#### DAVIDSON'S PROGRAM AND INTERPRETED LOGICAL FORMS

#### 1. Introduction

The "Interpreted Logical Form" (ILF) analysis of attitude ascriptions has been proffered, most notably by Higginbotham (1986) and Larson and Ludlow (1993), as a means of resolving within the framework of Davidson's semantic program the familiar problems posed by attitude ascriptions. In this paper I argue that only an analysis of attitude ascriptions along the lines of Davidson's (1968) "paratactic" analysis can resolve the problem posed by attitude ascriptions within the constraints of Davidson's semantic program. The ILF analyses, though following Davidson in providing recursive formal theories that entail statements of the truth conditions of sentences, violate theoretical constraints that Davidson takes pains to satisfy.

That the ILF analyses violate these constraints raises two theoretical questions: First, do the ILF analyses nonetheless adequately perform what Davidson calls the "central task of a theory of meaning", viz. explaining how "speakers of a language can effectively determine the meaning or meanings of an arbitrary expression" (Davidson, 1967)? Or does this violation serve to undermine their plausibility as explanations of our semantic competence? I will argue that because the ILF analyses violate the theoretical constraints of Davidson's semantic program their plausibility as explanations of our semantic competence is significantly undermined.

<sup>\*</sup> I am grateful to Mark Richard, Robert Stainton, Irene Heim, Brian Ulicny, Paul Pietroski, Andrew Botterell, James Higginbotham, Jason Stanley, Zoltan Szabo, Peter Ludlow, and several anonymous referees for helpful discussion and comments.

<sup>&</sup>lt;sup>1</sup> Versions of Higginbotham's proposal are endorsed by Segal (1989) and Pietrosky (1994). Under Higginbotham's (1986) proposal ILFs need not be actual "objects" of propositional attitudes, but need only be *similar* to the actual objects of the attitudes, whatever these might be. Segal (1989) criticizes this appeal to *similarity* as otiose, and a slightly amended version of Higginbotham's (1986) proposal is defended from Segal's criticism in Pietroski (1994) and Stainton (1994). The debate concerning similarity is irrelevant to the purposes of this paper. Larson and Ludlow's ILF analysis is also endorsed in Larson and Segal (1995).

The second, more general, theoretical question concerns the adequacy of Tarski-style truth theories for explaining our semantic competence. The only approach to resolving the problem posed by attitude ascriptions within the constraints of Davidson's semantic program is via analyses along the lines of Davidson's own (1968) "paratactic" analysis. Davidson's semantic program, as originally conceived, lives or dies with a paratactic analysis, but the paratactic analysis is inadequate. The ILF analyses of attitude ascriptions follow Davidson in providing Tarski-style truth theories, but they violate the constraints of Davidson's semantic program, and their plausibility as explanations of our semantic competence is thus significantly undermined. This raises the following general theoretical question: Should semantic theories attempting to explain out semantic competence be formulated within the framework of Tarski-style truth theories? Or do the shortcomings of Davidson's paratactic analysis and the ILF analyses indicate that our semantic competence with regard to attitude ascriptions, and thus our semantic competence in general, cannot be adequately explained within the framework of a Tarski-style truth theory? I will argue, albeit inconclusively, that the answer to this question is "yes." If the truth conditions of all sentences of natural language could be accounted for by something like a standard Tarski-style truth theory of the sort advocated by Davidson, then we would have the beginnings of an explanation of our semantic competence. Attitude ascriptions, however, patently cannot be adequately treated by standard Tarski-style truth theories. I suggest that the appropriate response to this is not to follow the ILF theorists in formulating ad hoc non-standard Tarski-style truth theories that, via various technical means, manage to assign appropriate truth conditions to attitude ascriptions. Rather the appropriate response is to abandon Davidson's idea that our semantic competence can be adequately explained within the framework of Tarski-style truth theories.

The paper proceeds as follows: In section I, Davidson's semantic program is sketched, and it is argued that if a formal semantic theory is to fall within Davidson's semantic program, it must satisfy a number of constraints, including the constraints of *extensional compositionality*, and *semantic innocence*. In section II, the problem attitude ascriptions pose for Davidson's program is described, and Davidson's own paratactic analysis is reviewed, and shown to fall within the constraints formulated in section I, though it is rejected for other reasons. In section III Larson and Ludlow's ILF analysis is explicated, and shown to violate the constraint of *extensional compositionality*. It is argued, moreover, that because Larson and Ludlow's ILF analysis violates the constraint of *extensional compositionality*, the plausibility of the explanation of semantic competence

provided by their analysis is undermined. Higginbotham's ILF analysis is explicated and developed in section IV. It is shown that Higginbotham's analysis satisfies neither the constraint of *semantic innocence* nor the constraint of *extensional compositionality*, and it is argued that as a result Higginbotham's analysis also fails to provide a plausible explanation of semantic competence. I conclude in section V by briefly considering the more general theoretical question concerning the explanatory adequacy of Tarski-style truth theories.

#### 2. DAVIDSON'S PROGRAM

Davidson maintains that the "central task of a theory of meaning" is to show how it is possible that "speakers of a language can effectively determine the meaning or meanings of an arbitrary expression" (Davidson, 1984, p. 35). That is, a semantic theory must account for the productivity of our linguistic competence; i.e., an adequate semantic theory must be capable of explaining how ordinary speakers, with limited cognitive resources, are able in principle to determine the truth conditions of an infinite number of sentences. Thus an adequate semantic theory must state what it is that speakers know which accounts for this ability. Moreover, Davidson maintains that in order to do this a semantic theory "must give an account of how the meanings of sentences depend upon the meanings of words" (1984, p. 17). Davidson famously proposed that these demands can be met by a theory which takes the form of a standard Tarski-style truth theory for a language.<sup>3</sup>

<sup>&</sup>lt;sup>2</sup> Davidson himself seems to think that a semantic theory is successful if it specifies a theory knowledge of which would enable an agent to interpret utterances of the relevant language. (See especially Davidson 1974.) But followers of Davidson, such as Higginbotham and Larson and Segal, maintain that a semantic theory is successful only if it specifies the theory an agent actually utilizes to interpret utterances of the relevant language. For example, Larson and Segal write that questions concerning semantic knowledge "should be approached from a **cognitivist perspective**, according to which knowledge of language is knowledge of a body of (largely unconscious) rules and principles that assign representations and meanings to the physical forms of signs" (Larson and Segal, 1995, p. 11). Larson and Segal attribute this cognitivist perspective to Higginbotham (1985).

<sup>&</sup>lt;sup>3</sup> It should be noted that the Tarski-style truth theories sketched by me and the ILF theorists go beyond what Davidson himself would endorse. This because Davidson endorses Quine's indeterminacy theses, and thus he resists assigning "absolute" semantic values to expressions. Thankfully the issue indeterminacy is not directly relevant to my concerns, and so I set it aside.

Consider a standard Tarski-style truth theory for a tiny fragment of English:<sup>4</sup>

## **Standard Theory Lexical Rules:**

- (LRa) Val([NP Twain]) = Twain
- (LRb) Val([NP Clemens]) = Twain
- (LRc) Val([NP Huck Finn]) = Huck Finn
- (LRd)  $Val([v \text{ wrote}]) = \{\langle x, y \rangle : x \text{ wrote } y\}$

## **Standard Theory Combinatorial Rules:**

- (CRa)  $Val([s[NP...[VP...]]) = true \text{ iff } \exists x \ (x = Val([NP...]) \& x \in Val([VP...])).$
- (CRb)  $Val([VP[V...]][NP/CP...])) = \{x: \exists y (y = Val([NP/CP...] \& \langle x, y \rangle \in Val([V...]))\}$
- (CRc) Val([s[s'.]]) and [s''.]]) true iff Val([s'.]) = true and Val([s''.]) = true

The Standard Theory above assumes that the syntactic structures interpreted by a semantic theory are "phrase structure markers," or "syntactic trees," that represent a sentence at the syntactic level known as LF. I will refer to such structures as trees or LFs.<sup>5</sup> The function Val() assigns semantic values to sections of the syntactic trees provided by syntax: the Lexical Rules specify the value of Val() for arguments which are terminal nodes (the leaves on the syntactic trees), and the Combinatorial Rules specify the value of Val() for arguments which are nonterminal nodes, where the value of Val() given a nonterminal node n as an argument is a function of the values of Val() for the terminal nodes that n nonterminal nodes n0 nonterminal nodes n1 nodes1 n2 note Val3 note Val4 note Val6 note Val9 note Val

<sup>&</sup>lt;sup>4</sup> In formulating the Simple Theory I mimic the sort of Tarski-style truth theories endorsed by the ILF theorists. For an excellent introduction to such theories, see Larson and Segal, 1995.

Segal, 1995.

<sup>5</sup> I follow Higginbotham and the other ILF theorists in identifying the objects which a Davidsonian theory interprets with *syntactic trees*, or *LFs*, as provided by a Chomskian theory of syntax. Davidson does not make this identification, but to my knowledge it is wholly compatible with Davidson's conception of a proper semantic theory. Also, syntactic trees, or *LFs*, will here be represented one-dimensionally; e.g., the LF of 'Twain wrote *Huck Finn*' is [S1[NP Twain][VP[V wrote][NP *Huck Finn*]]]. (I will use '[S1.]' to abbreviate this LF.)

<sup>&</sup>lt;sup>6</sup> This, of course, is a hyper-simplified semantic theory, though it will suffice for my purposes here.

 $Val([S_1[NP Twain][VP[V wrote][NP Huck Finn]]]) = true \text{ iff } \exists x (x = Twain & x \in \{y: } \exists z (z = Huck Finn & \langle y, z \rangle \in \{\langle v, w \rangle: v \text{ wrote } w\})\})$ 

 $Val([S_2[NP Clemens][VP[V wrote][NP Huck Finn]]]) = true iff \exists x (x = Twain & x \in \{y: \extstyre z | E Huck Finn & \langle y, z \rangle \in \{v, w \rangle: v wrote w\})\})$ 

Let us now consider how the Standard Theory explains our semantic competence with regard to the tiny fragment, with an eye toward determining a number of constraints that a more complex formal semantic theory would have to satisfy if it were to provide a similar explanation of semantic competence. The Standard Theory suggests that ordinary speakers with limited cognitive faculties are able to determine the truth conditions of an (in principle) infinite number of sentences within the tiny fragment because, first, they know, via the Lexical Rules, the unique "atomic" semantic values of the lexical items in the fragment, and second, they know, via the Combinatorial Rules, how to combine these semantic values, as directed by the LF, into "molecular" semantic values. Thus the Standard Theory illustrates the following elegant and powerful strategy for explaining the semantic competence of ordinary speakers: Upon perceiving a well formed utterance or inscription, a competent subject is able to determine the LF – represented by a syntactic tree – of that utterance. The subject comes to understand the utterance, i.e., determine its truth conditions, by working his way up the tree: he first assigns semantic values – via the Lexical Rules - to the terminal nodes of the tree, and he then moves up the tree, combining the "atomic" semantic values - via the Combinatorial Rules - and thereby eventually determines the conditions under which the root [S...]node is to be assigned the semantic value true. This elegant and powerful strategy lies at the heart of Davidson's semantic program; standard Tarskistyle truth theories can serve to explain our semantic competence because they illustrate in detail how this strategy is to be implemented for speakers of the language under investigation.

Let us now formulate a number of constraints that any formal Tarskistyle truth theory must satisfy if it is to illustrate this strategy for explaining our semantic competence. First, it must be such that the theorems that can be derived from it are *interpretive*; i.e., they must state correct truth conditions for the relevant object language sentences. Let us call this the *correctness* constraint.

Second, a semantic theory that illustrates the elegant strategy must be *finitely axiomatized*. A semantic theory with an infinite number of axioms, or rules, would be unknowable (or unlearnable) by beings like us with

limited cognitive resources. It is important to note, however, that being *correct* and being *finitely* axiomatized are *not* jointly sufficient for a semantic theory's being explanatorily adequate. For a theory could be correct and also finitely axiomatized, and yet fail, as Davidson puts it, to "give an account of how the meanings of sentences depend upon the meanings of words" (Davidson, 1984, p. 17). Consider the Absurdly Simple Theory that contains only the rule

```
(CRa) Val([s[NP...][VP...]]) = true \text{ iff}

\exists x \ (x = Val([NP...]) \& x \in Val([VP...]))
```

This Absurdly Simple Theory is obviously finitely axiomatized, and it is *correct*. From it one can derive correct theorems such as

```
Val([S_1[NP Twain][VP[V wrote][NP Huck Finn]]]) = true iff

\exists x \ (x = [NP Twain] \& x \in Val([VP[V wrote][NP Huck Finn]]))
```

But this Absurdly Simple Theory does not provide a plausible explanation of one's semantic competence with regard to the tiny fragment. To put the point crudely, a non English speaker could know that 'Twain wrote *Huck Finn*' is true if and only if the referent of 'Twain' satisfies the predicate 'wrote *Huck Finn*', and yet have no idea what the sentence means. The problem of course is that the Absurdly Simple Theory does not require its user to know the meanings of the words in the sentence, nor how the truth conditions of the sentence are a function of the meanings of those words.

If a semantic theory is to illustrate Davidson's explanatory strategy, it must satisfy a number of additional constraints. First, it must assign meanings, or semantic values, to the terminal nodes (words), and the semantic values assigned to the terminal nodes must be *fixed*; every occurrence of a terminal node, regardless of what syntactic structure it occurs in and where in a syntactic structure it occurs, must be assigned the same semantic value. (Thus Frege's (1892) trick of "shifting" the referents of terms inside the content clauses of attitude ascriptions is ruled out.) Following Davidson (1968), I will call the latter conjunct of this constraint the *semantic innocence* constraint. This constraint is stated more formally as follows:<sup>7</sup>

<sup>&</sup>lt;sup>7</sup> What follows are preliminary characterizations of semantic innocence and extensional compositionality. The definitions are adequate for theories that assign (non-indexical, non-ambiguous) expressions a single semantic value.

The theories of Larson and Ludlow and Higginbotham assign expresions multiple semantic values. (On these views, for example and speaking roughly, x is a semantic value of 'runs' iff x runs.) Definitions adequate to the general case are given in section IV. Since the present definitions are much closer to standard characterizations of compositionality and innocence I focus my discussion on them.

Let  $[\beta \dots [\alpha \dots]]$  and  $[\gamma \dots [\alpha \dots]]$  be any distinct syntactic structures that dominate occurrences of a smaller syntactic structure  $[\alpha \dots]$ . Semantic theory T is *innocent iff* if according to T Val $([\alpha \dots]) = x$  when  $[\alpha \dots]$  occurs in  $[\beta \dots [\alpha \dots]]$ , then according to T Val $([\alpha \dots]) = x$  when  $[\alpha \dots]$  occurs in  $[\gamma \dots [\alpha \dots]]$ .

The second additional constraint that a semantic theory illustrating the elegant strategy must satisfy is the constraint of *extensional compositionality*; a semantic theory must show how the truth conditions of a sentence are a *function of* the meanings of the words in the sentence. The Standard Theory is *extensionally compositional* because the Combinatorial Rules determine a *function* from the *extensions* of daughter nodes to the *extensions* of the immediately dominating mother nodes. Thus the semantic values appealed to in Davidson's semantic program are limited to *extensions*, as opposed to *intensions* of some sort, and this feature is often cited as the primary virtue of Davidson's semantic program. The constraint of *extensional compositionality* is stated more formally as follows:

Semantic theory T is extensionally compositional iff there is a function f such that for every node  $[\alpha[\beta,\ldots][\gamma,\ldots]]$  where  $[\gamma,\ldots]$  dominates immediate daughters  $[\beta,\ldots]$  and  $[\gamma,\ldots]$ , if according to T Val $([\beta,\ldots]) = x$  and Val $([\gamma,\ldots]) = y$  and Val $([\alpha,\ldots]) = z$ , then f(x,y) = z.

The elegant and explanatorily powerful strategy for explaining the semantic competence of ordinary speakers lies at the heart of Davidson's semantic program. This strategy is illustrated by standard Tarski-style formal truth theories that satisfy the constraints described above. Thus, Tarski-style truth theories that satisfy these constraints are theoretically motivated; i.e., it is clear how such formal theories provide at least first steps toward plausible explanations of our (in principle) infinite semantic competence. If a formal theory fails to satisfy one or more of these constraints, then it does not illustrate Davidson's strategy for explaining our semantic competence. Of course it does not *follow* that a formal theory that violates one or more of these constraints thereby fails to provide an adequate explanation of our semantic competence, but the question of just what explanatory strategy such a formal theory is meant to illustrate ought to be raised. If a formal truth theory is not illustrating how a speaker is

<sup>&</sup>lt;sup>8</sup> *Intensions* are typically functions from possible worlds to extensions in those worlds. For a classic presentation of this sort of program, see Montague (1974).

able to determine the truth conditions of an arbitrary sentence by computing a function from the fixed semantic values of terminal nodes to truth conditions, then what is the formal theory illustrating? If the formal theory does not illustrate some *other* plausible strategy for explaining our semantic competence, then it is a mere formal construction with no explanatory value. My objection against the ILF theories will be that not only do they violate one or more of the constraints imposed by Davidson's elegant explanatory strategy, but moreover they do not illustrate a plausible alternative explanatory strategy.

# 3. THE PROBLEM POSED BY ATTITUDE ASCRIPTIONS AND DAVIDSON'S PARATACTIC ANALYSIS

Standard Tarski-style truth theories, which satisfy all of the above constraints, cannot account for the truth conditions of attitude ascriptions. Or, more specifically, standard Tarski-style truth theories, which satisfy the constraints of *innocence* and *extensionally compositionality* do not entail theorems that state *correct* truth conditions for attitude ascriptions. Let us expand the tiny fragment of English under consideration by adding the noun phrase 'Odile', the complementizer 'that' and the transitive verb 'believes'. And let us expand the Standard Theory by adding the following rules:

```
 \begin{array}{ll} (LRe) & Val([_{NP}\ Odile]) = Odile \\ (LRf) & Val([_{V}\ believes]) = \{\langle x,y\rangle \colon x \ believes \ y\} \\ (LRg) & Val([_{C}\ that]) = \varnothing^{9} \\ \\ (CRd) & Val([_{CP}[_{C}\ that][_{S}\dots]]) = Val([_{S}\dots]) \\ \\ The \ difficulty \ is \ that \\ \end{array}
```

and

[S2[NP Clemens][VP[V wrote][NP Huck Finn]]]

 $[S_1[NP Twain][VP[V wrote][NP Huck Finn]]]$ 

<sup>&</sup>lt;sup>9</sup> The complementizer 'that' does not fall into any of the categories recognized by a Tarski-style truth theory. It is neither a predicate, nor a singular term, nor a functor, nor a connective, nor a quantifier. Thus it is treated as being semantically inert. It contributes no semantic value that is relevant to the truth conditions of sentences in which it occurs, but for technical reasons I assign it  $\varnothing$  as semantic value. Davidson's "paratactic" analysis rejects this treatment of 'that'.

will be assigned the same semantic value, *viz. true*. Consequently, because the Standard Theory is *innocent* and *extensionally compositional*, it will incorrectly assign the same semantic value to both

```
[S3[NP Odile][VP[V believes]CP
[C that][S1[NP Twain][VP[V wrote][NP Huck Finn]]]]].
```

and

 $[_{S4}[_{NP} \text{ Odile}][_{VP}[_{V} \text{ believes}]_{CP}[_{C} \text{ that}][_{S2}[_{NP} \text{ Clemens}][_{VP}[_{V} \text{ wrote}][_{NP} \text{ Huck Finn}]]]]].$ 

The problem is that because  $Val([NP Twain]) = Val([NP Clemens]), [S_1...]$ and [S2...] must contribute the same semantic value to the dominating [CP...] nodes, and thus in both sentences the same semantic values are "fed in" with the semantic value of [v believes] to determine the "molecular" semantic value of the dominating [VP[V] believes [CP...] node, and as a result the root nodes  $[s_3...]$  and  $[s_4...]$  must be assigned the same truth value. It seems undeniable that the semantic value of  $[_{NP}$  Twain] as  $[_{NP}$ Twain] occurs in *unembedded* [ $S_1$ ...] is Twain, and that the semantic value of [NP Clemens] as [NP Clemens] occurs in unembedded [S2...] is also Twain. But if this much is granted, then the constraint of *innocence* forces the claim that [NP Twain] and [NP Clemens] as they occur embedded in  $[s_3...]$  and  $[s_4...]$  have the same semantic value, viz. Twain. But now the constraint of extensional compositionality forces the incorrect result that  $[s_3...]$  and  $[s_4...]$  have the same truth conditions. (The same problem of course arises for any distinct content sentences [s...] and [s'...] that have the same truth value.)

How could a formal truth theory that has as theorems *correct* statements of the truth conditions of attitude ascriptions be formulated? One could solve the problem, i.e., formulate a truth theory that distinguishes the truth conditions of sentences such as  $[s_3...]$  and  $[s_4...]$ , by abandoning the constraint of extensional compositionality; e.g., one might deny that  $Val([c_P[C \text{ that}][s_1...]]) = Val([c_P[C \text{ that}][s_2...]])$ , even though  $Val([s_1...]) = Val([s_2...])$ . Abandoning *extensional compositionality* is the option taken by Larson and Ludlow (1993). Another option is to maintain extensional compositionality and abandon innocence; e.g. one might formulate a truth theory according to which  $[s_1...]$  (and/or  $[s_2...]$ ) is assigned different semantic values, depending upon what nodes dominate it – depending

 $<sup>^{10}</sup>$  Violating the constraint of *extensional compositionality* is also the option taken by Larson and Segal (1995).

upon whether or not it is embedded in the that-clause of an attitude ascription. This option is, more or less, what Frege (1892) proposes, and, as will be demonstrated below, this is also the option taken by Higginbotham (1986). Davidson's (1968) proposal for solving the problem, however, violates neither of these constraints.

Davidson's paratactic analysis is able to assign correct truth conditions to attitude ascriptions while satisfying the constraints of *innocence* and *simple compositionality* because Davidson denies that at the level of logical form attitude ascriptions are single sentences; according to Davidson's analysis, the ascription 'Odile believes that Twain wrote *Huck Finn*' is at the level of LF two sentences. The first has the LF

## $[_{S[NP]} Odile][_{VP}[_{V} believes][_{CP}[_{NP} that]]]]$

where the terminal [NP] that] node is a demonstrative that is assigned as a semantic value the *utterance* (or inscription) of the second sentence which instantiates the LF [ $s_1$ ...]. Both of these sentences can then be assigned truth conditions by an *innocent* and *extensionally compositional* theory, similar to the above-described Standard Theory. Because Davidson's proposed solution respects both *extensional compositionality* and *innocence* and it serves to illustrate the strategy for explaining our semantic competence.

The only way to solve the problem posed by attitude ascriptions within the constraints imposed by Davidson's semantic program is to adopt an analysis of attitude ascriptions along the lines of Davidson's paratactic analysis. If  $[s_3...]$  and  $[s_4...]$  are to be assigned distinct truth conditions, one must assign distinct values to  $[v_P[v]$  believes] $[c_P[c]$  that  $[s_2...]$ ] But given the constraints of *semantic innocence* and *extensional compositionality*, the only way that this can be accomplished is to posit an indexical feature within the  $[v_P[v]$  believes] $[c_P[c]$  that  $[s_2...]$ ] nodes that can be assigned distinct semantic values. Davidson suggests treating the [c] that [c] nodes as indexicals, but other proposals along these lines have been presented. Mark Richard (1990) has proposed treating the [c] believes nodes as indexicals, and Mark

<sup>&</sup>lt;sup>11</sup> Violating the constraint of *innocence* is also the option taken by Segal (1989), and Pietroski (1994).

An anonymous referee pointed out that the problem of attitude ascriptions could also be solved, or avoided, by denying the assumption that expressions *ever* have their "primary referents" as semantic values. This would in effect be to deny that, for example, "Twain' and 'Clemens' are coreferential. I will not here argue against this proposal, except to point out that the Davidson of "On Saying That" would clearly reject such a proposal: the proposal that words *always* give up their "pedastain references for the exotica" (1984, p. 108) is surely even more objectionable than Frege's claim that they *sometimes* do.

Crimmins (1992) has proposed (something like) treating the embedded [s...] nodes in their entirety as indexicals. Since the advocates of ILF theories must also reject these indexical analyses, and my central purpose here is to criticize the ILF analyses, I will not here present detailed arguments against the indexical analyses. I will, however, sketch the sorts of problems that indexical analyses face. <sup>14</sup>

First, there is the lexical problem that none of the relevant syntactic items are indexical in the appropriate way. For instance, with regard to Davidson's paratactic analysis, 'that' as it functions to introduce content sentences within attitude ascriptions is a complementizer, and not a referring NP.<sup>15</sup> And with regard to Richard's indexical theory, 'believes' and other propositional attitude verbs are not indexical in the appropriate way. Different occurrences of 'believes' do not designate different propositional attitude relations in the way that different occurrences of 'that' designate different objects.

Second, even granting that there is some relevant indexical feature in ordinary attitude ascriptions, there is the semantic problem that in making attitude ascriptions ordinary speakers could not be referring to, or quantifying over, the sorts of entities that must be posited by indexical analyses. All of the indexical analyses must in some way invoke appropriately individuated special entities, "modes of apprehension" or some such exotica, to explain the judgments of ordinary speakers concerning the truth conditions of ordinary attitude ascriptions. The problem is that, granting that such special entities exist, ordinary speakers are unable to identify and individuate these finely individuated special entities. Therefore such finely individuated special entities cannot plausibly be invoked to account for the semantic judgments of ordinary speakers concerning the truth conditions of ordinary attitude ascriptions. This second problem is simply a semantic version of the "problem of other minds."

The semantic problem arises directly for indexical theories such as Crimmins' "hidden indexical theory" (1992), which claim that token

<sup>&</sup>lt;sup>13</sup> The situation is more complicated than is suggested above because the analyses proffered by Richard and Crimmins assign "Russellian propositions" as the semantic values of sentences, and Richard's analysis violates *semantic innocence*. But despite these differences it is useful to view Richard's and Crimmins' analyses as utilizing the same general indexical strategy as Davidson's paratactic analysis.

<sup>&</sup>lt;sup>14</sup> In Clapp (1995) objections against Crimmins' (1992) "hidden indexical theory" are developed in detail.

<sup>&</sup>lt;sup>15</sup> Compelling arguments in support of this claim are presented in Higginbotham (1986), Burge (1986) and Segal and Speas (1986). But see Ian Rumfitt's (1993) for an ingenious version of Davidson's paratactic analysis of *oratio obliqua* that preserves the core idea of Davidson's analysis, yet avoids many of the difficulties faced by Davidson's analysis.

mental representations are somehow *referred* to by attitude ascriptions. Speakers cannot plausibly be interpreted as referring to entities they know nothing about. For example, I have no idea of what token "modes of presentation" Twain utilized in thinking about Socrates, or even if such entities existed. But despite my ignorance of such details concerning Twain's cognitive apparatus, I can understand, and engage in, discourse concerning Twain's beliefs concerning Socrates.

This semantic problem also arises, though less directly, for (extensions of) Davidson's paratactic analysis. Davidson initially applied his paratactic analysis to oratio obliqua, but it can perhaps be extended so that it also applies to attitude ascriptions. 16 According to Davidson's original proposal, an utterance of 'Odile said that Twain wrote Huck Finn' is true just in case the very embedded utterance token of 'Twain wrote Huck Finn' stands in the samesaying relation to some utterance made by Odile. Thus the relation of saying is analyzed in terms of the more basic samesaying relation and quantification over *utterances*. In order to apply this general paratactic strategy to an utterance of an *attitude* ascription such as  $[s_3...]$ , the relation of believing would have to be analyzed in terms of some more basic relation (sameRing) and quantification over some sort of appropriately individuated relata for this relation. That is, a paratactic analysis of  $[s_3...]$  would have to maintain that an utterance of  $[s_3...]$  is true if and only if the relevant utterance of the embedded token of  $[s_1...]$  stood in some analogy to the samesaying relation to some appropriately individuated relatum. There is an intimate relation between saying that p and the more basic relation of uttering an utterance, and it is plausible that this is a fact about the semantics of oratio obliqua. Roughly put, the truth of an utterance of the form 'X said that p' semantically presupposes the truth of an utterance of the form 'X uttered some utterance'. Consequently it is at least plausible for Davidson to claim that an utterance of the form 'X said that that p' is true only if X uttered a certain sort of utterance. But there is no analogous more basic relation in the case of belief, or regret, or hope, etc. The truth of an utterance of the form 'X regretted that p' does not require that X uttered an utterance. Moreover, in the case of regretting there is no other more basic relation to play the role of *uttering*, and there is no corresponding sort of relata to play the role of utterances; roughly put, there is no R such that regretting that p semantically presupposes Ring an R. And similarly for belief and the other propositional attitudes.

Of course one could *posit* some relation to play the role of the *uttering* relation, (perhaps "regret\*" or "REG") and one could *posit* some sort

<sup>&</sup>lt;sup>16</sup> A means of extending Davidson's paratactic analysis so that it applies to attitude ascriptions is proposed in Lepore and Loewer (1989, pp. 352–4).

of special appropriately individuated entity to play the role of *utterances* (perhaps "sentences of mentalese" or some such mental entities). But now it is apparent that paratactic analyses of *attitude ascriptions* also face the semantic version of the problem of other minds. Even granting that an appropriate relation – such as "REG" – exists, and that some sort of special appropriately individuated relata – such as "sentences of mentalese" – exist, such relations and relata are not within the ken of ordinary speakers. Thus it is implausible that in uttering attitude ascriptions speakers are (perhaps implicitly) attempting to utter embedded sentences that stand in some sort of *sameRing* relation to *some* such finely individuated mental entity. Semantic competence with regard to attitude ascriptions does not require detailed and highly theoretical knowledge (even implicit knowledge) of the processes underlying one another's psychological states.

Indexical proposals such as Davidson's paratactic analysis preserve the elegant and explanatorily powerful picture of our semantic competence, and that is why they are touted as being superior to Fregean analyses, which violate the constraint of *semantic innocence*. But in attempting to fit attitude ascriptions into this explanatory picture, they are forced to make implausible syntactic and semantic assumptions.

#### 4. LARSON AND LUDLOW'S ILF ANALYSIS

Larson and Ludlow claim that their theory can be "embed[ded] within a recursive theory of truth for natural language of the kind advocated by Davidson" (Larson and Ludlow, 1993, p. 305). This claim is correct in as much as their theory is a finitely axiomatized truth theory, but, as I will demonstrate here, Larson and Ludlow's theory deviates in a significant way from standard Tarski-style truth theories: Larson and Ludlow's ILF analysis, though it has as theorems statements of correct truth conditions for attitude ascriptions, violates the constraint of *extensional compositionality*. Thus Larson and Ludlow's ILF analysis violates constraints that Davidson himself took pains to satisfy.

<sup>17</sup> Even Rumfitt's (1993) revised paratactic analysis of *oratio obliqua* cannot be extended to apply to attitude ascriptions generally. On Rumfitt's version of the paratactic theory of *oratio obliqua* the indexical element refers, via deferred ostension, to an *act type* of making a certain type of utterance. Thus an utterance of sentence of the form 'Galileo said that the earth moves' is true if and only if Galileo performed a token act of uttering of the same type as the just performed action of uttering 'the earth moves'. The difficulty with extending this proposal to, e.g., regret ascriptions, is that there is no specific sort of action that accompanies *regretting* in the way that *uttering* accompanies *saying*.

<sup>&</sup>lt;sup>18</sup> More specifically, Larson and Ludlow's ILF analysis violates an appropriately amended version of *extensional compositionality*. Larson and Ludlow formulate their the-

Before considering Larson and Ludlow's ILF theory, it will be useful to consider a simpler ILF theory that is merely an amended version of the Standard Theory. The general strategy of the Simple ILF Theory is to assign distinct semantic values to the [CP[C] that  $[S_1...]$  node of  $[S_3...]$  and the [CP[C] that [S2...] of [S4...]. For if this can be accomplished, then the root [S3...] and [S4...] nodes could be assigned distinct semantic values as a function of their immediate daughter nodes. From a purely semantic point of view it does not matter what the semantic values assigned to the [CP[C] that [S...] nodes are, so long as they are individuated finely enough so that appropriate truth conditions are assigned. For example, one could simply identify the semantic value of an embedded [ $_{CP}[_{C}]$  that][ $_{S...}$ ]] node with that very syntactic subtree. Higginbotham even vaguely suggests that "the objects of the attitudes, at least when indicated by clauses, are objects similar to the interpreted structures of the clauses themselves" (1986, p. 39). Another option would be to order all the [CP[C] that [S...]nodes, say alphabetically, and assign each such ordered [CP[C] that][S...]node the corresponding natural number; e.g., if 'that Twain wrote Huck Finn' appears as the 30,324th entry on the alphabetically ordered list of that-clauses, then Val([CP[C that][S1...]]) = 30,324, and thus

$$Val([s_3...])$$
 is *true* iff  $\langle Odile, 30,324 \rangle \in \{\langle w, v \rangle: w \text{ believes } v\}.^{19}$ 

Larson and Ludlow, however, opt for a different kind of semantic value. They maintain that the [CP[C that][S...]] nodes of attitude ascriptions are to be assigned the *interpreted logical form* (ILF) which is determined by those nodes. The ILF determined by a node is obtained by replacing every node of the LF with the ordered pair whose first member is the syntactic item (terminal node or subtree), and whose second member is the semantic value of that syntactic item. (Bear in mind that a nonterminal node is a *tree*, an LF.) The ILF of an LF is then an isomorphic tree structure whose nodes are occupied by ordered pairs of syntactic items and their semantic

ory in a non standard way, and as a result the principle of *extensional compositionality* must be slightly amended so that it applies to their theory. My objections against Larson and Ludlow's (1993) ILF theory also apply to Larson and Segal's (1995) ILF theory.

<sup>&</sup>lt;sup>18</sup> I see no reason why the suggested "numerical analysis" is any less plausible than the ILF analyses: the only motivation for identifying the semantic values of 'that'-clauses with ILFs is that doing so yields correct statements of truth conditions for attitude ascriptions, but the numerical analysis would also yield these. (Of course ambiguous content sentences such as 'Everybody loves somebody' require that amendments be made to the numerical suggestion. But 'that'-clauses containing demonstratives and indexicals require that ad hoc amendments be made to the ILF analysis.)

values.<sup>20</sup> But, to reiterate, it does not matter what ILFs are taken to be, so long as ILFs are appropriately individuated so the theory yields correct statements of the truth conditions for attitude ascriptions.

How can this appeal to ILFs distinguish the truth conditions of  $[s_3...]$  and  $[s_4...]$ ? Let  $\mathfrak{I}()$  be a function from LFs to the unique ILFs they determine. Thus  $\mathfrak{I}([s_1...]) \neq \mathfrak{I}([s_2...])$  because  $\mathfrak{I}([s_1...])$  contains  $[s_n]$  Twain] where  $\mathfrak{I}([s_2...])$  contains  $[s_n]$  Clemens]. And because  $\mathfrak{I}([s_1...]) \neq \mathfrak{I}([s_2...])$  the problem of distinguishing the truth conditions assigned to  $[s_3...]$  and  $[s_4...]$  can seemingly be solved rather easily. All that needs to be done is to replace (CRd) in the Standard Theory with,

$$(CRD') Val([CP[C that][S...]]) = \Im([S...])$$

If this substitution is made, then the Simple ILF theory yields the following theorems:

Val(
$$[s_3...]$$
) = true iff  $\exists x (x = Odile & x \in \{y: \exists z (z = \Im([s_1...]))$ and  $\langle y, z \rangle \in \{\langle w, v \rangle: w \text{ believes } v\}\}$ ).

Val(
$$[s_4...]$$
) = true iff  $\exists x (x = Odile \& x \in \{y: \exists z (z = \Im([s_2...]))$ and  $\langle y, z \rangle \in \{\langle w, v \rangle: w \text{ believes } v\}\}$ ).

Thus, because  $\mathfrak{I}([s_1...]) \neq \mathfrak{I}([s_2...])$ ,  $[s_3...]$  and  $[s_4...]$  will be assigned distinct truth conditions.

The Simple ILF Theory sketched above, however, does not solve the problem posed by attitude ascriptions within the constraints imposed by Davidson's semantic program. The difficulty is that the Simple ILF Theory above does not respect *extensional compositionality*: under the rules of the

And, for example, the ILF of the node [VP[V] wrote][NP] is the following tree structure:

$$<[_{VP}[_{VWTOte}][_{NP}Huck\ Finn]], \ \{x: x\ wrote\ Huck\ Finn\}> \\ / \\ <[_{VWTOte}], \ \{: x\ wrote\ y\}> \\ <[_{NP}Huck\ Finn], \ Huck\ Finn> \\$$

<sup>&</sup>lt;sup>20</sup> For example, the ILF of the root [ $_{S...}$ ] node of [ $_{S1}$ [ $_{NP}$  Twain][ $_{VP}$ [ $_{V}$  wrote][ $_{NP}$  Huck Finn]]] is the following tree structure:

Simple ILF Theory  $Val([CP[C] that][S_1...]) \neq Val([CP[C] that][S_2...])$ , even though  $Val([S_1...]) = Val([S_2...])$ , and Val([C] that]) = Val([C] that]). Therefore, though the Simple ILF Theory has as theorems *correct* statements of the truth conditions of  $[S_3...]$  and  $[S_4...]$ , it blatantly violates the constraint of *extensional compositionality*.

But why is it problematic that the Simple ILF Theory violates *extensional compositionality*? A defender of the Simple ILF Theory might concede that an adequate semantic theory must be compositional in some sense, yet deny that satisfying *extensional compositionality* is necessary. For, though the Simple ILF Theory violates *extensional compositionality*, it satisfies other, weaker, principles of compositionality. Why is satisfying some weaker principle of compositionality not sufficient? Let us consider some (relatively plausible) weaker principles of compositionality that the Simple ILF Theory satisfies:

A Tarski-style truth theory T is *vacuously compositional iff* for every node  $[\alpha[\beta,\ldots][\gamma,\ldots]]$  where  $[\alpha,\ldots]$  dominates immediate daughters  $[\beta,\ldots]$  and  $[\gamma,\ldots]$ , *if* according to T Val( $[\beta,\ldots]$ ) = X and Val( $[\gamma,\ldots]$ ) = X and Val( $[\alpha,\ldots]$ ) = X and  $[\alpha,\ldots]$  =  $[\alpha,\ldots$ 

A Tarski-style truth theory T is *lexically compositional iff* there is a function f such that for every node  $[\alpha[\beta,\ldots][\gamma,\ldots]]$  where  $[\alpha,\ldots]$  dominates immediate daughters  $[\beta,\ldots]$  and  $[\gamma,\ldots]$ , *if* according to T Val $([\alpha,\ldots]) = z$ , then  $f(\langle [\beta,\ldots], [\gamma,\ldots] \rangle) = z$ .

A Tarski-style truth theory T is *hybrid compositional* iff there is a function f such that for every node  $[\alpha[\beta...][\gamma...]]$  where  $[\alpha...]$  dominates immediate daughters  $[\beta...]$  and  $[\gamma...]$ , *if* according to T Val( $[\beta...]$ ) = x and Val( $[\gamma...]$ ) = y and Val( $[\alpha...]$ ) = z, then  $f(\langle\langle x, [\beta...] \rangle, \langle y, [\gamma...] \rangle\rangle)$  = z.

The Simple ILF Theory satisfies all of these compositionality principles (and many other, more bizarre, principles). Why is satisfying one of these weaker compositionality principles not sufficient? Consider first *vacuous compositionality*. This principle is so-called because it is tautological, and therefore places no constraints whatsoever on Tarski-style truth theories. To see this, consider a ridiculous Tarski-style truth theory according to which Val([NP] Twain]) = Val([NP] Clemens]) = Twain, and  $Val([VP] VWOTE)[NP] Huck Finn]]) = \{y: \exists z \ (z = Huck Finn \& \langle y, z \rangle \in \{\langle v, w \rangle: v \ wrote \ w\})\}$ , and  $Val([S_1...]) = true$ , though  $Val([S_2...]) = false$ . This Tarski-style truth theory is ridiculous because it assigns distinct truth values to  $[S_1...]$ 

and  $[s_2...]$ , though it assigns the same semantic values to their parts. But despite this the ridiculous theory satisfies *vacuous compositionality*. There is *some* function f such that

$$f(\langle \text{Twain}, \{y: \exists z \ (z = \text{Huck Finn \& } \langle y, z \rangle \in \{\langle v, w \rangle: v \text{ wrote } w\})\}\rangle) = \text{true}$$

and some (other) function  $f^*$  such that

$$f^*(\langle \text{Twain}, \{y: \exists z \ (z = Huck \ Finn \ \& \langle y, z \rangle \in \{\langle v, w \rangle: v \ \text{wrote } w\})\}\rangle) = false$$

Does it *follow* from the fact that vacuous compositionality is tautological that it is not sufficiently strong? No, it does not follow. The goal is to explain our semantic competence. If such an explanation could be illustrated by a Tarski-style truth theory that violates every remotely plausible compositionality constraint except for *vacuous compositionality*, then so much the worse for non-vacuous compositionality constraints. So why could the Simple ILF Theory not be defended on the grounds that it satisfies *vacuous compositionality* and *vacuous compositionality* is compositionality enough? The problem is that it is not at all clear what explanatory strategy the Simple ILF Theory is illustrating in virtue of being *vacuously compositional*. That is, it is not at all clear what explanatory strategy supports *vacuous compositionality* in the way that Davidson's elegant and powerful explanatory strategy supports *extensional compositionality*.

Suppose that the Simple ILF Theory is defended on the grounds that it satisfies vacuous compositionality. It is therefore allowed that the semantic value of a dominating mother node need not be a function of the semantic values assigned to its immediate daughter nodes, and consequently there are no longer any theoretical constraints on what should be assigned as semantic values to daughter nodes, including lexical nodes. So it does not matter what semantic values are assigned to lexical nodes. Indeed, since it does not matter what values are assigned to lexical nodes, there is no reason to assign semantic values to lexical nodes at all. Let us extrapolate from the Simple ILF Theory, and consider a semantic theory according to which all transitive verbs designate relations between agents and ILFs. Hence, just as the belief relation is taken by the ILF Theory to be a (rather contrived) relation between agents and ILFs, on the extrapolated theory the wrote relation is identified with a relation between agents and ILFs: just as x holds the belief relation toward ILF y iff, roughly, x believes the proposition expressed by the LF corresponding to ILF y, so x holds

the *wrote* relation toward ILF y iff, roughly, x wrote the thing designated by the LF corresponding to the ILF y. In this way attitude ascriptions are covered by the same general rule which applies to all transitive verbs, viz.

(CRb") 
$$Val([v_P[v...][N_{P/CP}...]]) = \{x: \exists y (y = \Im([N_{P/CP}...] \& \langle x, y \rangle \in Val([v...]))\}^{21}$$

If the constraint of *extensional compositionality* is maintained – as it must be if a Tarski-style truth theory is to illustrate Davidson's elegant explanatory strategy – then this ILF infected theory can be rejected on the grounds that it is not *extensionally compositional*. But if *extensional compositionality* is abandoned in favor of *vacuous compositionality* then we have no reason to reject this ILF infected theory. Indeed, parsimony dictates that the proponent of the Simple ILF Theory endorse the proposed ILF infected theory, as it contains fewer combinatorial rules.

But if the combinatorial rule (CRb") is accepted, then why look to nodes dominated by [NP...] and [NP...] nodes at all? That is, if parsimony compels us to identify the semantic values of the objects of transitive verbs with ILFs, then parsimony would also dictate that the Simple ILF Theory be replaced by a theory that contains only the following ILF infested combinatorial rule:

(CRa\*) Val(
$$[s[NP...][VP...]$$
) = true iff  $\Re(\Im([NP...]), \Im([VP...]))$ .

(where  $\mathfrak{R}(.)$  is a, rather contrived, relation which holds between  $?([_{NP}...])$ , and  $\mathfrak{I}([_{NP}...])$ ) iff, roughly, the referent of  $[_{NP}...]$  has the property designated by  $([_{VP}...])$ . But the formal semantic theory wholly comprised by the rule (CRa\*) is not significantly different from the Absurdly Simple Theory whose only rule is

(CRa) 
$$Val([_S[_{NP}...][_{VP}...]]) = true \text{ iff } \exists x \ (x = Val([_{NP}...]) \& x \in Val([_{VP}...])).$$

<sup>21</sup> It might be objected that this cannot be a proper analysis of transitive verbs, because, e.g., the things that are *written* are books, papers, etc., and *not* ILFs. I think that this objection has some force, but note that it applies with equal force against the ILF analysis of attitude ascriptions: If 'Odile believes that snow is white' can be analyzed as asserting that Odile stands in the believes relation to the ILF determined by 'Twain smoked in virtue of this ILF bearing some further relation to Odile's mental state, then why not analyze' 'Twain wrote *Huck Finn*' as asserting that Twain stands in the wrote relation to the ILF determined by '*Huck Finn*' in virtue of this ILF bearing some further relation to *Huck Finn*?

 $<sup>^{22}\,</sup>$  Again, one might object that  $\mathfrak{R}(,)$  is a rather contrived relation. But it is no more contrived than the relation alleged to hold between believers and ILFs.

However, as was shown in section I, the Absurdly Simple Theory fails to illustrate an adequate explanation of a speaker's semantic competence with regard to the tiny fragment. A non-English speaker could know (CRa\*), and yet not be able to "effectively determine the meaning or meanings of an arbitrary expression" within the fragment.

The defender of the Simple ILF Theory might respond to this difficulty by simply stipulating that, though the semantic values assigned to mother nodes need not be a function of the semantic values assigned to daughter nodes, nonetheless an adequate formal semantic theory must somehow assign semantic values to lexical nodes. The problem with this response is that it is wholly unmotivated. It is undoubtedly correct that a competent speaker is able to "effectively determine the meaning or meanings of an arbitrary expression" in part because he knows the meanings of individual words. Davidson's semantic program, which endorses standard, extensionally compositional, Tarski-style truth theories to illustrate the elegant and powerful explanatory strategy, is not only compatible with this truism; it explains this truism. But if extensional compositionality is rejected in favor of vacuous compositionality, then simply stipulating that an adequate formal semantic theory must assign semantic values to lexical nodes is gratuitous. If extensional compositionality is rejected, then Davidson's elegant explanatory strategy is rejected along with it. But then it is not at all clear why a hearer needs to know the meaning of the words in a sentence to comprehend the sentence.

The general problem then with defending the Simple ILF Theory by rejecting extensional compositionality in favor of vacuous compositionality is this: Because the Simple ILF Theory violates extensional compositionality it fails to illustrate Davidson's explanatory strategy. Moreover, the fact that Simple ILF Theory satisfies vacuous compositionality would be theoretically significant only if it were shown that an alternative plausible strategy for explaining our semantic competence supports vacuous compositionality in the way that Davidson's explanatory strategy supports extensional compositionality. But it is not all clear what this alternative strategy could be. In the absence of such an alternative explanatory strategy, the Simple ILF Theory is a mere formal theory, and it is not clear how it contributes to explaining our semantic competence.

In response to this, the defender of the Simple ILF Theory might concede that *vacuous compositionality* is too weak, yet claim that *hybrid compositionality* is sufficiently strong.<sup>23</sup> There are two problems with this defense of the Simple ILF Theory.

<sup>&</sup>lt;sup>23</sup> I here focus on *hybrid compositionality*, as opposed to *lexical compositionality*, because *lexical compositionality* would be too restrictive for any language with indexicals.

First, this defense merely issues a promissory note. Unless and until a plausible alternative strategy for explaining our semantic competence is formulated, the fact that the Simple ILF Theory satisfies *hybrid compositionality* is no more significant than is the fact that it satisfies *vacuous compositionality*.

Second, if extensional compositionality is rejected in favor of hybrid compositionality, then the motivation for formulating semantic theories as Tarski-Style truth theories is undermined. Moreover, the explanatory strategy that is vaguely suggested by hybrid compositionality would not be properly illustrated by a (non-standard) Tarski-style truth theory. The central idea of Davidson's semantic program is "that a satisfactory theory of meaning must give an account of how the meanings of sentences depend upon the meanings of words" (Davidson, 1984, p. 17). Slightly more precisely, at the core of Davidson's semantic program lies the following elegant and powerful strategy for explaining semantic competence: Upon perceiving an utterance a competent subject is able to determine the logical form of that utterance. The subject is able to understand the utterance because he knows the semantic values of the words and he knows how these semantic values "combine," as directed by the logical form of the sentence, to determine the conditions under which the sentence is true. Thus Davidson utilizes an explanatory strategy according to which, roughly, the truth conditions of a sentence are a function of only (i) the structure, or logical form, of the sentence, and (ii) the extensions of the words appearing in the sentence. Tarski-style truth theories are perfectly suited to illustrate this explanatory strategy because they display via a finite number of axioms how the truth conditions of sentences are a function of only these two factors.

The explanatory strategy vaguely suggested by the principle of *hybrid compositionality*, however, is that comprehension of an utterance of a declarative sentence is a function of not only (i) the structure, or logical form of the sentence, and (ii) the extensions of the words appearing in the sentence, but also (iii) those very words themselves. Moreover, the words that contribute to the truth conditions of a sentence need not themselves be the extensions (semantic values) of any words in the sentence. Thus *hybrid compositionality* suggests that an utterance does not merely fix semantic values and the order in which combinatorial rules will be applied to these semantic values. Rather in at least some cases phonological features of the utterance itself are relevant to determining the truth conditions of the utterance. Our goal is a general explanatory strategy that makes clear *how* 

Cases similar to Kripke's "Paderewski" cases raise additional problems, but these issues are beyond my concerns here.

and *why* such features, which are *not* the semantic values of any of the words in the sentence, are relevant to determining truth conditions. The point of presenting a *formal* semantic theory would be to illustrate this general explanatory strategy.

The second problem for the Simple ILF Theory is that because it is formulated as a Tarski-style theory, it is ill suited for this task. For Tarski-style truth theories illustrate, roughly, how the truth conditions of a sentence are a function of only (i) the structure, or logical form of the sentence, and (ii) the extensions of the words appearing in the sentence. The Simple ILF Theory manages to satisfy the constraint of *correctness* (for the tiny fragment) because, via (CRd'), the embedded words (perhaps relevant phonological features thereof) become parts of the semantic values of [CP...] nodes. But no explanation is provided as to how or why these words get into the semantic values of [CP...] nodes; they are not the semantic values of terminal nodes, and they are not determined by the semantic values of the terminal nodes. (It is because of this latter fact that the Simple ILF Theory violates extensional compositionality.) Hence Tarski-style truth theories are ill suited to illustrate the sort of general explanatory strategy that would support hybrid compositionality. This brief discussion by no means establishes that Tarski-style truth theories are inadequate for illustrating plausible explanations of our semantic competence, but it does justify a demand for a defense of Tarski-style theories: If extensional compositionality is rejected in favor of hybrid compositionality, then what motivation is there for formulating semantic theories as (even non-standard) Tarski-style truth theories?

It remains for me to demonstrate that Larson and Ludlow's ILF analysis also violates the constraint of *extensional compositionality*, and thus is no better off than is the Simple ILF Theory. That Larson and Ludlow's analysis also violates *extensional compositionality* is somewhat difficult to discern because the semantic rules of their theory are formulated in a nonstandard way. Apparently Larson and Ludlow want to avoid assigning sets as the semantic values of predicates, and so in formulating their theory they eschew *functions* such as Val() and instead utilize a *binary relation* SV(,) that applies to a pair  $\langle x, [\alpha...] \rangle$  iff x is a semantic value of  $[\alpha...]$ . Thus instead of assigning as *the* semantic value of a predicate  $[\alpha...]$  the *set* of things that satisfy  $[\alpha...]$ , Larson and Ludlow assign as *a* semantic value of  $[\alpha...]$  any thing that satisfies  $[\alpha...]$ . So, for example, in Larson and Ludlow's theory *anybody* who runs is a semantic value of [vP] runs and anybody who believes that Twain runs is a semantic value of [vP] runs believes [sS] [NP] Twain][vP] runs]]].

Simplifying somewhat, Larson and Ludlow's semantic theory is as follows:

## **Larson and Ludlow Lexical Rules:**

```
(LRa') SV(x, [NP Odile]) iff x = Odile
```

(LRb') 
$$SV(x, [NP Twain])$$
 iff  $x = Twain$ 

(LRc') 
$$SV(x, [NP Clemens])$$
 iff  $x = Twain$ 

(LRe') 
$$SV(\langle x, y \rangle, [v \text{ believes}])$$
 iff x believes y

## **Larson and Ludlow Combinatorial Rules:**

(CRa') SV(true, 
$$[S[NP...][VP...]$$
) iff  $\exists x$  (SV(x,  $[NP...]$ ) & SV(x,  $[VP...]$ ))

(CRb') 
$$SV(x, [v_P[v \text{ believes}][s...]) \text{ iff}$$
  
 $\exists y(SV(\langle x, y \rangle, [v \text{ believes}]) \& y = \Im([s...]))$ 

(Note that Larson and Ludlow take the complement of [v] believes] to be an [S...] node, and not a [v] node. I follow them in making this assumption to stay as close as possible to Larson and Ludlow's formulation of their rules.)

Let us now use Larson and Ludlow's theory to determine the truth conditions of

[S6[NP Odile][VP[V believes][S5[NP Twain][VP runs]]]]

and

$$[_{S8}[_{NP} \text{ Odile}][_{VP}[_{V} \text{ believes}][_{S7}[_{NP} \text{ Clemens}][_{VP} \text{ runs}]]]]$$

To solve the problem of attitude ascriptions, the theory must assign distinct truth conditions to these sentences. First, consider  $[s_6...]$ . By (CRa'),

1.  $SV(true, [_{S6}[_{NP} Odile][_{VP}[_{V} believes][_{S5}[_{NP} Twain][_{VP} runs]]]] iff$  $\exists x (SV(x, [_{NP} Odile]) & SV(x, [_{VP}[_{V} believes][_{S5}[_{NP} Twain][_{VP} runs]]])))$ 

By 1, (LRa') and (CRb'),

2.  $SV(true, [_{S6}[_{NP} Odile][_{VP}[_{V} believes][_{S5}[_{NP} Twain][_{VP} runs]]]] iff$  $\exists x (x = Odile & <math>\exists y(SV(\langle x, y \rangle, [_{V} believes]) & y = \Im([_{S5}[_{NP} Twain][_{VP} runs]])))^{24}$ 

 $<sup>^{24}</sup>$  Because Larson and Ludlow do not assign unique semantic values to nodes, they must adopt a more complicated "construction algorithm" for ILFs. That is, because on

And from 2 and LRe' it follows that

3.  $SV(true, [_{S6}[_{NP} Odile][_{VP}[_{V} believes][_{S5}[_{NP} Twain][_{VP} runs]]]]$  iff Odile believes  $\Im([_{S5}[_{NP} Twain][_{VP} runs]]))$ 

Now consider [ $s_8...$ ]. By (CRa'),

4.  $SV(true, [_{S8}[_{NP} Odile][_{VP}[_{V} believes][_{S7}[_{NP} Clemens][_{VP} runs]]]]$  iff  $\exists x \ (SV(x, [_{NP} Odile]) \& SV(x, [_{VP}[_{V} believes][_{S7}[_{NP} Clemens][_{VP} runs]]])))$ 

By 4, (LRa'), and (CRb'),

5.  $SV(true, [_{S8}[_{NP} Odile][_{VP}[_{V} believes][_{S7}[_{NP} Clemens][_{VP} runs]]]] iff$  $\exists x (x = Odile & <math>\exists y(SV(\langle x, y \rangle, [_{V} believes])) & y = \Im([_{S7}[_{NP} Clemens][_{VP} runs]]))$ 

And from 5 and LRe' it follows that

6.  $SV(true, [_{S8}[_{NP} Odile][_{VP}[_{V} believes][_{S7}[_{NP} Clemens][_{VP} runs]]]]$  iff Odile believes  $\Im([_{S7}[_{NP} Clemens][_{VP} runs]]))$ 

Assuming that Odile can hold the attitude of belief toward the ILF  $\mathfrak{I}([s_7[NP Clemens]]_{VP} runs]])$ , yet not the ILF  $\mathfrak{I}([s_5[NP Twain]]_{VP} runs]])$ , Larson and Ludlow's theory succeeds in distinguishing the truth conditions of  $[s_6...]$  and  $[s_8...]$ . (And I assume that a slightly more sophisticated version of the theory could distinguish the truth conditions of  $[s_3...]$  and  $[s_4...]$ ).

The constraint of *extensional compositionality* has been formulated in terms of the Val() function, and consequently it does not even apply to Larson and Ludlow's semantic theory, which utilizes the binary relation SV(,). So before I can demonstrate that Larson and Ludlow's theory violates the constraint of *extensional compositionality* I must reformulate the constraint so that it applies to Larson and Ludlow's semantic theory; i.e. I must state the constraint in terms of the binary relation SV(,).

their theory nodes do not have unique semantic values, the ILF of an LF cannot be a structure of ordered pairs, the second member of which is *the* semantic value of the first member. For Larson and Ludlow's "construction algorithm" see Larson and Ludlow *ibid.*, p. 312. I overlook the details of ILF construction because, as I have intimated above, it does not much matter what ILFs are, so long as (i) LFs determine ILFs, and (ii) ILFs are appropriately individuated.

Hence these amended versions of *semantic innocence* and *extensional compositionality*:

Let  $[\beta \dots [\alpha \dots] \dots]$  and  $[\gamma [\alpha \dots] \dots]$  be any distinct syntactic structures that dominate occurrences of a smaller syntactic structure  $[\alpha \dots]$ . Semantic theory T is  $innocent^*$  iff if according to T SV( $\langle x, [\alpha \dots] \rangle$ ) when  $[\alpha \dots]$  occurs in  $[\beta \dots [\alpha \dots] \dots]$ , then according to T SV( $\langle x, [\alpha \dots] \rangle$ ) when  $[\alpha \dots]$  occurs in  $[\gamma [\alpha \dots] \dots]$ .

Semantic theory T is extensionally compositional\* iff there is a function f such that for every node  $[\alpha[\beta,\ldots][\gamma,\ldots]]$  where  $[\alpha,\ldots]$  immediately dominates  $[\beta,\ldots]$  and  $[\gamma,\ldots]$ ,  $f(\langle \{x: according to <math>T \ SV(\langle x,[\beta,\ldots]\rangle)\}, \{y: according to <math>T \ SV(\langle y,[\gamma,\ldots]\rangle)\}) = \{z: according to <math>T \ SV(\langle z,[\alpha,\ldots]\rangle)\}.$ 

Larson and Ludlow's ILF analysis does not satisfy *extensional compositionality*\*. Suppose that  $[_{S6}[_{NP} \text{ Odile}][_{VP}[_{V} \text{ believes}][_{S5}[_{NP} \text{ Twain}][_{VP} \text{ runs}]]]]$  is true, and that  $[_{S8}[_{NP} \text{ Odile}][_{VP}[_{V} \text{ believes}][_{S7}[_{NP} \text{ Clemens}][_{VP} \text{ runs}]]]]$  is false. So (by (LRa'), (CRa') and (CRb')) according to Larson and Ludlow's ILF analysis Odile is a member of  $\{z: SV(\langle z, [_{V} \text{ believes}][_{S5}\dots]]\rangle)\}$ . So according to Larson and Ludlow's analysis  $\{z: SV(\langle z, [_{V} \text{ believes}][_{S5}\dots]]\rangle)\}$   $\neq \{z: SV(\langle z, [_{V} \text{ believes}][_{S7}\dots]]\rangle)\}$ . But according to their analysis  $\{x: SV(\langle x, [_{V} \text{ believes}]\rangle)\} = \{\langle x, y \rangle: x \text{ believes }y\}$ . And also  $\{y: SV(\langle y, [_{S5}\dots]\rangle)\} = \{y: SV(\langle y, [_{S7}\dots]\rangle)\} = \{true\}$ . So Larson and Ludlow's ILF analysis violates *extensional compositionality*\*. There is no function f such that

```
f(\langle \{\langle x, y \rangle: x \text{ believes } y\}, \{true\} \rangle) = \{z: SV(\langle z, [v \text{ believes}][s_5...]] \rangle)\}
```

and

```
f(\langle \{\langle x, y \rangle: x \text{ believes } y\}, \{true\} \rangle) = \{z: SV(\langle z, [v \text{ believes}][s_7...]] \rangle)\}
```

because, to repeat, according to Larson and Ludlow's ILF analysis  $\{z: SV(\langle z, [_V \text{ believes}][_{S5}\dots]]\rangle)\} \neq \{z: SV(\langle z, [_V \text{ believes}][_{S7}\dots]]\rangle)\}$ . So Larson and Ludlow's ILF analysis violates *extensional compositionality*\*. <sup>25</sup>

Larson and Ludlow's theory, like the Simple ILF Theory, abandons the idea that truth conditions of a sentence are a function of its syntactic structure, or LF, and the meanings of the words in the sentence, and consequently Larson and Ludlow's ILF analysis no more illustrates a plausible strategy for explaining our semantic competence than does the Simple ILF Theory.

#### 5. HIGGINBOTHAM'S FREGEAN ILF THEORY

Higginbotham's ILF proposal takes the option of violating *innocence*, yet preserving a relativized version of *extensional compositionality*. The essential idea behind Higginbotham's proposal is that Frege's proposal — where ILFs will play the role played by Frege's *thoughts* — be implemented within the framework of a Tarski-style truth theory. Higginbotham implements semantic value shifting within a Tarski-style truth theory by replacing the one-place Val() function used in the Standard Theory with a *two-place* function, Val\*(,), from nodes and dominating (or identical) nodes to semantic values.  $^{26}$  ("Val\*([ $\alpha$ ...],[ $\beta$ ...])" is to be read, "the se-

If  $\alpha$  is an expression,  $\beta_1, \beta_2, \ldots, \beta_n$  are immediate constituents of  $\alpha$ , and for some  $x, y_1, y_2, \ldots, y_n$ ,  $[SV](x, \alpha)$ ,  $[SV](y_1, \beta_1)$ ,  $SV(y_2, \beta_2)$ , ...  $[SV](y_n, \beta_n)$ , then, for some function f,  $x = f(y_1, y_2, \ldots, y_n)$ , (Larson and Ludlow, 1993, p, 344–5).

This principle, however, is also vacuous and thus imposes no constraints on a semantic theory whatsoever. I suspect that Larson and Ludlow intended their compositionality principle to be equivalent to *extensional compositionality*. But because they formulate their theory in a non standard way, they failed to formulate the principle correctly, and consequently they overlooked the fact that their ILF analysis violates *extensional compositionality*.

<sup>26</sup> I am simplifying here. Higginbotham actually employs the relation x is a semantic value of  $[\alpha \dots]$  relative to  $[\beta \dots]$ . Strictly speaking, the sorts of modifications, in the definitions of semantic innocence and extensional compositionality, introduced in the last section need to be carried over to the discussion of Higginbotham. To keep matters maximally simple, I have reformulated Higginbotham's view so that (for example) relative to  $[\alpha \dots]$ ,  $[\beta \dots]$  receives a unique semantic value. Nothing substantive hangs on this.

Thus what for Frege is "the primary referent" of an expression is for Higginbotham the value of  $Val^*(,)$  where the same expression fills both argument places; e.g., the primary semantic value of a root sentence node [s...] is *true iff*  $Val^*([s...], [s...]) = true$ . Higgin-

 $<sup>^{25}</sup>$  In a footnote Larson and Ludlow assert that their theory satisfies the following principle of compositionality:

mantic value of  $[\alpha...]$  relative to  $[\beta...]$ .") Thus, where  $[\beta...]$  either dominates or is identical to  $[\alpha...]$ , and  $[\gamma...]$  dominates  $[\alpha...]$  and is not identical with  $[\beta...]$ , it is possible that  $Val^*([\alpha...], [\beta...]) = x$  while  $Val^*([\alpha...], [\gamma...]) = y$  (where  $x \neq y$ ). In particular, Higginbotham's proposal allows that  $Val^*([CP[C] that][S_1...])$ ,  $[CP[C] that][S_1...]) = true$ , while  $Val^*([CP[C] that][S_1...])$ ,  $[S_3...]) = \Im([S_1...])$ .

By invoking the two-place Val\*(,) function and allowing the semantic values of nodes to vary relative to identical or dominating nodes, Higgin-botham is able to provide, at least in principle, a semantic theory that is able to preserve something like *extensional compositionality*. But since Higginbotham rejects the one-place Val() function in favor of the two-place Val\*(,) function, he must be interpreted as rejecting *extensional compositionality* as a constraint on formal semantic theories, and endorsing instead the following principle of *relativized extensional compositionality*:

There is some function f such that for every node  $[\alpha[\beta...][\gamma]]$ , where  $[\alpha...]$  dominates immediate daughters  $[\beta...]$  and  $[\gamma...]$ , and every node  $[\delta...]$ , if according to the theory  $Val^*([\beta...], [\delta...]) = x$  and  $Val^*([\gamma...], [\gamma...], [\delta...]) = y$  and  $Val^*([\alpha...], [\delta...]) = z$ , then  $f(\langle x, y \rangle) = z$ .

How does Higginbotham propose to utilize the technique of relativizing semantic values to solve the problem posed by attitude ascriptions? Under the Standard Theory, the [CP[C] that  $[S_1...]$  node of  $[S_3...]$  is assigned the semantic value true. This is because the immediate daughter node  $[S_1...]$  is assigned the semantic value true, and the other immediate daughter node [C] that [C] is semantically vacuous. Under Higginbotham's proposed solution, however, the [CP[C] that  $[S_1...]$  node of  $[S_3...]$  is not assigned, relative to  $[S_3...]$ , the semantic value true, but is instead assigned, relative to  $[S_3...]$ , the semantic value  $[S_3...]$ . Under Frege's proposed solution semantic value shifting allows the nodes [CP[C] that  $[S_1...]$  and [CP[C] that  $[S_2...]$  to be assigned distinct  $[S_3...]$ , and this allows Frege to preserve some sort

botham defines  $Val^*(\ ,\ )$  so that  $Val^*([\alpha\ldots],[\beta\ldots])$  is defined only if  $[\beta\ldots]$  is identical to or dominated by  $[\alpha\ldots]$ . Thus the *primary* semantic value of a node can be identified with the semantic value of a node relative to itself, there is no "lower" node relative to which a node has a semantic value.

<sup>&</sup>lt;sup>27</sup> Higginbotham does not identify ILFs with the complex structures that Larson and Ludlow identify with ILFs, though it is not entirely clear what ILFs are on Higginbotham's account. Higginbotham states that he follows Harman's (1972) suggestion that "the objects of the attitudes, at least when indicated by clauses, are objects similar to the interpreted structures of the clauses themselves" (Higginbotham, 1986, p. 39). As I suggested above, however, it does not matter what ILFs *are*; all that matters from the point of view of semantics is how they are individuated.

of compositionality while also predicting that 'Odile believes that Twain wrote *Huck Finn*' and 'Odile believes that Clemens wrote *Huck Finn*' will have different truth conditions. Similarly, Higginbotham's relativization of semantic values to dominating nodes allows the nodes [CP[C] that][SI] and [CP[C] that][SI] to be assigned distinct ILFs as semantic values in keeping with the constraint of *relativized extensional compositionality*, and this in turn allows Higginbotham to assign distinct truth conditions to [SI] and [SI] in keeping with *relativized extensional compositionality*.

Higginbotham can thus be interpreted as proposing a semantic theory that incorporates the following relativized Combinatorial Rules:

```
\begin{split} & Val^*([{}_{CP}[{}_{C} \; that][{}_{S'}...]], \\ & [s[{}_{NP}...][{}_{VP}[[{}_{V} \; believes][{}_{CP}[{}_{C} \; that][{}_{S'}...]]]]) = \mathfrak{I}([{}_{S'}...]). \\ & Val^*([{}_{VP}[[{}_{V} \; believes][{}_{CP}[{}_{C} \; that][{}_{S'}...]]]], \\ & [s[{}_{NP}...][{}_{VP}[[{}_{V} \; believes][{}_{CP}[{}_{C} \; that][{}_{S'}...]]]]) = \{x: \; \exists y \; (y = \mathfrak{I}([{}_{S'}...]) \; \& \; \langle x, \, y \rangle \in \{\langle v, \, w \rangle \colon v \; believes \; w\})\}. \end{split}
```

Though Higginbotham also maintains that

```
Val*([s_1...], [s_1...]) = true iff Twain \in \{x: \exists z \ (z = it \ Huck \ Finn \& \langle x, z \rangle \in \{\langle v, w \rangle: v \ wrote \ w\})\}.
```

And thus a properly formulated theory incorporating Higginbotham's proposal would entail the following relativized theorems:

```
\begin{split} & \text{Val}^*([_{S3}\dots],\,[_{S3}\dots]) = \textit{true} \text{ iff} \\ & \text{Odile} \in \{x\colon \exists z\ (z=\mathfrak{I}([_{S1}\dots]) \text{ and } \langle x,\,z\rangle \in \{\langle w,\,v\rangle\colon w \text{ believes } v\})\}. \\ & \text{Val}^*([_{S4}\dots],\,[_{S4}\dots]) = \textit{true} \text{ iff} \\ & \text{Odile} \in \{x\colon \exists z\ (z=\mathfrak{I}([_{S2}\dots]) \text{ and } \langle x,\,z\rangle \in \{\langle w,\,v\rangle\colon w \text{ believes } v\})\}. \end{split}
```

So Higginbotham's proposal is similar to Larson and Ludlow's proposal in that  $[s_3...]$  and  $[s_4...]$  will be assigned distinct truth conditions so long as  $\Im([s_1...]) \neq \Im([s_2...])$ .

As it stands, however, Higginbotham's proposal is not adequate. Consider the following passage, in which Davidson criticizes Frege's analysis of attitude ascriptions:

the languages Frege suggests as models for natural languages ... are [not] amenable to theory in the sense of a truth definition meeting Tarski's standards. What stands in the way ... is that every referring expression has an infinite number of entities it may refer to, depending on the context, and there is no rule that gives the reference in more complex contexts on the basis of the reference in simpler ones. (Davidson, 'On Saying That', 1984, p. 99)

The problem with Frege's proposal, according to Davidson, is that it cannot be incorporated into a formal Tarski-style theory of truth because "there is no rule that gives the reference [of a term] in more complex contexts on the basis of the reference [of the term] in simpler ones." (The fact that Frege's theory abandons semantic innocence is an additional problem.) Higginbotham suggests that by invoking the two-place Val\*(,) function it is in principle possible to formulate a semantic theory that respects at least relativized extensional compositionality, but he does not formulate a finite semantic theory that states or entails exactly what semantic value(s) are to be assigned to a given occurrence of an expression. If Higginbotham's proposal is to preserve at least relativized extensional compositionality at the cost of *innocence*, yet also satisfy Davidson's demand that it be "amenable to theory in the sense of a truth definition meeting Tarski's standards", then Higginbotham's theory must assign new semantic values to the nodes inside that-clauses, and not just shift the semantic values of that-clauses alone. That is, Higginbotham must show us how the ILF assigned as the semantic value of a [CP[C that][S...]] node is a function of the semantic values assigned to nodes within this [CP[C] that [S...] node.

Though neither Higginbotham, nor his followers<sup>28</sup> has shown how this is to be done within the framework of a Tarski-style truth theory, it can be done.<sup>29</sup> The resulting formal theory, however, is rather baroque. What the formal theory does is to assign to each node  $[\alpha...]$ , relative to every node  $[\beta...]$  that either dominates or is identical to  $[\alpha...]$ , a semantic value. The theory does so in such a way that the semantic value of  $[\alpha...]$  relative to nodes that are neither identical to nor dominated by a [CP...] node is  $[\alpha...]$ 's "ordinary" semantic value. But the semantic value of  $[\alpha...]$  relative to a node that is either identical to or dominated by a [CP...] node is an ILF. This, in effect, assigns semantic values to all relevant node pairs, while allowing the semantic values of nodes within that-clauses to

<sup>&</sup>lt;sup>28</sup> Both Segal (1989) and Pietroski (1994) endorse Higginbotham's proposal, but neither demonstrates how it can be fleshed out so that it meets Davidson's challenge.

<sup>&</sup>lt;sup>29</sup> I was once skeptical that Higginbotham's proposal could be fleshed out to meet Davidson's challenge, but Josep Macia-Fabrega showed me that it could be done.

"shift" semantic values. After stating the theory, I will use it to derive truth conditions for  $[s_3...]$ .

## **Basic Lexical Rules:**

- (LR1)  $Val^*([NP Twain], [NP Twain]) = Twain$
- (LR2)  $Val^*([NP Clemens], [NP Clemens]) = Twain$
- (LR3)  $Val^*([NP Odile], [NP Odile]) = Odile$
- (LR4)  $Val^*([NP Huck Finn], [NPHuck Finn]) = Huck Finn$
- (LR5)  $Val^*([v \text{ wrote}], [v \text{ wrote}]) = \{\langle x, y \rangle : x \text{ wrote } y\}$
- (LR6)  $Val^*([v believes], [v believes]) = \{\langle x, y \rangle : x believes y\}$
- (LR7)  $Val^*([c that], [c that]) = \emptyset$

## **Recursive Lexical Rule:**

- (RLR) If  $[\gamma \dots]$  is a lexical (terminal) node, and  $[\alpha \dots]$  immediately dominates  $[\beta \dots]$ , and either  $[\gamma \dots] = [\beta \dots]$  or  $[\beta \dots]$  dominates  $[\gamma \dots]$ , then
  - (i) if  $[\alpha...]$  is not a [CP...], then  $Val^*([\gamma...], [\alpha...]) = Val^*([\gamma...], [\beta...])$
  - (ii) if  $[\alpha \dots]$  is a  $[CP \dots]$ , then  $Val^*([\gamma \dots], [\alpha \dots]) = \mathfrak{I}^*([\gamma \dots])$ .

## **Combinatorial Rules:**

- (CR1) If  $[\nu]$  is of the form  $[\nu]$   $[\nu]$ , then
  - (i) if  $Val^*([v...], [a...])$  is not an ILF, then  $Val^*([v...], [a...])$ =  $\{x: \langle x, Val^*([v...], [a...]) \rangle \in Val^*([v...], [a...])\}$
  - (ii) if  $Val^*([v...], [\alpha...])$  is an ILF, then  $Val^*([v...], [\alpha...]) = \mathfrak{I}^*([v...])$
- (CR2) If  $[\gamma \dots]$  is of the form  $[\nabla P[\nabla \dots]][\nabla P[\nabla \dots]]$ , then
  - (i) if  $Val^*([v...], [a...])$  is not an ILF, then  $Val^*([v...], [a...])$ =  $\{x: \langle x, Val^*([cp...], [a...]) \rangle \in Val^*([v...], [a...])\}$
  - (ii) if  $Val^*([v...], [a...])$  is an ILF, then  $Val^*([v...], [a...]) = \mathfrak{I}^*([v...])$

- (CR3) If  $[\nu]$  is of the form  $[S[\nu]]$ , then
  - (i) if  $Val^*([NP...], [\alpha...])$  is not an ILF, then  $Val^*([\gamma...], [\alpha...]) = true$  iff  $Val^*([NP...], [\alpha...]) \in Val^*([VP...], [\alpha...])$
  - (ii) if  $Val^*([NP...], [\alpha...])$  is an ILF, then  $Val^*([\gamma...], [\alpha...]) = \mathfrak{I}^*([\gamma...])$

(CR4) If 
$$[\gamma \dots]$$
 is of the form  $[CP[C \dots][S \dots]]$ , then  $Val^*([\gamma \dots], [\alpha \dots]) = Val^*([S \dots], [\alpha \dots])$ 

Let us now apply these rules to derive the truth conditions of

 $[_{S3}[_{NP} \ Odile][_{VP1}[_{V} \ believes]]_{CP}[_{C} \ that][_{S1}[_{NP} \ Twain]]_{VP2}[_{V} \ wrote][_{NP} \ Huck \ Finn]]]]]$ 

Let us first utilize the basic lexical rules and the recursive lexical rule to determine the semantic values of the terminal nodes of  $[s_3...]$ :

- 1.  $Val^*([NP Odile], [S3...]) = (by (RLRi)) Val^*([NP Odile], [NP Odile]) = (by (LR3)) Odile$
- 2.  $Val^*([v believes], [s_3...] = (by (RLRi)) Val^*([v believes], [vp_1...] = (by (LR6)) {\langle x, y \rangle: x believes y}$
- 3.  $Val^*([C that], [S_3...]) = (by (RLRi)) Val^*([C that], [VP1...]) = (by (RLRii)) \mathfrak{I}^*([C that])$
- 4.  $Val^*([NP Twain], [S3...]) = (by (RLRi)) Val^*([NP Twain], [VP1...]) = (by (RLRi)) Val^*([NP Twain], [CP...]) = (by (RLRii)) <math>\mathfrak{I}^*([NP Twain])$
- 5.  $Val^*([v \text{ wrote}], [s_3...]) = (by (RLRi)) Val^*([v \text{ wrote}], [v_{P1}...]) = (by (RLRi)) Val^*([v \text{ wrote}], [c_{P}...]) = (by (RLRii)) \mathfrak{I}^*([v \text{ wrote}])$
- 6.  $Val^*([NP \ Huck \ Finn], [S3...]) = (by (RLRi)) Val^*([NP \ Huck \ Finn], [VP1]) = (by (RLRi)) Val^*([NP \ Huck \ Finn], [CP...]) = (by (RLRii)) <math>\mathfrak{I}^*([NP \ Huck \ Finn])$

Let us now utilize the combinatorial rules, together with the determined semantic values of the terminal nodes above, to determine the semantic values of the root  $[s_3...]$  node:

- 7.  $Val^*([s_3...], [s_3...]) = (by (CR3))$ 
  - (7i) true iff  $Val^*([NP Odile], [S3...]) \in Val^*([VP1...], [S3...])$ , if  $Val^*([NP Odile], [S3...])$  is not an ILF.
  - (7ii)  $\mathfrak{I}^*([s_3...], [s_3...])$ , if  $Val^*([s_1 Odile], [s_3...])$  is an ILF.

But from step 1 above, we know that  $Val^*([_{NP}\ Odile],[_{S3}\dots])$  is Odile, and not an ILF.

So we have,

(7i') 
$$Val^*([s_3...], [s_3...]) = true \text{ iff Odile } \in Val^*([v_{P1}...], [s_3...])$$

8. 
$$Val^*([v_{P1}...], [s_3...]) = (by (CR2))$$

- (8i)  $\{x: \langle x, Val^*([CP...], [S3...]) \in Val^*([V...], [S3...]) \}, if Val^*([V] believes], [S3...]) is not an ILF$
- (8ii)  $\mathfrak{I}^*([v_{P1}...])$ , if  $Val^*([v_{P1}...])$  is an ILF

But from step 2. above we know that  $Val^*([v] believes], [s_3...])$  is  $\{\langle x, y \rangle : x believes y\}$  and *not* an ILF. So we have

(8i') 
$$\begin{aligned} & \text{Val*}([_{\text{VP1}}\dots],[_{\text{S3}}\dots]) = \{x: \langle x, \text{Val*}([_{\text{CP}}\dots],[_{\text{S3}}\dots]) \rangle \in \{\langle x,y \rangle: \\ & x \text{ believes } y \} \} \end{aligned}$$

Putting (8i') together with (7i') yields,

(7i") 
$$Val^*([s_3...], [s_3...]) = true \text{ iff Odile } \in \{x: \langle x, Val^*([c_P...], [s_3...])\rangle \in \{\langle x, y \rangle: x \text{ believes } y\}\}$$

9. 
$$Val^*([CP...], [S3...]) = (by (CR4)) Val^*([S1...], [S3...])$$

10. 
$$Val^*([s_1...], [s_3...]) = (by (CR3))$$

- (10i)  $true ext{ iff Val}^*([_{NP} ext{ Twain}], [_{S3}...]) \in Val^*([_{VP2}...], [_{S3}...]), if Val^*([_{NP} ext{ Twain}], [_{S3}...]) is not an ILF$
- (10ii)  $\mathfrak{I}^*([s_1...])$ , if  $Val^*([s_N Twain], [s_3...])$  is an ILF

But from step 4 above we know that  $Val^*([_{NP}\ Twain], [_{S3}\dots])$  is  $\mathfrak{I}^*([_{NP}\ Twain])$  which is an ILF. So we have

(10ii') Val\*([s<sub>1</sub>...], [s<sub>3</sub>...]) = 
$$\mathfrak{I}$$
\*([s<sub>1</sub>...])

From (10ii') and (7i") it follows that

```
(7i''') Val*([s<sub>3</sub>...], [s<sub>3</sub>...]) = true iff Odile \in \{x: \langle x, \mathfrak{I}^*([s_1...]) \rangle \in \{\langle x, y \rangle: x \text{ believes } y\}\}
```

A similar derivation would yield that

Val\*([S4...], [S4...]) = true iff Odile 
$$\in \{x: \langle x, \mathfrak{I}^*([S2...]) \rangle \in \{\langle x, y \rangle: x \text{ believes } y\}\}$$

Since  $\mathfrak{I}^*([s_2...])$  is distinct from  $\mathfrak{I}^*([s_1...])$ ,  $[s_3...]$  and  $[s_4...]$  are assigned distinct truth conditions.

This formalization of Higginbotham's proposal not only respects relativized extensional compositionality, but it also meets Davidson's challenge to Frege's proposal. The theory determines when semantic value shifting is to take place and it does so with a finite number of rules. Moreover, the theory meets Davidson's objection against Frege's proposal: it determines  $Val^*([\gamma \dots], [\alpha \dots])$  for every node  $[\gamma \dots]$  and for every node  $[\alpha...]$  that either dominates or is identical to  $[\gamma...]$ .<sup>30</sup> But does the theory illustrate a plausible explanation of our semantic competence? What is wanted is not merely a formal Tarski-style truth theory that somehow manages to preserve some sort of compositionality constraint. What is wanted is a theory that illustrates a plausible explanation of the semantic competence of ordinary speakers. Higginbotham's theory does not illustrate Davidson's strategy for explaining our semantic competence because it violates semantic innocence, and it rejects extensional compositionality in favor of relativized extensional compositionality. It does not follow from this alone that no plausible explanatory strategy is illustrated by Higginbotham's formal theory, as the theory may illustrate an alternative strategy. There are, however, two reasons suggesting that the formal apparatus of Higginbotham's ILF proposal is not motivated by any such alternative strategy.

First, the only explanatory strategy that could motivate the above formalization of Higginbotham's proposal is something like this: A competent speaker is able to determine the truth conditions of an arbitrary sentence because he knows the basic semantic values of the words in the sentence – he knows what the semantic value of each word relative to itself is. From this knowledge, together with his knowledge of the LF of the sentence, he is able to determine for each word its non-basic semantic value relative

 $<sup>^{30}</sup>$  The theory also yields correct truth conditions for attitude ascriptions with multiple embeddings, though I will not demonstrate this here.

to each node that dominates the word. And finally, from his knowledge of all these non-basic semantic values, together with his knowledge of the LF of the sentence, he is able to determine the conditions under which the sentence, relative to itself, is true. It is an empirical question as to whether or not this explanatory strategy reflects how competent speakers actually come to understand utterances of natural language sentences, and I concede that it is possible that our semantic competence is due to some process roughly like this. But I submit that this explanatory strategy is extremely implausible.

It is obvious that if a speaker is to comprehend an utterance of  $[S_1]_{NP}$ Twain][ $_{VP}[_{V} \text{ wrote}][_{NP} \text{ Huck Finn}]]]$  he must know the semantic value of [NP Huck Finn]. But it seems just as obvious that he need not know, in addition to this, the semantic value of [NP Huck Finn] relative to [VP[V wrote][ $_{NP}$  Huck Finn]], and also relative to [ $_{S1}[_{NP}$  Twain][ $_{VP}[_{V}$  wrote][ $_{NP}$ *Huck Finn*]]]. One might argue that evidence in support of this seemingly implausible explanatory strategy is garnered by the fact that this strategy would allow us to state a formal semantic theory for attitude ascriptions within the framework of a Tarski-style truth theory. But this fact would constitute evidence in support of the strategy only if there were independent reasons for supposing that an adequate semantic theory must be formalized as a Tarski-style truth theory. (It would be circular to argue, first, that Higginbotham's baroque formal theory is explanatory because it illustrates a seemingly implausible explanatory strategy, and then to argue, second, that the explanatory strategy is not implausible after all, because it is illustrated by the baroque formal theory.) But, given the demise of Davidson's semantic program, there are no independent reasons for supposing that an adequate semantic theory must be formalized as a Tarski-style truth theory.

Second, Higginbotham's ILF theory egregiously violates *semantic innocence*. The principle of *semantic innocence* requires, roughly, that occurrences of a node be assigned the same semantic value (or set of semantic values) regardless of the larger syntactic environments of the occurrences. Thus Higginbotham's proposal of relativizing semantic value, which allows that an occurrence of a node may have a distinct semantic value relative to itself and every node that dominates it, directly contradicts *semantic innocence*. This is problematic because the principle of *semantic innocence* is supported by two important theoretical considerations.<sup>31</sup>

<sup>31</sup> These two theoretical considerations are reflected in Larson and Segal's (1995) definition of innocence: "We prefer the semantics of clausal-complement constructions to be innocent: the values assigned in embedded and unembedded contexts should be the same;

The first consideration concerns the plausibility of a semantic theory being known and especially *learned* by ordinary people. As Larson and Segal state, "On an innocent theory, a child acquiring language has only one semantic-valuation statement to learn for any given lexical item . . . . By contrast, on a noninnocent semantics, a learner is obliged to acquire a number of valuation statements for a given item, mastering the context where each applies" (1995, p. 436). Because Higginbotham's ILF theory egregiously violates *semantic innocence*, it is relatively implausible that it is the theory actually learned by young children, and thus it is relatively implausible that this theory explains the semantic competence of ordinary people.

The second consideration in support of the principle of semantic innocence concerns theoretical parsimony. Larson and Segal explain that "the semantic analysis of the simpler parts of the grammar [ought to be motivated] from data available in those simpler parts," and "the analysis of the simpler parts [ought not to be motivated] from data concerning the more complex parts" (1995, p. 437). In other words, complexities introduced into a semantic theory to account for attitude ascriptions ought not affect how the theory treats simple declarative sentences.<sup>32</sup> This follows from general considerations of parsimony: if there are two semantic theories, both of which correctly predict the truth conditions of simple declarative sentences, then the simpler of the two theories is to be preferred for this class of sentences. But semantic theories of the sort preferred by Davidson, which are designed to account for the truth conditions of simple declarative sentences, correctly predict the truth conditions of these sentences, and these theories are much simpler than Higginbotham's ILF theory, which is designed to account for attitude ascriptions. The above formalization makes it clear that Higginbotham's relativization of semantic value adds a great deal of complexity to the entire semantic theory. Moreover, the complex relativizing machinery in Higginbotham's theory is present, but idle, when the theory is applied to simple declarative sentences. Consequently, though Higginbotham's ILF theory may be preferable to a simple extensional theory when applied to attitude ascriptions, simple extensional semantic theories are preferable to Higginbotham's ILF theory when ap-

furthermore, the analysis of the unembedded parts should not depend on evidence from the embedded parts" (Larson and Segal, 1995, p. 437).

<sup>&</sup>lt;sup>32</sup> Higginbotham also discusses an interesting problem concerning the intensional operator 'unless' and he proposes a relativized treatment of the operator. I think Higginbotham's relativized treatment is problematic, and that a much more plausible analysis of 'unless' is suggested by Kratzer (1979), (1986) and Lewis (1975). Unfortunately I do not have the space to discuss the issue here.

plied to simple declarative sentences. It is not clear which theory provides a superior *general* explanation of our semantic competence.

# 6. ARE TARSKI-STYLE TRUTH THEORIES EXPLANATORILY ADEQUATE?

The above remarks constitute evidence, by no means conclusive, that Tarski-style truth theories are inadequate as means of illustrating the knowledge and abilities that are responsible for our semantic competence. The motivation for formal Tarski-style truth theories was that such theories could illustrate Davidson's explanatory strategy for explaining our semantic competence. Though the ILF analyses considered above are finitely axiomatized truth theories that entail correct statements of truth conditions for attitude ascriptions, they deviate from standard Tarski-style truth theories in important ways, and as a result they do not illustrate Davidson's elegant explanatory strategy. Moreover, they do not illustrate plausible alternative explanatory strategies. This suggests, but does not entail, that our semantic competence concerning attitude ascriptions, and thus our semantic competence generally, cannot be adequately accounted for in terms of Tarski-style truth theories. Standard Tarski-style truth theories of the sort advocated by Davidson are well motivated theoretically, but do not entail theorems stating the correct truth conditions of attitude ascriptions. Deviant Tarski-style truth theories, of the sorts advocated by Larson and Ludlow and Higginbotham, can entail theorems stating the correct truth conditions of attitude ascriptions, but do not illustrate plausible explanatory strategies.

The best way to demonstrate the inadequacy of Tarski-style truth theories would be to present a superior alternative, but I cannot attempt to do that here. I will suggest, however, that "Discourse Representation Theory" (DRT), as proposed by Hans Kamp (1990, 1993), Nicolas Asher (1986), and others, may prove to be a superior alternative. DRT provides a formal framework in which the fundamental idea behind the ILF proposals can be utilized in a more plausible way. The fundamental idea behind the ILF proposals is that the truth conditions of a sentence, in particular an attitude ascription, are not solely a function of the syntactic structure of, and semantic values invoked by, the sentence, but are also a function of the very words in the sentence. The only way that this insight can be incorporated into a Tarski-style truth theory is if the words themselves are somehow added into the semantic values invoked by the sentence, and this can be done only by violating extensional compositionality and/or semantic innocence. But, as I have argued, if these principles are violated, then the theoretical

motivation for formulating semantic theories as Tarski-style truth theories in the first place is undermined. DRT, however, presents a theoretically motivated way in which the actual words in a sentence can affect the truth conditions of a sentence even though these words are not in any way added into the semantic values invoked by the sentence. For in DRT, which applies to entire discourses and not isolated sentences, intersentential semantic relations can affect the truth conditions of occurrences of sentences. These intersentential relations are partly determined by the very words in the sentence. As a result, in DRT the truth conditions of a sentence are in part a function of the words in the sentence, though in order to affect the truth conditions the words need *not* be incorporated into the semantic values of the expressions in the sentence. Moreover, these intersentential relations include not only pronominal anaphora, but also intersentential relations between names, other NPs and perhaps even predicates. For this reason DRT, or a similar theoretic approach, may offer a more plausible semantics for attitude ascriptions, and a superior explanation of our semantic competence in general.

#### REFERENCES

Asher, N.: 1986, 'Belief in Discourse Representation Theory', *Journal of Philosophical Logic* **15**, 137–189.

Burge, T.: 1986, 'On Davidson's "Saying That" ', in E. LePore, (ed.), Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson, Blackwell, Oxford, pp. 190-2

Clapp, L.: 1995, 'How to be Naive and Innocent: A Criticism of Crimmins and Perry's Theory of Attitude Ascriptions', *Linguistics and Philosophy* 18, 529–565.

Crimmins, M.: 1992, Talk About Beliefs, MIT Press, Cambridge.

Davidson, D.: 1967, 'Truth and Meaning', *Synthese* **17**, 304–323. Also printed in Davidson (1984).

Davidson, D.: 1968, 'On Saying That', *Synthese* 19, 130–146. Also printed in Davidson (1984).

Davidson, D.: 1974, 'Belief and the Basis of Meaning', *Synthese* **27**, 309–323. Also printed in Davidson (1984).

Davidson, D.: 1984, *Inquiries into Truth and Interpretation*, Oxford University Press, Oxford

Frege, G.: 1892, 'Sinn und Bedeutung', Zeitschrift für Philosophie und philosphische Kritik 100, 25–50 (published as 'Sense and Reference' in P. Geach and M. Black, (eds.), 1952, Translations from the Philosophical Writings of Gottlob Frege, Basil Blackwell, Oxford, pp. 56–78).

Harman, G.: 1972, 'Deep Structure as Logical Form', in D. Davidson and G. Harman (eds.), *Semantics of Natural Language*, pp. 25–47, D. Reidel, Dordrecht, Holland.

Higginbotham, J.: 1985, 'On Semantics', Linguistic Inquiry 16, 547–593.

Higginbotham, J.: 1986, 'Linguistic Theory and Davidson's Program in Semantics', inE. LePore, (ed.), *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, pp. 29–48, Blackwell, Oxford.

Higginbotham, J.: 1991, 'Belief and Logical Form', Mind and Language 6, 344–369.

Kamp, H.: 1990, 'Prolegomena to a Structural Account of Belief and Other Attitudes', in C. A. Anderson and J. Owens (eds.), *Propositional Attitudes*, pp. 27–90, CSLI, Stanford.

Kamp, H. and Reyle, U.: 1993, From Discourse to Logic: Introduction to Model Theoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory, Kluwer, Dordrecht.

Kratzer, A.: 1979, 'Conditional Necessity and Possibility', in R. Baeuerle, U. Egli, and A. V. Stechow (eds.), *Semantics from Different Points of View*, Springer, Berlin.

Kratzer, A.: 1986, 'Conditionals', in A. M. Farley, P. Farley, and K. E. McCullough (eds.), Papers from the Parasession on Pragmatics and Grammatical Theory, Chicago Linguistic Society.

Larson, R. and Ludlow, P.: 1993, 'Interpreted Logical Forms,' Synthese 95, 305–355.

Larson, R. and Segal, G.: 1995, *Knowledge of Meaning: An Introduction to Semantic Theory*, MIT Press, Cambridge.

LePore, E. and Loewer, B.: 1989, 'You Can Say That Again,' in P. French, T. Uehling, and H. Wettstein (eds.), *Midwest Studies in Philosophy Volume XIV: Contemporary Perspectives in the Philosophy of Language II*, University of Notre Dame Press, Notre Dame.

Lewis, D.: 1975, 'Adverbs of Quantification', in E. Keenan (ed.), *Formal Semantics of Natural Language*, Cambridge University Press, Cambridge.

Montague, R.: 1974, Formal Philosophy, Yale University Press, New Haven.

Pietroski, P.: 1994, 'Defending Similarity in Semantics', presented at the Canadian Philosophical Association Meeting, Calgary, Alberta, June 8th 1994.

Richard, M.: 1990, *Propositional Attitudes: An Essay on Thoughts and How we Ascribe Them*, Cambridge University Press, Cambridge.

Rumfitt, I.: 1993, 'Content and Context: The Paratactic Theory Revisited and Revised', Mind 102, 429-454.

Segal, G.: 1989, 'A Preference for Sense and Reference', *Journal of Philosophy* **86**, 73–89. Segal, G. and Speas, M.: 1986, 'On Saying [thet]', *Mind and Language* (Summer), pp. 124-132.

Stainton, R.: 1994, 'How to Higgy a Frege-Church (Comments on Pietroski)', presented at the Canadian Philosophical Association meeting, Calgary, Alberta, June 8th 1994.

Lenny Clapp
Department of Philosophy
Illinois Wesleyan University
Bloomington Illinois
E-mail: lclapp@titan.iwu.edu

Copyright of Linguistics & Philosophy is the property of Kluwer Academic Publishing / Academic and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.