument position a thematic role (or theta role or thematic relation) which holds between the referent of that argument and the event (or, possibly, state) in which that referent participates. A thematic role shows how it is that the argument's referent participates in the event.

### 3.3 Parsons' Dream Machine Argument

In Chapter 5 of his *Events in the Semantics of English*, Terence Parsons offers an argument for the correctness of the neo-Davidsonian treatment of truth-conditions (4) over the Davidsonian treatment (3). The argument turns on determining how many positions of the metalanguage predicate associated with a verb must be taken to be quantified over in any use of the verb. Davidson and Parsons both distinguish arguments from non-arguments on the basis of their obligatoriness. The number of syntactic elements required by a predicate for well-formedness is its syntactic adicity; the number of variable places over that must be quantified over in the truth axiom of an element is its semantic adicity. Parsons employs such notions to determine when expressions should be represented as arguments of a metalanguage predicate associated with a verb and when they should be represented as arguments of conjoined thematic relation predicates. Parsons proceeds by considering how many things must be quantified over in using the verb "stab" in order to report bizarre dreams without contradicting oneself.

Parsons argues as follows:

In trying to describe...a dream, I may say "In a dream last night, I was stabbed, although in fact nobody had stabbed me, and I wasn't stabbed with anything." I do *not* mean this to be a report that...the stabbing had taken place earlier than the events in the dream... Such a report raises

no interesting issues at all. I mean this to be a report of an incoherent dream, one in which, say, I am bewildered by the fact that I have been stabbed but not by anyone or anything. Such testimony should *not* be analyzed as containing an explicit contradiction, as in

(47) I was stabbed, but not by anybody  $\leftrightarrow \exists e(e \text{ is a stabbing of me by somebody} \ \& \ e \text{ is not by anybody})$

In my report I use an “agentless passive”, a construction in which the agent role is unoccupied. Any analysis that attempts to analyze this example by existentially quantifying over an agent role will be wrong; it will attribute to me what I do not intend. I have not said anything from which it should be inferred that in the dream I was stabbed by somebody. My dream may have been incoherent, but I am not, and what I am saying should not contain a self-contradictory logical form.

We might, however, ask why his report shouldn't be given a self-contradictory logical form? It is not true that we don't report dreams with contradictions. After all, doesn't one report dreams with such obvious contradictions as “The old monk in my dream both was and wasn't Bob Dole”? Here an explicit contradiction is and, presumably, is meant to be expressed. The incoherent nature of the dream sometimes calls for this. On the other hand, perhaps Parsons believes that one can't have expressed, or be attributed the expression of, a contradiction unless one intends to express one, but this is obviously false. One expresses a contradiction by denying something logically equivalent to something one has affirmed. One can therefore express a contradiction without intending to if one fails to recognize that something one denies is logically equivalent to something one affirmed. The issue, then, is whether or not Parsons has expressed a contradiction, whatever his intentions were.

Parsons intends to use such dream reports as the one above to argue that all metalanguage predicates are one-place predicates of events. I reconstruct his argument as follows: (i) The metalanguage predicate associated with a verb must have only as

many arguments as entities to which one is committed to quantify over in using the verb. But (ii), dream reports and other reports can include well-formed uses of verbs which commit one to quantifying over fewer entities than reports of actual events without being self-contradictory. Therefore, (iii) only the event position is an argument of the metalanguage predicate associated with an eventive verb; all the other arguments are optional and so should be represented in thematic relation conjuncts.

Parsons employs the English verb “to stab” as an example. “Stab”, according to Parsons’ first argument, need denote no more than a dyadic predicate with an event argument and a Patient argument because one can truly report a dream in which one was stabbed but by no one, with nothing and with no other arguments specified by means of the sentence:

(48) I was stabbed but by no one and with nothing.

This dream report is intended to convey that the agent argument is optional. Moreover, Parsons also supposes that quantification over the patient of the stabbing is optional because he believes it is grammatical to say:

(49) Brutus stabbed

He thinks that the truth-conditions of (49) do not quantify over something that was stabbed. Parsons take this sentence to be true just in case Brutus made as if to stab something or someone but missed. Thus, since Parsons supposes that one need not be committed to quantifying over an agent or patient or instrument in reporting a stabbing, then the metalanguage predicate “stab” must be considered a one-place predicate of events alone. I think Parsons is incorrect in taking (49) to be grammatical. In any case, no general fact follows from this example.

In that Parsons supposes neither the agent nor the patient of the stabbing to be required syntactically (since syntactic representation entails semantic quantification except in the cases of pleonastic elements “there” and “it”), it would follow that a

stabbing event could be meaningfully reported with “stab” as the matrix verb of a sentence with no overt arguments. Although there are meaningful, noncontradictory sentences with no overt arguments, “stab” is not one of them. Consider:

(50) It is raining. ( $\exists e(\text{raining}(e) \ \& \ \text{Present}(e))$ )

It snowed. ( $\exists e(\text{snowing}(e) \ \& \ \text{Past}(e))$ )

But not:

(51)\* It is stabbing. ( $\exists e(\text{stabbing}(e) \ \& \ \text{Present}(e))$ )

\* There stabbed. ( $\exists e(\text{stabbing}(e) \ \& \ \text{Present}(e))$ )

where “it” and “there” in (51) are pleonastic, non-referential elements. It is not open to Parsons to consider some complicated scheme in which one argument is optional only if another argument is represented. This would not argue for metalanguage *stabbing* as a one-place predicate of an event.

Perhaps what Parsons has in mind, however, is a dream report such as

(52) There was a stabbing, but no one stabbed anyone, and nothing stabbed anyone.

Again, this is not intended to convey that the stabbing occurred before I became aware of the course of events in the dream and that the participants exist but are unknown to me. However, the report (52) indicates nothing about the adicity of English “stab” since “stabbing” is not a matrix verb here but a noun. There is no contradiction here because the noun “stabbing” is presumed not to require any object-language arguments at all. Nothing follows about the adicity of the metalanguage predicate correlated with the verb from the adicity of the metalanguage predicate correlated with its related gerundive noun. At least, Parsons hasn’t provided a theory linking them.

In any case, what would one be reporting when one reported such a dream by (52)? How does the “unreal world” Parsons describes by this dream report differ

from one related by a dream report of a participantless kissing event, or of a participantless walking event? All such participantless event descriptions would seem to be true if any was. The truth-conditions of such sentences would always and everywhere be satisfied. Moreover, talk of such participantless events would allow one to say, truly, that every actual spatio-temporal region contained countless participantless events, such as agentless walkings and agent- and patientless kissings, in addition to the events and states involving real things occurring or instantiated there. This is absurd. Because we don't judge it true that there are participantless stabbings in every spatio-temporal region, such uses of "stab" as in (51) are ruled out. "Stab" requires two referential arguments syntactically, and thus, the truth axiom for "stab" will require quantification over at least two places. In addition, "stab" requires that an event position be quantified over. Stabbings may require, metaphysically, an instrument; but this need not be represented in the truth axiom of "stab". The number of necessary features or participants involved in an event of a certain kind must be distinguished from both the adicity of the object-language predicate and the metalanguage predicates of events and thematic relations that provide its truth-conditions. Learning all of the necessary features of events of a certain kind goes beyond learning the truth-axiom for the verb with which one reports such an event.

Parsons has thus not provided a successful general argument for the neo-Davidsonian treatment of truth-conditions on the basis of this example. Some of the syntactic premises upon which Parsons bases his argument are simply false. Neither the patient nor agent argument of "stab" is syntactically optional. Thus, the number of things quantified over in the truth-axiom of "stab" must not be less than two, since no referential NP of a sentence contributes vacuously to its truth-conditions.<sup>11</sup> Therefore, Parsons's dream report does express a logical contradiction even if it doesn't seem contradictory on the surface. Something must account for the incoherence of the dream, after all. In that Parsons' argument fails, the neo-Davidsonian treatment of truth-conditions should be defended by arguing that neo-Davidsonian truth-

---

<sup>11</sup>Cf. Chomsky's discussion of the Principle of Full Interpretation in 1986.

conditions more completely represents what the speaker knows about the various expressions and constructions of the language.

It is clear that the agent argument of “stab” is not merely optional semantically. This argument place can’t simply be replaced by a pleonastic element *salva congruitate*, thereby reducing the number of entities quantified over in its truth-conditions.

(53) Brutus stabbed Caesar

\*It (pleonastic) stabbed Caesar.

(Cf. Caesar was stabbed.)

However, if “stab” requires two non-pleonastic arguments, then what happens to the Agent argument in the passive construction? That is, we assume that the same verb “stab” appears in both active and passive constructions. We further assume that the argument structure of a verb doesn’t vary with context (Chomsky’s Projection Principle); this is a simplifying assumption perhaps motivated by the fact that it simplifies language learning if true. Therefore, since “stab” requires both an agent and patient arguments syntactically in the active construction, it must do the same in the passive construction. What, then, discharges the agent argument in a passive construction?

Saying that the agent argument is “implicitly represented” in the syntax doesn’t get us very far, though. This notion of implicit representation is obscure. If it is syntactically represented, where does it appear in the syntax? If there is no answer to this question, then in what sense is the argument represented at all? Nevertheless, appeal to the implicitness of the agent argument is sometimes invoked as an explanation of the phenomenon of “control”, e.g.. the phenomenon of understood reference of the unpronounced subject *PRO* of infinitival phrases. (That there must be a subject for infinitival phrases is itself a consequence of the Projection Principle.) For example, in the sentences

(54) The boat was sunk [PRO to collect the insurance].

\*The boat sank [PRO to collect the insurance].

we can understand the agent of the sinking in the first sentence to be coreferential with the subject of the purpose clause “PRO to collect the insurance”; the second sentence is ill-formed because it lacks just such an argument. This is frequently taken to show that the Agent argument must be somehow represented in the syntax in order to explain the understood coreference without, however, saying where it is represented.

This is not an especially convincing argument for the implicit representation of the agent argument, however, since it is not clear that control works this way. Howard Lasnik (1988) points out that if this account of control were right, then we should find this sentence grammatical, but we don't:

(55)\*The boat was sunk [PRO to become a hero].

(Cf. John<sub>i</sub> sank the boat [PRO<sub>i</sub> to become a hero],

meaning that John sank the boat in order to become a hero.)

Since we don't, Lasnik suggests that the controller is the event of the sinking. That is, we should understand the sentence as meaning that the purpose of the *sinking* event was for collecting the insurance, rather than that the agent of the sinking did it in order to collect the insurance. At most, this example of a control phenomenon argues for the syntactic representation of the event argument, not the agent argument, in order to explain under what configurations event control occurs.

Instead of being implicitly represented, one might suppose that the agent argument of “stab” and the like is discharged by an explicit element of the syntactic structure. Indeed, in a treatment representative of current theory, Baker, Johnson, and Roberts (1989, extending Osvaldo Jaeggli's work in Jaeggli (1986)), suggest just that: the passive construction involves the discharge of the “external” argument<sup>12</sup> by

---

<sup>12</sup>So called because it appears outside of the VP projected by the verb.

the passive morpheme *-en*, which affixes to the verb in the passive construction. The direct object is then forced to occupy the subject position for syntactic reasons (in order to receive Case). Thus, BJR would analyze the passive construction “Mary was kissed” as follows:

(56) [<sub>IP</sub> [<sub>NP</sub> Mary<sub>j</sub>] [<sub>I'</sub> [<sub>I</sub> was] [<sub>VP</sub> [<sub>V</sub> kiss+en<sub>i</sub>] [<sub>VP</sub> [<sub>NP</sub> t<sub>j</sub>]]]]]]  
 Mary<sub>j</sub> was kiss+en<sub>i</sub> (=kissed) t<sub>j</sub>  
 ↑—————↓

Here the internal (or “direct”) argument, “Mary”, has been raised to subject position, and the passive morpheme *-en*, generated in I, has been demoted to affix to the verb and assigned the role of the external argument, here the Agent.

On the face of it, BJR’s suggestion seems rather ad hoc: in passive constructions, an element disappears (the external argument) and another element appears (the passive morpheme), so everything works out if we make the introduced element discharge the duties of the lost element. However, what does it mean to say that the passive morpheme is an argument. What semantic significance does it have, if any? If *-en* is an argument, would the sentence “Mary was kissed” be given the following truth-conditions?

(57)  $\exists e(\text{kissing}(e) \ \& \ \text{Agent}(e, \text{-en}) \ \& \ \text{Patient}(e, \text{Mary}))$

If so, how is “Agent(*e*, *-en*)” to be interpreted? What does the affix denote such that it can be an Agent, a concrete participant? Indeed, perhaps the passive morpheme is only an argument *syntactically* but, like a pleonastic element, it does not invoke quantification over an entity.

Further evidence for the non-pleonastic semantics of *-en* is given by a further element of BJR’s analysis. BJR point out that binding considerations explain the ill-formedness of


(58)\* He was kissed by himself

(Cf. He kissed himself  
 He was kissed by HIMSELF.)

just in case *-en* is given a referential index, as in

(59)\* He<sub>i</sub> was kissed (kiss+en<sub>i</sub>) t<sub>i</sub> by himself<sub>i</sub>

The structure (59) is an example of the violation of syntactic binding conditions called “Strong Crossover” which rules out structures of the form

(60) X<sub>i</sub> Y<sub>i</sub> t<sub>i</sub>  


where X<sub>i</sub> c-commands Y<sub>i</sub> and Y<sub>i</sub> c-commands t<sub>i</sub>; and t<sub>i</sub> is the trace left by the movement of X<sub>i</sub> to a position above Y<sub>i</sub>. Thus, assigning the passive morpheme a referential index *i* allows us to suppose that the Agent of a passive sentence’s truth-conditions is the *i*-th element of the sequences that satisfy the sentence. This is the semantic significance of coindexation.<sup>13</sup>

In any case, the suggestion here is that *syntactically*, the passive morpheme discharges an argument position, as shown by this data.

---

<sup>13</sup>Critics of BJR’s proposal argue deny, however, that the illformedness of (59) can only be explained as a Strong Crossover violation. Some claim, for example, that the simplest explanation is that reflexives are ruled out from *by*-phrases across the board. (This response was brought to my attention by Danny Fox.) Consider, for example, these nominal constructions:

- (61) John’s portrayal of himself as the wronged party was met with skepticism.  
 \* John’s portrayal by himself as the wronged party was met with skepticism.

The explanation of this data would seem to be that the genitive in a nominal phrase such as these and *John’s firing of himself*\* *John’s firing by himself* is interpreted in terms of the thematic role of the external argument as is the object of the *by*-phrase, so the bad examples violate some principle excluding expressing a thematic role in both of these ways. In any case, it is too strong to say that reflexives are ruled out from *by*-phrases across the board since the example above, *John was kissed by HIMSELF*, is fine with intonational stress. Also fine are examples with the focus adverbs *even* and *only*:

- (62) Mary has only ever been admired by herself.  
 (63) Many parodied Hemingway, but Truman Capote was even parodied by himself.

(Shawn Kern pointed out these examples.) Detailing the mechanisms of focus at work here is beyond the scope of this paper. I will simply assume that focus alters the scope relations of the elements sufficiently to avoid the violations of unfocused reflexives in *by*-phrases.

(64)\* Mary kiss

\* Mary was kiss

Mary was kissed

Furthermore, the truth axiom of verbs such as “stab” and “kiss” involves quantification over both an Agent and a Patient, since, as noted above, a non-referential element cannot discharge either argument.

(65)\* Mary kisses it (pleonastic)

\* It (pleonastic) kissed Mary

Thus, the truth axiom for a verb such as “stab” is given along these lines:

(66)  $\text{Val}(\langle x,y,e \rangle, [{}_V \text{stab}]) \leftrightarrow (\text{stabbing}(e) \ \& \ \text{Agent}(e,x) \ \& \ \text{Patient}(e, y))$

All of the variable places in the truth-axiom must be quantified over if the truth-conditions are to be well-formed. Baker, Johnson, and Roberts’ proposal, then, comes down to thinking of the passive morpheme as being a sort of clitic of the verb expressing an unspecified participant. Thus, “Mary was kissed” is to be glossed as “Mary was somebody-kissed” (and with the *by*-phrase: “Mary was somebody-kissed by John”) rather than as “Mary was the patient of a kissing which may or may not have had an agent”.

All in all, the spirit of BJR’s proposal for the passive seems more satisfactory than Parsons’. Incorporating BJR’s view into the neo-Davidsonian framework, there is one verb “stab” which semantically requires three arguments: an Agent, a Patient, and an event. When the Agent is not discharged overtly in the passive sentence, it is because the discharge of the Agent role is brought about by the passive morpheme. Discharge of the Agent role means that something is quantified over as Agent in the truth-conditions of the sentence. Therefore, the passive entails a sentence which overtly quantifies over the external argument. That is, it follows on this approach that “Caesar was stabbed” entails “Someone stabbed Caesar” analytically.

Parsons, on the other hand, considers arguments not visible at the surface of a sentence to be really absent. Thus, Parsons considers the truth-conditions of the sentence “I was stabbed, but I was stabbed by nobody and with nothing” to be

$$(67) (\exists e)(\text{stabbing}(e) \ \& \ \text{Patient}(e, I) \ \& \ \neg(\exists x)\text{Agent}(e, x) \ \& \ \neg(\exists y)(\text{Instrument}(e, y)))$$

So, on Parsons’ view, it is not analytic that “I was stabbed” entails “someone stabbed me”. However, as we have said, the view that the passive simply involves quantification over one less argument than the active cannot be maintained.

Thus, if we assume with BJR that the same verb “stab” appears in both “Brutus stabbed Caesar” and “Caesar was stabbed” and if we assume that verbs have their adicities essentially, then we can assign the affixation of the passive morpheme the following truth-axiom:

$$(68) (\text{Val}(\langle x_0, x_1 \dots x_n, e \rangle, \text{Verb}) \leftrightarrow (V\text{-ing}(e) \ \& \ \theta_0(e, x_0) \dots \theta_n(e, x_n))) \rightarrow \\ ((\text{Val}(\langle x_0, x_1 \dots x_n \rangle, \text{Verb} + \text{-en}_i) \leftrightarrow (V\text{-ing}(e) \ \& \ \theta_0(e, -\text{en}_i) \dots \theta_n(e, x_n))) \ [\& \ \text{By}(e, \text{NP}_i)])$$

(This assumes that the passive morpheme remains affixed to the verb at LF.) Here the external argument of the active verb is assigned the index of the passive morpheme, and may be “doubled” by the appending of a *by*-phrase with a coindexed object. For example, the truth-conditions of “Caesar was stabbed” and “Caesar was stabbed by Brutus” are given in this framework by:

$$(69) \exists e(\text{stabbing}(e) \ \& \ \text{Agent}(e, -\text{en}_i) \ \& \ \text{Patient}(e, \text{Caesar}))$$

$$(70) \exists e(\text{stabbing}(e) \ \& \ \text{Agent}(e, -\text{en}_i) \ \& \ \text{Patient}(e, \text{Caesar}) \ \& \ \text{By}(e, \text{Brutus}_i))$$

These truth-conditions clearly entail those of “Someone stabbed Caesar” by nothing more than Conjunction Elimination. Furthermore, the “long passive” (70) with the *by*-phrase entails a sentence indicative of the thematic role of active subject (here,

“Brutus did something”, perhaps) because the object of the *by*-phrase is coindexed with the passive morpheme discharging the Agent thematic relation.

Furthermore, Parsons takes the truth-conditions of “Brutus stabbed Caesar” and “Caesar was stabbed by Brutus” to be formally identical, thus assuming that “by” encodes the Agent thematic relation. This can’t be correct, however. The preposition “by” must not express Agency since arguments bearing other thematic roles can appear in the *by*-phrase of a passive sentence as long as that argument is the external argument of the active sentence. Thus we have:

(71) Mary was loathed by John. (Experiencer)

The intersection was approached by five cars. (? Theme=change of place)

The dot was enclosed by a circle. (?)

The crate was received by the firm. (Goal)

The package was sent by the firm. (Source)

Cf. John had the firm send the crate to the firm.

The butter was spread by a knife. (Instrument)

Cf. John spread the butter with the knife. (Adapted from Marantz, 1984).

The idea here is that although “the knife” and “the firm” play pseudo-agentive roles in these situations, there are sentences describing the situation more fully in which the agent role is more clearly occupied by something else. It is assumed that no two distinct things can bear the same thematic role. If “by” does not express Agency, then it is not clear how Parsons will explain how “Caesar was stabbed by Brutus” entails either “Brutus stabbed Caesar” or “Brutus did something”.

The thesis that the external argument is discharged in a passive construction just in case the passive morpheme is affixed to the verb supports the prediction that for any made-up passive construction, e.g., “John was glooped (gloop+en)”, speakers will accept the entailment of “Therefore, someone glooped John”. This seems to be right. Therefore, despite the passive construction, we can say that the number of grammatically necessary expressions in the active represents the real adicity of the

verb.

What we have seen, then, is that the syntactic adicity of a verb is independent of the number of metaphysically necessary participants in that kind of event. Parsons' suggestion that determining the optional expressions in reports of real or unreal events of a certain kind in order to determine the adicity of a verb is misguided. The adicity of a verb is a syntactic matter determined by grammaticality judgments about active sentences. Reflections on whether we could coherently report a bizarre dream using the verb are not relevant.

Before moving on to the next topic, it is worth mentioning that some linguists have taken the passive construction to involve a switch from an eventive interpretation to a stative interpretation. Certain data (see Z. Harris, 1975) suggests that the passive morpheme has such aspectual significance:

(72)\* Jones caught the fish freshly.

The fish was freshly caught (by Jones)

The freshly caught (catch + -en) fish.

( $\exists$ s(being-caught(s) & Subject(s, the fish) & Fresh(s)))

(73) ?\* Jones reinterpreted the doctrine newly

The doctrine was newly reinterpreted (by Jones)

The newly reinterpreted doctrine

( $\exists$ s(being-reinterpreted(s) & Subject(s, the doctrine) & New(s)))

Here, the truth-conditions are not given correctly by taking the adverb to modify an event of catching or reinterpreting or retrieving. It is not the catching that is fresh; rather it is the state of being caught that is fresh. It is not the reinterpreting that is new, rather it is the state of being reinterpreted that is new.

We need not suppose that it follows that all passivized verbs (verbs with the passive morpheme affixed) denote states, however. There is a well-known distinction between *adjectival* and *verbal* passives. Adjectival passives are verbs to which the

passive morpheme has been affixed that can appear in adjectival positions as in, e.g., “the opened door”. There are tests for adjectival passives. These include the ability to take the prefix *un-* as in “opened/unopened”; “tutored/untutored”; “loved/unloved”. This affix attaches to adjectives such as “friendly”, “happy”, “lovely”, but not to verbs, as in “kiss”/\* “unkiss”.<sup>14</sup> Furthermore, adjectival passives can appear as the complement of verbs like “seem”, “remain”, and “look”, which take only adjectival complements. Finally, verbal passives, unlike adjectival passives, can permit a secondary adjective as in

(74) They were educated young

(75) They seem educated (\* young) (Roeper & Siegel, 1978)

Thus, perhaps the passivized “caught”, and “reinterpreted” in (72), and (73) are adjectival rather than verbal passives, which explains the shift in their semantics from events to states. Indeed, in the sentences above, we can replace *caught* and *reinterpreted* with clearly adjectival *uncaught* and *unreinterpreted* showing that the passives here are adjectival. Thus:

(76) The still uncaught fish frustrated John.

(77) The yet unreinterpreted doctrine caused a legal crisis.

Verbal passives remain eventive when their source is eventive.

### 3.4 Event Identity and Thematic Roles

The neo-Davidsonian analysis of logical form encodes a “planetary” view of events. Events are conceptualized as individuals of certain kinds (e.g., stabbings, kissings,

---

<sup>14</sup>The prefix “un-” has a different meaning in verbs like “to unload”. To unload something is not simply to abstain from loading it; it is to undo the loading done previously.

etc.) with participants distinguished by the “thematic relation” they uniquely bear to that event (e.g., Agent, Patient, etc.). This category, “event”, has always been viewed with some metaphysical suspicion. What are its identity conditions? When can two event descriptions or quantifications be of the same event? The planetary view of events puts only one constraint on event identity: no two sentences can quantify over the same event if they represent the referents of two different expressions as bearing the same thematic relation to that event.<sup>15</sup> This may be the metaphysical correlate of Chomsky’s Theta Criterion: the requirement that every argument must be assigned a unique thematic role and that every thematic role must be uniquely assigned to an argument (Chomsky, 1986). Although John may be both Agent and Patient of the same event when he shaves himself, if John fights Bill and Bill fights John there must be two events, two fightings, even if one speaks of only one fight. John and Bill cannot both be the Agent of the same event; nor can they both be the Patient of that event. This constraint is the only one built into neo-Davidsonian theory. The question of when two event descriptions or quantifications are of the same event is not answered by the theory itself.

Traditional puzzles about event identity include worries about whether Jones’ swimming across the channel was Jones’ catching of the cold; whether the heating up of the sphere is identical to its rotating (since they occupy the same spatio-temporal region); whether Mary’s writing the check was Mary’s paying the bill; whether Brutus’ stabbing Caesar was Brutus’ killing him; whether Socrates’ dying was Xantippe’s becoming his widow, and so on. One makes distinctions among these by means of Leibniz’s Law: two event descriptions refer to different events if there is something true of the one event that is not true of the other. For example, if the swim caused the catching of the cold but catching the cold didn’t cause the catching of the cold, then the swimming and the catching of the cold can’t be identical. If the rotating was at 20 radians per second and the heating was not, then the events must, again,

---

<sup>15</sup>For simplicity’s sake, I will be concerned only with singular NPs. For an extension of the neo-Davidsonian view to sentences with plural NPs, see Higginbotham and Schein, 1988 and Schein, forthcoming.

not be identical.

I would like to examine the interaction of thematic relation theory with the theory of event identity. Identification of events in the neo-Davidsonian scheme can lead, by instantiation and rearrangement of arguments, to an unsound entailment. For example:

(78) Romeo kissed Juliet  $\leftrightarrow \exists e(\text{kissing}(e) \ \& \ \text{Agent}(e, \text{Romeo}) \ \& \ \text{Patient}(e, \text{Juliet}))$

(79) Juliet kissed Romeo  $\leftrightarrow \exists e(\text{kissing}(e) \ \& \ \text{Agent}(e, \text{Juliet}) \ \& \ \text{Patient}(e, \text{Romeo}))$

If Romeo's kissing Juliet were identical to Juliet's kissing Romeo, then it would follow that

(80) Romeo kissed Romeo/himself  $\leftrightarrow \exists e(\text{kissing}(e) \ \& \ \text{Agent}(e, \text{Romeo}_i) \ \& \ \text{Patient}(e, \text{Romeo/himself}_i))$

and also, of course, that "Juliet kissed Juliet". Neither inference is sound; what is assumed to be true by hypothesis formally entails something false. It follows that these are non-identical events. Similar examples include hitting the 8-ball into the corner pocket and the 9-ball into the side pocket with one stroke. These can't be the same event since one didn't hit the 8-ball into the *side* pocket, although if the two events were identified, this would be a sound inference. Thus, too, writing the check (i.e. for some  $e$ ,  $\text{writing}(e) \ \& \ \text{Theme}(e, \text{the check})$ ) and paying the bill (i.e. for some  $e$ ,  $\text{paying}(e) \ \& \ \text{Theme}(e, \text{the bill})$ ) can't be identical since one didn't write the bill (for some  $e$ ,  $\text{writing}(e) \ \& \ \text{Theme}(e, \text{the bill})$ ).

I would like to examine some apparent problems for the neo-Davidsonian theory: cases in which we seem to want to say that the events quantified over are identical but in which it seems that the thematic relations borne by the participants rule out their being identified. An example of this is the pair of VPs "saddling a horse" and "putting a saddle on a horse". Intuitively, these can describe the same event, but this assumption seems to be blocked by the thematic relations of the arguments.

Suppose that Roy is in the stable preparing his trusty steed Trigger for riding. Can these two sentences quantify over the same event?

(81) Roy put a saddle on Trigger.

(82) Roy saddled Trigger.

This is not to ask whether the two sentences have the same meaning. Nor is it to ask whether all and only puttings on of saddles are saddlings. Obviously they are not. If I put a saddle on a horse upside down and backwards I have not thereby saddled it. The question is: are these events identifiable contingently? Can the event position quantified over in the logical forms corresponding to these sentences be identical? Must we say that he did at least two distinct things: saddled Trigger and put a saddle on him?

It is natural to think that these sentences involve the same event. For one, either of these sentences are possible answers to "What did Roy do?": "Roy put a saddle on Trigger"; "Roy saddled Trigger". The conjunction of them sounds redundant as a reply, however: "Roy saddled Trigger and put a saddle on him." Furthermore, both sentences are possible answers to "What happened to Trigger?" Again, the conjunction of (81) and (82) seems redundant.

Certainly if events are individuated by their spatio-temporal boundaries (à la Quine), then the boundaries of the first event (whatever they are<sup>16</sup>) are the boundaries of the second. Furthermore, if events are individuated by their causes and effects (à la Davidson), then the causes and effects of the first event would seem to be all and only the causes of the second event. Certainly, neither is a cause of the other. Roy's

---

<sup>16</sup>The idea that events have definite spatio-temporal boundaries seems to me indefensible. World War II was certainly an event, something that happened, but where were its spatio-temporal boundaries? Spatially, were its boundaries just where there were combatants? Would the spatio-temporal boundaries individuating the war have been different if a single soldier had made a further step to the left in Italy on a certain date than he did? Consider, too, the event of Bernard's preparing for the priesthood. When does it begin? With his first consideration of the idea or when he first enters the seminary? Do the spatio-temporal boundaries of this event coincide with the spatio-temporal boundaries of where Bernard is during the period in which he is preparing to become a priest or only, perhaps, with those times and places in which he is having priestly thoughts?

saddling Trigger didn't cause Roy to put a saddle on Trigger; nor did Roy's putting a saddle on Trigger cause him to saddle Trigger. It would seem, rather, that the causes of both the saddling and the putting on of the saddle consisted of such things as Roy's desires to ride and so on. Furthermore, both the saddling and the putting on of a saddle resulted in his being able to ride.

No adverbial modifiers distinguish the two events. If Roy saddled Trigger deliberately /slowly /stubbornly /or in a huff... then Roy put a saddle on Trigger deliberately /slowly /stubbornly /in a huff..., and vice versa. Adverbial modifiers don't tease apart our intuitions of event identity here as they do when we observe that a certain sphere rotates quickly or at 30 radians per second but doesn't heat up slowly or at 30 radians per second. Thus, even though the heating up and the rotating are spatio-temporally identical, if something can't heat up at 30 radians per second, then we must suppose that the heating and the rotating are different events.

It is sometimes argued that if two clauses involve an identical event, and one sentence expresses the manner in which the other sentence occurred, then they cannot be the same event. Thus, we have

(83) Roy saddled Trigger by putting a saddle on him. (True)

(84) Roy put a saddle on Trigger by saddling him. (False)

One is, therefore, to conclude that the events can't be identical since they can't be substituted for one another *salva veritate*. This argument is fallacious, though, since no one will deny that saddling the horse is identical to saddling the horse but

(85) Roy saddled Trigger<sub>i</sub> by saddling Trigger<sub>i</sub> (/him<sub>i</sub>).

is false; the second clause does not express the manner in which Trigger was saddled. One might use such a sentence with emphasis on the second phrase to show that there was no particular manner in which Roy saddled the horse, but such a use doesn't legitimately report a manner in which the horse was saddled.

Furthermore, not all sentences involving identicals exhibit perfect symmetry. The reduction of thermodynamics to statistical mechanics involves the claim that the heat of a gas is nothing other than the mean kinetic energy of its molecules. But consider:

(86) The heat of a gas is reducible to the mean kinetic energy of the molecules of a gas (True)

(87) The mean kinetic energy of the molecules of a gas is reducible to the heat of the gas. (False)

It need not count against the identity of an event and the event in the manner clause, therefore, that one can't always inter-substitute the two phrases. For example, cooking spaghetti and boiling seem to be identifiable despite the fact that one says:

(88) John cooked the spaghetti by boiling it.

But not:

John boiled the spaghetti by cooking it.

Thus, the asymmetry of manner phrases doesn't count against the identifiability of two events.

On the standard neo-Davidsonian view, the sentences at issue, (81) and (82), would be given these truth-conditions.

(89)  $\exists e(\text{saddling}(e) \ \& \ \text{Agent}(e, \text{Roy}) \ \& \ \theta(e, \text{Trigger}))$

Roy saddled Trigger

(90)  $\exists e(\text{putting}(e) \ \& \ \text{Agent}(e, \text{Roy}) \ \& \ \text{Theme}(e, \text{a saddle}) \ \& \ \text{Location}(e, \text{on Trigger}))$

Roy put a saddle on Trigger.

What thematic relation does "Trigger" bear in (89)? Certainly, Theme seems to be the most likely choice. David Dowty's (1991) theory of thematic roles, for example, would assign "Trigger" the Theme role since it (i) undergoes change of state,

(ii) is causally affected by another participant, (iii) is inactive relative to the other argument's activity. Further, according to Terence Parsons, "the direct object of (noncausative) transitive verb in English is always a Theme" (p. 74). None of the other canonical roles listed above seems to be more appropriate.

However, if  $\theta$  in (89) is "Theme", then, despite the reasons just given, the events quantified over in (89) and (90) cannot be identical. If they were identified, then "Roy saddled Trigger" (82) and "Roy put a saddle on Trigger" (81) would entail

(91) Roy saddled a saddle.

( $\exists e(\text{saddling}(e) \ \& \ \text{Agent}(e, \text{Roy}) \ \& \ \text{Theme}(e, \text{a saddle}))$ )

meaning that Roy put one saddle on top of another. This entailment isn't valid in English, of course, so either the events quantified over aren't the same event or "Trigger" is not a Theme in (89).

If "Trigger" isn't a Theme in (89), then what thematic relation does "Trigger" bear? "Trigger" in (89) certainly doesn't seem to be a Location. It can't appear alone as the Location argument of "put".

(92)\* Roy put a saddle Trigger.

Nor does any other thematic relation seem appropriate for "Trigger". It is not an Agent; nor is it a Goal as in "John gave a book to Mary (=Goal)"; nor is it a Benefactive as in "we threw John (=Benefactive) a party"; nor an Instrument; nor an Experiencer; nor a Patient, since the Patient and Theme roles were virtually equivalent.<sup>17</sup>

---

<sup>17</sup>Alec Marantz points out that this consideration isn't conclusive unless we adopt the unmotivated claim that only prepositional phrases are Locations. He suggests that one might view "Trigger" as a Location in both sentences. The verb "put" simply requires a preposition with its Location for unknown reasons. Of course, one might dispute the details of the analysis of this example, but the problem raised by this example is a general one. Even if the problems addressed don't apply to this example, one would need to show that no two verbs in a language could encode the same entity as bearing two different thematic relation to contingently identical events.

Perhaps the best solution is to suppose that contrary to the assumptions of standard neo-Davidsonian theory, “Trigger” doesn’t semantically bear a thematic relation to an event in (82). “To saddle” requires a direct object argument and an external argument syntactically. However, this example shows that we must reject the idea that for every sentence:

(93) [<sub>S</sub> arg1 V arg2...argN (adjunct1 adjunct2...)]

there must be corresponding truth-conditions of the form:

(94)  $\exists e(V\text{-ing}(e) \ \& \ \theta_1(e, \text{arg1}) \ \& \ \theta_2(e, \text{arg2}) \ \dots \ \& \ \theta_n(e, \text{argN}) \ \dots)$

The argument position of “Trigger” is syntactically obligatory in “Roy saddled Trigger”, but the referent of “Trigger” does not bear a canonical thematic role semantically.

The fact that the semantics of “to saddle” seems to depart from the standard neo-Davidsonian account is not crippling, however. We want our semantic theory to account for the fact that “Roy saddled Trigger” entails but is not entailed by “Roy put a saddle on Trigger”. It seems that the only way to account for this in the present framework is to suppose that the truth-conditions of “saddle” are given by

(95)  $\text{Val}(\langle x,y,e \rangle, [{}_V \text{ saddle}]) \leftrightarrow (\text{putting}(e) \ \& \ \text{Agent}(e,x) \ \& \ (\exists z(\text{Theme}(e,z) \ \& \ \text{saddle}(z))) \ \& \ \text{Location}(e, \text{on } y) \ \& \ e \text{ is done in the relevant manner.})$

The verb “to saddle”, then, still requires two arguments syntactically, but the referent of its direct object doesn’t itself bear a thematic relation to an event. Nevertheless, the truth-conditions of the sentence are still given in terms of thematic relations to events. This solution preserves our intuitions about event identity and entailment while requiring us to give up the view that for every argument of an event verb, there is a canonical thematic relation that it bears to the event. Syntactically, “Trigger” seems to be a Theme since it appears in a position typically occupied by Themes.

However, the identity of saddlings with puttings on of saddles prevents this analysis. The formula (89) isn't the correct form for the truth-conditions of (82). Rather, the truth-conditions of "Roy saddled Trigger" should be given by:

(96)  $\exists e(\text{putting}(e) \ \& \ \text{Agent}(e, \text{Roy}) \ \& \ \exists z(\text{Theme}(e, z) \ \& \ \text{Saddle}(z)) \ \& \ \text{Location}(e, \text{on Trigger}) \ \& \ e \text{ is done in the relevant manner.})$

Thus, "Roy saddled Trigger" entails, but is not entailed by, "Roy put a saddle on Trigger".

Furthermore, if we assume that the object of a *with*-phrase is coreferential with the entity quantified over as the thing functioning as a saddle in (95), then we can account for this entailment formally:

(97) Roy saddled Trigger with Dale's favorite saddle.

Therefore, Roy put Dale's favorite saddle on Trigger

The truth-conditions of the premise would be:

(98)  $\exists e(\text{putting}(e) \ \& \ \text{Agent}(e, \text{Roy}) \ \& \ \exists z_i(\text{Theme}(e, z_i) \ \& \ \text{Saddle}(z_i)) \ \& \ \text{Location}(e, \text{on Trigger}) \ \& \ e \text{ is done in the relevant manner} \ \& \ \text{With}(e, [\text{Dale's favorite saddle}]_i)$

The treatment of "with" here is like that of "by" in the passives considered above. The *with*-phrase allows the adjoined object to bear a thematic relation obliquely through coindexation.

According to recent work of Ken Hale and Jay Keyser (Hale and Keyser, 1991, 1992), the verb "to saddle" is actually derived from a structure of roughly the form "put a saddle on".<sup>18</sup> In their work, Hale and Keyser have attempted to explain gaps in the lexicon by postulating a level of syntactic projection they call "lexical relational

---

<sup>18</sup>See Clark and Clark, 1979, for a thorough survey of verbs related to nouns. Not all the examples related to this discussion involve "put". Witness: *take the skin off the rabbit/skin the rabbit*.

structure". At this level, H&K claim that thematic relations can be reduced to syntactically specified relations. Syntactic principles constrain the derivation of words from structures at this level. Thus, an explanation is given as to why, as a general phenomenon, natural languages licenses

(99) The mare had a foal.

The mare foaled.

But not:\* It (-pleonastic) mared a foal.

The verb "to foal" is seen to be derived at a deep level of syntax from a structure similar to that of "have a foal" while syntactic constraints rule out incorporating "the mare" down into the verb. Given the support of Hale and Keyser's theory as to the relatedness of "to saddle" and "put the saddle on", the fact that the satisfaction conditions of "to saddle" are those of "put a saddle on" is predictable. The standard neo-Davidsonian account is generally valid, but the exceptions, derived verbs like "to saddle", are predictable by syntactic means.

The constructions "saddle" and "put a saddle on" are not the only examples of ways of expressing what Roy did to Trigger. Consider the sentence

(100) Roy gave Trigger a (quick) saddling

(Cf: Roy gave Trigger a beating)

This sentence may be true of the same event as putting a saddle on Trigger or saddling him. The verb "to give", however, is usually thought of as having an Agent, a Theme, and a Goal argument, which may be distributed in an active construction in either of two ways.

(101) John gave a book to Mary.

John gave Mary a book.

The truth-conditions of (100) would be given as follows:

(102)  $\exists e(\text{giving}(e) \ \& \ \text{Agent}(e, \text{Roy}) \ \& \ \text{Theme}(e, \text{a saddling}) \ \& \ \text{Goal}(e, \text{Trigger}))$

Then, assuming that the saddling and the giving of a saddling are identical, both sentences would entail:

(103)\* Roy put a saddling on Trigger.

Or:

Roy gave Trigger a saddle.

$\leftrightarrow \exists e(\text{giving}(e) \ \& \ \text{Agent}(e, \text{Roy}) \ \& \ \text{Theme}(e, \text{a saddle}) \ \& \ \text{Goal}(e, \text{Trigger}))$

Again, we would have a case in which a sentence which seems to involve the same event as “Roy saddled Trigger” or “Roy put a saddle on Trigger” leads, by argument switching, to a falsehood.

In her paper “Light Verbs in English” (1988), Kate Kearns proposes that “give” in (100) doesn’t have the semantic properties of “give” in (101). Notice that the syntactic properties of these constructions are different. For example, although (100) is acceptable, this alternation, acceptable for “give” ordinarily, is less so:

(104) \* Roy gave a quick saddling to Trigger.

Kearns supposes that in certain “light verb” constructions (following Jespersen’s terminology),

the semantic content of the predicate is provided not by the verb but by the action nominal complement; i.e. “John made an inspection of the premises” means “John inspected the premises”. (p. 3)

Such light verb constructions include:

(105) take a walk; give a sigh; give the floor a sweep; make an offer; have a bite.

Following Kearns’ suggestion, then, we can give the truth-conditional contribution of “give NP a V-ing” as:

(106)  $\text{Val}(\langle x,y,e \rangle, [v', \text{give NP a V+ing}]) \leftrightarrow \text{Val}(\langle x,y,e \rangle, [v', \text{V NP}])$

That is, the truth-conditions of “give Trigger a saddling” are simply those of “saddle Trigger”, which, in turn, entail those of “put a saddle on Trigger”. The fact that the semantics of these “give NP a V-ing” constructions are not compositional in the same way as ordinary “give” constructions suggests that learning the meaning of these constructions would be harder than learning the meaning of the ordinary construction.

Finally, we have the sentence

(107) Roy got Trigger saddled.

This sentence is clearly entailed by “Roy saddled Trigger”. What thematic relation, if any, does “Trigger” bear in (107)? Clearly, it would be a mistake to suppose that the truth-conditions are parallel to that of “Roy got a new car” where the new car is a Theme. If this were so, we would have as the truth-conditions of (107):

(108)  $\exists e(\text{getting}(e) \ \& \ \text{Agent}(e, \text{Roy}) \ \& \ \text{Theme}(e, \text{Trigger}) \ \& \ \theta(e, \text{being-saddled}))$

On the assumption that a specific instance of getting Trigger saddled is identical to putting a saddle on him, this would allow us to infer that

(109) Roy got a saddle saddled

$\exists e(\text{getting}(e) \ \& \ \text{Agent}(e, \text{Roy}) \ \& \ \text{Theme}(e, \text{a saddle}) \ \& \ \theta(e, \text{being-saddled}))$

Again, this inference must be blocked. How should we represent the truth-conditions of (107)? Clearly, “Roy” is an Agent, and clearly, the verb quantifies over an event, since the verb meets the diagnostics for event-quantification outlined above.<sup>19</sup>

---

<sup>19</sup>“Saddled” isn’t an adjectival passive here since “Roy got Trigger unsaddled in two minutes” means that Roy reversed the saddling and it took two minutes rather than that Roy brought it about that Trigger was not saddled during that interval.

Perhaps we should first consider the truth-conditions of *Trigger got saddled*; this seems to be a simple passive construction with “get” as the auxiliary verb instead of “be” as in “Trigger was saddled”. Notice that the long passive form is available, too: we can say *Trigger got saddled by Roy*. As a passive, the truth-conditions of this sentence would be given by:

(110)  $\exists e(\text{putting}(e) \ \& \ \text{Agent}(e, -en_i) \ \& \ \exists z(\text{Theme}(e,z) \ \& \ \text{Saddle}(z)) \ \& \ \text{Location}(e, \text{on Trigger}) \ (\& \ \text{By}(e, \text{Roy}_i)))$

*Trigger got saddled by Roy.*

The transitive verb “get” simply adds an external argument to this. Notice that one can say both *Roy himself got Trigger saddled*, meaning that Roy performed the saddling of Trigger, and also *Roy got Trigger saddled by the stableboy*, meaning that Roy was responsible for initiating a chain of events that resulted in the stableboy saddling Trigger. Thus, the subject of transitive “get”, as in (107), can be assigned a role we might call “Initiator”, where Initiators differ from Agents in that an Agent must perform the action while an Initiator need only bring it about that it was performed. Thus, “Roy got Trigger saddled (by the stableboy) (with Dale’s favorite saddle)” would be assigned the truth-conditions:

(111)  $\exists e(\text{putting}(e) \ \& \ \text{Initiator}(e, \text{Roy}) \ \& \ \exists -en_i(\text{Agent}(e, -en_i)) \ \& \ \exists z_j(\text{Theme}(e,z_j) \ \& \ \text{Saddle}(z_j)) \ \& \ \text{Location}(e, \text{on Trigger}) \ \& \ e \text{ is done in the relevant way} \ (\& \ \text{By}(e, \text{the stableboy}_i)) \ (\& \ \text{With}(e, [\text{Dale’s favorite saddle}]_j)))$

Thus, *Roy got Trigger saddled by the stableboy with Dale’s favorite saddle* would formally entail: *The stableboy saddled Trigger*, *The stableboy put a saddle on Trigger*, *The stableboy put Dale’s favorite saddle on Trigger*, *Trigger was saddled by the stableboy*, and so on. However, pragmatic knowledge would be required to infer *Roy saddled Trigger* from *Roy got Trigger saddled* since the Agent is not actually specified in the truth-conditions. One makes the inference on the basis of knowing that if Roy initiated the saddling and no one else is mentioned who performed it, then Roy must have done it himself.

What all of this shows is that the neo-Davidsonian account need not require that intuitively identical events must be distinguished because of the thematic roles assigned to the arguments of these sentences.

### 3.5 Conclusions

In this paper, I have argued that (i) a class of verbs can be distinguished which require quantification over an event-position in their logical forms; (ii) thematic relations must be employed in the truth-conditions of eventive verbs if facts about entailments and syntactic alternations are to be explained by the grammar; (iii) that metaphysical considerations about how I might report unreal events do not establish the adicity of a verb; this is a syntactic matter determined by the obligatoriness of arguments of active sentences; and (iv) neo-Davidsonian accounts of truth-conditions need not violate our intuitions about event identity: saddling a horse, putting a saddle on it, giving it a saddling, and so on, may be identifiable events despite the apparent incompatibility of their thematic projections. The only thing that must be rejected is the idea that the thematic roles of arguments are somehow encoded by their syntactic positions. Thus, crucially, the referent of an NP argument need not bear a canonical thematic relation in that sentence's truth-conditions. Procedural approaches "linking" syntactic arguments to their thematic relations (as in Dowty, 1991) on the basis of diagnostic tests applied to the various arguments may, as we have seen, lead to invalid inferences when their thematic relation assignments are inconsistent. The thematic relations of a sentence's truth-conditions need not be predictable from the surface syntax of a construction. If Hale and Keyser are right, departures from the standard neo-Davidsonian picture may be predicted on the basis of a level of syntax far removed from the surface.

I have not yet provided a general criterion of event identity, a way of filling out the formula:

(112)  $\forall e_1 \forall e_2 (e_1 = e_2 \leftrightarrow \dots)$

We have given a sufficient condition for the non-identity of events: no two sentences with singular NP arguments can quantify over identical events if they have distinct objects bearing the same thematic relation. Thus, writing the check cannot be paying the bill if “the check” and “the bill” are both Themes; otherwise, the identity of these events would entail, invalidly, that one wrote the bill. What necessary conditions are there for two sentences which involve the same entities bearing the same thematic relations to quantify over the same event? For example, is Brutus’s stabbing of Caesar his killing of Caesar? In both of the relevant sentences, “Brutus” is Agent and “Caesar” is Theme or Patient. Thus, in identifying the events, we do not license invalid inferences of the switch-argument sort we examined above.

Identity conditions for events are not persistence conditions, as sought, for example, for the metaphysical category of persons and material objects. In providing persistence conditions, necessary and sufficient conditions are proposed for an element of a category to persist while undergoing various changes. Thus, one proposes that a person P at time t is the same person as at time t' just in case C, where C spells out the conditions that are necessary and sufficient for a person to persist over a period of time despite various sorts of changes. Similarly, persistence conditions are given for material objects and organizations and institutions and so on.

Persistence conditions for events are inappropriate since token events do not exist *through* a period of time, as persons and objects do; instead, they occupy a certain period of time. Time-slices of that period do not contain the event itself; rather, they are the temporal constituents of that event. Further, as a period-bound individual, the event that was Brutus’s stabbing Caesar could not have happened other than when it did; as particulars distinguished in part by their temporal location, token events cannot recur. If Brutus had stabbed Caesar twice, the second token stabbing that Brutus performed would not be the same event as one that occurred earlier. Similarly, no two token playings of the Star Spangled Banner are the same event. While the same object can exist at different times, the same token event cannot

occupy different times.

Furthermore, token events do not persist through changes; they *are* changes. Thus, we cannot contemplate an event being different in some way, such as changing with respect to its Theme, and remaining the same event. Rather the event involves a change in the participants of that event. Thus, a raining involves a change in the position of the water droplets involved; a stabbing involves a change in its Theme, as does a killing.

Since we need a criterion of event identity and since the persistence of events across time and change is not at issue, then we don't need anything more for a theory of event identity than Leibniz's Law plus the metalinguistic resources of the neo-Davidsonian theory. Leibniz's Law states that two expressions refer to the same thing just in case every extensional property true of the first expression is true of the second expression as well. Thus, the truth-conditions of two expressions involve the same event if no extensional property of the first event is not true of the second event and vice versa. That is, the truth-conditions of two expressions involve the same event just in case no switching of conjuncts leads to a falsehood. Leibniz's Law entails that two events are identical just in case they occupied the same period of time, had all and only the same participants, and these participants played all and only the same thematic roles. Two events descriptions pick out the same event just in case all and only the same facts concerning temporal properties, participants, and thematic roles are the same for each event. Thus, two events are identical only if all and only the same properties and thematic relations are true of both events.<sup>20</sup>

This, I think, covers the data adequately. The event of Brutus stabbing Caesar, which involves a knife and involves Caesar's back and takes place in Rome is distinguishable from the event of Brutus killing Caesar by means of Leibniz's Law. That

---

<sup>20</sup>Facts about the manner in which an event was accomplished are excluded from these considerations for the reasons given above; perhaps manner phrases express facts about the Agent, not the event itself. For example, "How did John cook the spaghetti?" is answered "By boiling it." This suggests that John was the agent of the spaghetti-cooking because he was the agent of the spaghetti-boiling, and the boiling was the cooking.

is, the events quantified over in these truth-conditions:

(113)  $\exists e(\text{stabbing}(e) \ \& \ \text{Agent}(e, \text{Brutus}) \ \& \ \text{Patient}(e, \text{Caesar}) \ \& \ \text{Instrument}(e, \text{a knife}) \ \& \ \text{Internal-Location}(e, \text{in the back}) \ \& \ \text{Frame-Location}(e, \text{in Rome})\dots)$

(114)  $\exists e(\text{killing}(e) \ \& \ \text{Agent}(e, \text{Brutus}) \ \& \ \text{Patient}(e, \text{Caesar}) \ \& \ \text{Instrument}(e, \text{a knife})$

can't be identified, because while it is true that Brutus stabbed Caesar in the back, he didn't kill him in the back. The stabbing event and the killing event have properties which distinguish them. Similarly, the heating of a sphere and its rotating are not identical events despite the identity of their spatio-temporal boundaries and Themes (the sphere) because (*pace* Quine) the rate of the rotating (30 radians per second) is not the rate of the heating (10 degrees C per second). Further, the widowing of Xantippe is not the same event as the dying of Socrates because the Theme of the widowing (Xantippe) is not identical to the Theme of the dying (Socrates). On the other hand, John's cooking the spaghetti and boiling it seem to be identical events since every property true of the first is true of the second. The cooking and the boiling both took place during the same interval of time, in the same location, had the same Agent, and so on. <sup>21</sup>.

The neo-Davidsonian approach thus facilitates the discussion of event identity by laying bare the predicate-object structure involved in expressions that quantify over events. Any two event descriptions or expressions quantify over the same event just in case all and only the same thematic relations and non-attributive properties of the events are the same.

---

<sup>21</sup>Boiling the spaghetti vigorously need not entail cooking the spaghetti vigorously, however, if "vigorously" is interpreted attributively as "vigorous for a boiling"

## References

- Baker, M., Johnson K., and Roberts, I. 1989: "Passive Arguments Raised". *Linguistic Inquiry*, Vol. 20, no. 2, Spring, 1989, pp. 219-251.
- Bennett, J. 1988: *Events and Their Names*. Indianapolis: Hackett Publishing Company.
- Boghossian, P. 1989: "The Rule-Following Considerations". *Mind*, Vol. XCVIII, No. 392, October, 1989, pp. 507-49.
- Forthcoming: "Analyticity", In C. Wright and B. Hale, eds, *Blackwell's Companion to the Philosophy of Language*. Cambridge, MA: Basil Blackwell, Inc.
- Boolos, G. 1981: "For Every A There is a B", *Linguistic Inquiry*, Vol. 12, No. 3, 1981, pp. 465-7.
- 1984: "To Be Is To Be A Value of a Variable (Or To Be Some Values of Some Variables)". *MIT Cognitive Science Center Occasional Papers*.
- Bostock, D. 1990: "Logic and Empiricism", *Mind*, Vol. XCIX, No. 396, October, 1990, pp. 571-582.
- Bowerman, M. 1982: "Evaluating Competing Linguistic Models with Language Acquisition Data: Implications of Developmental Errors with Causative Verbs", *Quaderni di Semantica* vol. 3, pp. 5-66.
- Burge, T. 1986: "Intellectual Norms and the Foundations of Mind". *Journal of Philosophy*, Vol. LXXXIII, no. 12, December, 1986, pp. 697-720.
- 1992: "Philosophy of Language and Mind: 1950-1990". *Philosophical Review*, Vol. 101, no. 1, January, 1992, pp. 3-52.
- Carter, R. 1976: "Some Constraints on Possible Words", *Semantikos*, Vol. 1, No. 2, pp. 27-66.

- 1988: *On Linking: Papers by Richard Carter*, B. Levin and C. Tenny, eds., MIT Lexicon Project Working Papers no. 25.
  
- Castañeda, H-N. 1967: "Comments". In N. Rescher, ed. *The Logic of Decision and Action*, Pittsburgh: U. of Pittsburgh Press.
  
- Chomsky, N. 1975: *Reflections on Language*. New York: Pantheon Books.
  
- 1986: *Knowledge of Language*. New York: Praeger.
  
- 1987: "Language and Problems of Knowledge". ms. MIT.
  
- 1988: *Language and Problems of Knowledge: The Managua Lectures*. Cambridge, MA: The MIT Press.
  
- 1991: "Language from an Internalist Perspective". ms. MIT.
  
- 1992: "Language and Interpretation: Philosophical Reflections and Empirical Inquiry". In J. Earman, ed., *Inference, Explanation, and Other Philosophical Frustrations*. Berkeley: U. of California Press, 1992.
  
- 1992: "A Minimalist Program For Linguistic Theory". *MIT Occasional Papers in Linguistics*, No.1, 1992.
  
- Clark, E. V. and Clark, H. H. 1979: "When Nouns Surface as Verbs", *Language*, vol. 55, pp. 767-811.
  
- Crane, T., and Mellor, D.H., 1990: "There is No Question of Physicalism". *Mind*, Vol. XCIX, no. 394, April, 1990, pp. 157-206.
  
- Davidson, D. 1965: "Theories of Meaning and Learnable Languages". In *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press, 1984.
  
- 1967: "Truth and Meaning". In *Inquiries into Truth and Interpretation*.
  
- 1967: "Reply to Castañeda on Agent and Patient". In *Essays on Actions and Events*.

- Oxford: Clarendon Press.
- 1967: “The Logical Form of Action Sentences”. In *Essays on Actions and Events*.
- 1986: “A Nice Derangement of Epitaphs”. In E. LePore 1986.
- Descartes, R., *Meditations on First Philosophy*
- Dowty, D. 1979: *Word Meaning and Montague Grammar*. Dordrecht: Reidel.
- 1989: “On the Semantic Content of the Notion of ‘Thematic Role’”. In G. Chierchia, B. Partee, and R. Turner, eds., *Properties, Types, and Meaning II*. Dordrecht: Kluwer.
- 1991: “Thematic Proto-Roles and Argument Selection”. *Language*, Vol. 67, no. 3, pp. 547-619.
- Dummett, M. 1970: “Wang’s Paradox”. In *Truth and Other Enigmas*. Cambridge, MA: Harvard University Press, 1978.
- 1973: “The Significance of Quine’s Indeterminacy Thesis”. In *Truth and Other Enigmas*.
- 1976: “Is Logic Empirical?”. In *Truth and Other Enigmas*.
- Engel, P. 1991: *The Norm of Truth*. Toronto: University of Toronto Press.
- Etchemendy, J. 1990: *The Concept of Logical Consequence*. Cambridge, MA: Harvard University Press.
- Evans, G. 1976: “Semantic Structure and Logical Form”. In G. Evans and J. McDowell, eds. *Truth and Meaning: Essays in Semantics*. Oxford: Oxford University Press.
- Fodor, J. 1983: *The Modularity of Mind: An Essay on Faculty Psychology*. Cambridge, MA: The MIT Press/Bradford Books.
- 1987: *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge, MA: The MIT Press/Bradford Books.

- 1990: “A Theory of Content, I: The Problem”, in Fodor, 1990.
  - 1990: “A Theory of Content, II: The Theory” in Fodor, 1990.
  - 1990: “Stephen Schiffer’s Dark Night of the Soul: A Review of *Remnants of Meaning*” in Fodor, 1990.
  - 1990: *A Theory of Content and Other Essays*. Cambridge, MA: The MIT Press/Bradford Books.
- Frawley, E. 1992: *Linguistic Semantics*. Hillsdale, NJ : Lawrence Earlbaum.
- George, A. 1986: “Whence and Whither the Debate between Quine and Chomsky?”, *Journal of Philosophy*, Vol. LXXIII, No. 9, September, 1986.
- Greenwood, J. D. 1990: “Analyticity, Indeterminacy and Semantic Theory: Some Comments on ‘The Domino Theory’”, *Philosophical Studies*, vol. 58. pp. 41-49.
- Grimshaw, J. 1990: *Argument Structure*. Cambridge, MA: The MIT Press.
- Gropen, J., S. Pinker, M. Hollander, and R. Goldberg 1991: “Affectedness and Direct Objects: The Role of Lexical Semantics in the Acquisition of Verb Argument Structure”, *Cognition*, Vol. 41, 1991, pp. 153-195.
- Haack, S. 1978: *Philosophy of Logics*. New York: Cambridge University Press.
- Hale, K. 1986: “Notes on World View and Semantic Categories: Some Warlpiri Examples”, In P. Muyskens and H. van Riemsdijk, eds., *Features and Projections*. Dordrecht: Foris Publications.
- Hale, K. and Keyser, J. 1991: *On the Syntax of Argument Structure*, MIT Lexicon Project Working Papers.
- 1992: “On Argument Structure and the Lexical Representation of Syntactic Relations”, ms, MIT.

- Harris, Z. 1982: *A Grammar of English on Mathematical Principles*. New York: Wiley.
- Hellman, G. 1980: "Quantum Logic and Meaning", *PSA 1980*, Vol. 2, pp. 493-511.
- Higginbotham, J. 1985: "On Semantics", *Linguistic Inquiry*, Vol. 6, No. 4, pp. 547-93.
- 1986: "Linguistic Theory and Davidson's Program in Semantics". In LePore, 1986.
- 1987: "Is Semantics Necessary?", *Proceedings of Aristotelian Society*, Vol. 88, p. 230.
- 1990: "The Place of Natural Language". ms. MIT.
- 1991: "Parsons and Dowty on Thematic Roles". ms. MIT.
- forthcoming: "Grammatical Form and Logical Form", *Philosophical Perspectives*, Vol. 7-8.
- and Schein, B. 1988: "Plurals", *NELS Proceedings*.
- Hornstein, N. 1988: "The Heartbreak of Semantics", *Mind and Language*, Vol. 3, No. 1, Spring, 1988, pp. 9-27.
- Horwich, P. 1987: *Asymmetries in Time*. Cambridge, MA: The MIT Press.
- Hughes, R. I. G. 1981: "Quantum Logic", *Scientific American*, Vol. 245, October, 1981, pp. 202-213.
- Jackendoff, R. 1972: *Semantic Interpretation in Generative Grammar*. Cambridge, MA: The MIT Press.
- 1990: *Semantic Structures*. Cambridge, MA: The MIT Press.
- Jaeggli, O. 1986: "Passive", *Linguistic Inquiry*, Vol. 17, no. 4, pp. 587-622.
- Johnston, M. 1988: "The End of the Theory of Meaning", *Mind and Language*, Vol. 3, No. 1, Spring, 1988, pp. 28-43.

- Katz, J. 1986: *Cogitations*. Oxford: Oxford University Press.
- Kearns, K. 1988: "Light Verbs in English", ms. MIT.
- 1991: *The Semantics of the English Progressive*, PhD Dissertation, MIT.
- Kripke, S. 1982: *Wittgenstein on Rules and Private Language*. Cambridge, MA: Harvard University Press.
- Lasnik, H. 1988: "Subjects and the  $\Theta$ -Criterion", *Natural Language and Linguistic Theory*, Vol. 6, pp. 1-17.
- Layzer, D. 1990: *Cosmogogenesis: The Growth of Order*. Oxford: Oxford University Press.
- LePore, E. 1986: *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*. Cambridge, MA: Basil Blackwell, Inc.
- Loewer, B. and Rey, G. 1991: "Editors' Introduction", *Meaning in Mind*. Cambridge, MA: Basil Blackwell, Inc.
- Lycan, W. 1984: *Logical Form in Natural Language*. Cambridge, MA: The MIT Press/Bradford Books.
- Marantz, A. 1984: *On the Nature of Grammatical Relations*. Cambridge, MA: The MIT Press.
- Mathews, R. 1989: "The Plausibility of Rationalism". In R. J. Mathews and W. Demopoulos, eds., *Learnability and Linguistic Theory*. Dordrecht: Kluwer.
- Millikan, R. 1984: *Language, Thought, and Other Biological Categories: New Foundations for Realism*. Cambridge, MA: The MIT Press/Bradford Books.
- Parsons, T. 1990: *Events in the Semantics of English*. Cambridge, MA: The MIT Press/Bradford Books.

- Penrose, R. 1991: *The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics*. New York: Penguin Books.
- Pesetsky, D. 1992: *Zero Syntax: Experiencers and Cascades*, Version 2.0 (September, 1992). ms. MIT
- Pinker, S. 1989: *Learnability and Cognition: The Acquisition of Argument Structure*. Cambridge, MA: The MIT Press/Bradford Books.
- 1990: "Language Acquisition", In D. Osherson and H. Lasnik, eds., *Language: An Invitation to Cognitive Science, Vol. 1*. Cambridge, MA: The MIT Press/Bradford Books.
- Putnam, H. 1975: "The Logic of Quantum Mechanics", In *Mathematics, Matter, and Method: Philosophical Papers, Vol. 1*. New York: Cambridge University Press.
- 1983: "Analyticity and Apriority: Beyond Wittgenstein and Quine", In *Realism and Reason: Philosophical Papers, Vol. 3*. New York: Cambridge University Press.
- 1983: "Quantum Mechanics and the Observer," n *Realism and Reason*.
- 1983: "Vagueness and Alternative Logic", In *Realism and Reason*.
- Quine, W. 1953: "Two Dogmas of Empiricism", *From A Logical Point of View and Other Essays*. New York: Harper and Row.
- 1960: *Word and Object*. Cambridge, MA: The MIT Press.
- 1970: *Philosophy of Logic*. Cambridge, MA: Harvard University Press.
- 1985: "Events and Reification", In LePore, E. and McLaughlin, B. eds., *Actions and Events: Perspectives on the Philosophy of Donald Davidson*. Cambridge, MA: Basil Blackwell, Inc.
- 1986: "Reply to Jules Vuillemin". In L. E. Hahn and P. A. Schilpp, eds. *The Philosophy of W. V. Quine*. La Salle, IL: Open Court.

- 1991: “Two Dogmas in Retrospect”, *Canadian Journal of Philosophy*, Vol. 21, No. 3, September, 1991, pp. 265-74.
- 1992: *Pursuit of Truth, Revised Edition*. Cambridge, MA: Harvard University Press.
- Rappaport, M. and Levin, B. 1986: “What to Do with  $\Theta$ -Roles”. In *Syntax and Semantics 21: Thematic Relations*.
- Redhead, M. 1987: *Incompleteness, Nonlocality, and Realism: A Prolegomenon to the Philosophy of Quantum Mechanics*. Oxford: Clarendon Press.
- Roeper, T. and Siegel, M. 1978: “A Lexical Transformation for Verbal Compounds”, *Linguistic Inquiry*, Vol. 9, No. 2, pp. 199-260.
- Sainsbury, R. M. 1987: *Paradoxes*. New York: Cambridge University Press.
- Schein, B. Forthcoming: *Plurals and Events*. ms. USC.
- Schiffer, S. 1987: *Remnants of Meaning*. Cambridge, MA: The MIT Press/Bradford Books.
- Smith, B. C. 1992: “Understanding Language”. *Proceedings of the Aristotelian Society*, Vol. 93, pp. 109-141.
- Sorensen, R. 1991: “Vagueness Within the Language of Thought”, *Philosophical Quarterly*, Vol. 41, No. 165, October, 1991, pp. 389-413.
- Stalnaker, R. 1990: “Mental Content and Linguistic Form”, *Philosophical Studies*, Vol. 58, pp. 129-146.
- Wilson, N. L. 1967: “Linguistical Butter and Philosophical Parsnips”, *Journal of Philosophy*, Vol. LXIV, No. 2, February, 1967, pp. 55-67.
- Wright, C. 1976: “Language Mastery and the Sorites Paradox”. In G. Evans and J. McDowell, eds. *Truth and Meaning: Essays in Semantics*. Oxford: Oxford University Press.