

CHAPTER 4

**Beyond Sense and Reference:
An Alternative Response
to the Problem of Opacity ***

Lenny Clapp

Department of Philosophy, Illinois Wesleyan University, Bloomington, IL 61702-2900, USA

Contents

1. Introduction	44
2. The general difficulty with Frege's strategy	46
3. The general difficulty and Richard's overt indexical analysis	55
4. An alternative strategy for responding to the problem of opacity	65
References	74

*I am grateful to Mark Richard, Sylvain Bromberger, Alex Byrne, Irena Heim, Cara Spencer, Michael Glansberg, Rob Streifer, Robert Stainton, Andrew Botterell, Jason Stanley, Jessica Wilson, Harry Deutsch, Mary Reed and the philosophy graduate students at MIT for helpful comments and criticisms.

PRAGMATICS OF PROPOSITIONAL ATTITUDE REPORTS
Current Research in the Semantics/Pragmatics Interface, Vol. 4
Edited by K.M. Jaszczolt
© 2000 Elsevier Science Ltd. All rights reserved

1. Introduction

Semantic theorists working within the traditions of Davidson and Montague have followed Frege (1893) in presupposing what I will call the principle of *sentential compositionality*: the truth conditions of an occurrence of a declarative sentence must be determined by (i) the syntactic structure of the sentence,¹ and (ii) the semantic values assigned to the words and other semantically relevant syntactic elements in the sentence. In this theoretical tradition the task of a semantic theory is to illustrate how the truth conditions of a sentence are a *function* of its syntactic structure and the semantic values it invokes. More precisely, each semantically relevant syntactic element of a sentence, including perhaps *phonetically unrealized* elements, is mapped, relative to a context of utterance, to a semantic value. The theory is directed by this mapping, and the syntactic structure of the sentence, to determine how all the semantic values invoked by the occurrence are combined to determine its truth conditions.² Though it is widely recognized that due to context sensitive expressions such as indexicals, demonstratives and tensed verbs, context plays a significant role in determining the truth conditions of an utterance, the role played by context has been limited to that of determining the semantic values of context sensitive expressions. Thus, though it is widely recognized that most if not all declarative sentences determine truth conditions only relative to a context, it is nonetheless presupposed that context is relevant *only* for determining what semantic values are invoked by an occurrence; *sentential compositionality* is preserved.

It is within the constraints imposed by sentential compositionality that *the problem of opacity* for propositional attitude ascriptions arises. Consider an all too familiar example:

- (1) John believes that Twain wrote *Huckleberry Finn*.
- (2) John believes that Clemens wrote *Huckleberry Finn*.

Sentences (1) and (2) seem to have the same syntactic structure. Moreover, since 'Twain' and 'Clemens' refer to the same man, it at least seems that (1) and (2) invoke the same semantic values in the same order. But if occurrences of (1) and (2) have the same syntactic structure, and invoke the same semantic values in the same order, then *sentential compositionality* dictates that they have the same truth conditions. This is problematic because competent speakers judge that some occurrences of (1) and (2) have distinct truth conditions.

¹ In contemporary syntactic theory the relevant level of syntactic representation is called "LF". There is not universal agreement, however, concerning precisely what information appears at LF. Such controversies are not directly relevant to my purposes here, and so I merely assume that those committed to sentential compositionality presuppose the existence of *some* syntactic structure that interacts with a semantic theory as described above.

² Commitment to, and motivation for, sentential compositionality is clear in recent textbooks on semantics. See Heim and Kratzer (1998), pp. 1–2, and Larson and Segal (1995), pp. 11–16.

If sentential compositionality is taken for granted, and it is assumed that (1) and (2) have the same syntactic structure, then there are only two plausible strategies for responding to this problem.³ First, one can maintain that though (1) and (2) have the same syntactic structure, despite initial appearances they do not invoke the same semantic values. This is the strategy adopted by Frege. Frege claimed that expressions occurring with the complement clauses of attitude ascriptions do not have their ordinary referents as semantic values, but instead they shift semantic values so that they have extraordinary entities – senses – as their semantic values. So according to Frege embedded occurrences of 'Twain' and 'Clemens' refer to distinct senses. In this way Frege solves the problem of opacity, for on Frege's analysis occurrences of (1) and (2) invoke different semantic values, and thus it is not problematic that they have distinct truth conditions. I shall say that any proposed solution to the problem of opacity that follows Frege in claiming that, despite appearances, ascriptions such as (1) and (2) do not invoke the same semantic values utilizes *Frege's strategy*. The second strategy of response is to deny that the problem ever really arises. According to this strategy, ascription pairs such as (1) and (2) cannot differ in truth value, and our judgments to the contrary are due to a conflation of *semantic* factors which are limited to considerations of reference and truth value, and mere *pragmatic* factors such as the *appropriateness* of the utterance, and whether or not it is *misleading*. As this second strategy is developed in detail by Salmon (1986) I will call it *Salmon's strategy*.⁴

In my view neither of these strategies of response is adequate, and the only adequate response is to reject the general theoretical constraint that gives rise to the problem: sentential compositionality must be rejected. Salmon's strategy is inadequate not so much because it is implausible, but because it is incompatible with the central task of a semantic theory. The central task of a semantic theory is to explain and predict the judgments of competent speakers concerning the truth conditions of sentences. Denying the veracity of such judgments is not a viable option. Consequently I will here be concerned primarily to demonstrate that Frege's strategy cannot succeed. In what follows I first explain in general terms why *any* proposed solution to the problem of opacity that utilizes Frege's strategy will be inadequate. I then demonstrate in detail how this general problem arises for Richard's (1990)

³ I will assume throughout this paper that at the relevant level of syntactic representation (1) and (2) have the same syntactic structure. This assumption is incompatible with the claim that the difference in truth conditions is a result of a *structural ambiguity*. If one grants that (1) and (2) have same syntactic structure, one cannot explain the difference between (1) and (2) as being some sort of *de re/de dicto* distinction resulting from differences in quantifier scope.

⁴ It might be maintained that "paratactic" analyses, as developed for *oratio obliqua* by Davidson (1968) and extended to attitude ascriptions by LePore and Lowerer (1989), constitute a third plausible strategy. Though I will not consider such paratactic analyses here, they clearly fall within the bounds of Frege's strategy. This because the essential feature of paratactic analyses is their assignment of appropriate semantic values – *utterances* and/or mental event tokens – to the complementizers of complement clauses. Similar remarks apply to the "Interpreted Logical Form" analyses of attitude ascriptions proposed by Larson and Ludlow (1993), and others.

overt indexical analysis of attitude ascriptions. After thus completing my objection to Frege's strategy, I sketch a *discourse holistic* account of the truth conditions of attitude ascriptions that does not presuppose sentential compositionality.

2. The general difficulty with Frege's strategy

Though none of the more recently proposed solutions to the problem of opacity simply posits *senses* and claims that they are the semantic values of occurrences of embedded terms, many recently proposed solutions follow Frege in positing some sort of extraordinary entity to distinguish the truth conditions of ascriptions such as (1) and (2) in keeping with sentential compositionality. That is, though these proposals differ from Frege's in various ways, they all utilize Frege's strategy for solving the problem of opacity. I will briefly explicate two such proposals.

Mark Richard's (1990) "overt indexical" analysis invokes mental representations, expressions of mentalese, to play a role similar to the role played by Frege's senses. But Richard's ingenious analysis differs from Frege's in that Richard analyzes attitude verbs as indexicals; on Richard's analysis attitude verbs designate different relations in different contexts. This added overt indexical feature of Richard's analysis is designed to account for the context sensitivity of attitude ascriptions – it is designed to account for the fact that the same attitude ascription can have different truth conditions in different contexts of utterance. The details of Richard's analysis are somewhat complex, and I will present a more detailed explication of his analysis below, but the basic idea is this: An occurrence of (1) is true if and only if John holds the relation designated by this particular occurrence of 'believes' toward the sentence (type) 'Twain wrote Huck Finn' that occurs embedded in the complement clause. And John holds this particular relation toward 'Twain wrote Huck Finn' if and only if John has a sentence of mentalese Σ on his "belief blackboard" such that *in the context of utterance* 'Twain wrote Huck Finn' *appropriately translates* Σ . It is the semantic value of 'believes' in a context that determines what counts as *appropriate translation* in that context. Different occurrences of 'believes' determine more or less stringent *constraints* on what qualifies as an appropriate translation. If an occurrence of 'believes' determines very lax constraints, then it designates a more transparent *believes* relation, and if it determines very stringent constraints, then it designates a more opaque *believes* relation. So on Richard's analysis an attitude ascription uttered in a context c is true if and only if the sentence embedded in the complement clause *appropriately translates* in c a sentence of mentalese instantiated by the subject, and what qualifies as an *appropriate translation* in c is determined by what relation is designated by the attitude verb in c . How does Richard's analysis purport to solve the problem of opacity? Suppose that in a context c (1) is true, and (2) is false. Richard's analysis explains this possibility, in keeping with sentential compositionality, as follows: in c 'believes' designates a relation such that 'Twain wrote Huck Finn'

appropriately translates some sentence of mentalese on John's belief blackboard, whereas 'Clemens wrote *Huck Finn*' does not appropriately translate some sentence of mentalese on John's belief blackboard. So John does hold the relation designated by 'believes' in *c* toward 'Twain wrote *Huck Finn*', though he does not hold this relation toward 'Clemens wrote *Huck Finn*'. So in *c* (1) is true and (2) is false.

Crimmins and Perry's (1989, 1992) "hidden indexical" analysis posits *cognitive particulars* to play a role very similar to the role played by Frege's senses. But Crimmins and Perry's analysis differs from Frege's and Richard's in that their analysis is *semantically innocent*. According to Frege's and Richard's analyses, terms embedded inside complement clauses of attitude ascriptions *shift* semantic values. On these analyses embedded terms do not have their ordinary referents as semantic values, but instead have senses, expressions of mentalese, or the very expressions themselves, as their semantic values. Crimmins and Perry's analysis, however, does not rely on embedded terms *shifting* semantic values in this way. Rather Crimmins and Perry reject "the doctrine of full articulation," which maintains that every semantic value invoked by an occurrence must be "the content of some [phonetically realized] expression in the sentence" (1992, p. 10). Rejection of full articulation allows Crimmins and Perry to maintain that, while embedded articulated terms retain their ordinary referents, *cognitive particulars* are referred to by a "hidden indexical" element of an ascription. As I will interpret Crimmins and Perry's analysis, it really does posit a "hidden indexical". That is, I will interpret their analysis as positing in the syntactic structure of attitude ascriptions (perhaps at the level of LF) *phonetically unrealized* indexical elements that have cognitive particulars (or properties thereof) as semantic values.⁵ Occurrences of propositional attitude ascriptions are thus analyzed as asserting that a three-place relation obtains between an agent, a Russellian proposition, and a complex cognitive particular (or *kind* of cognitive particular), where the complex cognitive particular (or kind) is "tacitly referred to" by phonetically unrealized syntactic elements. And in this way Crimmins and Perry's analysis also purports to distinguish the truth conditions of occurrences of (1) and (2) in keeping with sentential compositionality: An occurrence of (1) is true if and only if the three-place *believes* relation is satisfied by (John, *P*, Σ_1), where *P* is the Russellian proposition that Twain wrote *Huckleberry Finn*, and Σ_1 is a complex cognitive particular in John's

⁵ Crimmins (1992) seems to allow, ironically enough, that there be no "hidden indexical" elements. That is, Crimmins seems to allow that an utterance of an attitude ascription might invoke cognitive particulars as semantic values, even though these cognitive particulars are not the semantic values of any phonetically realized or even phonetically unrealized syntactic elements. Thus Crimmins allows (something like) that an occurrence as a whole somehow "tacitly refers" to cognitive particulars. But this allowance is merely an unmotivated attempt to preserve the spirit of sentential compositionality, while avoiding the (implausible) commitment to phonetically unrealized syntactic elements that could have cognitive particulars as semantic values. There is no reason to think that what Crimmins calls "tacit reference" is a sort of *reference*. I therefore ignore Crimmins' allowance, and interpret the "hidden indexical" analysis as described above.

brain that includes one of John's Twain-ish cognitive particulars. An occurrence of (2), in contrast, is true if and only if the three-place *believes* relation is satisfied by $(\text{John}, P, \Sigma_2)$, where Σ_2 is a complex cognitive particular in John's brain that includes one of John's Clemens-ish cognitive particulars.

Both of these analyses utilize Frege's strategy; i.e. they each attempt to distinguish the truth conditions of occurrences of (1) and (2) in keeping with sentential compositionality by positing an extraordinary sort of entity and appealing to such entities to distinguish the semantic values invoked by the occurrences. But because these analyses utilize Frege's strategy, they fall prey to the same general difficulty. The difficulty, in essence, is that any analysis that utilizes Frege's strategy will be pulled in two incompatible directions. The pull in one direction arises from the requirement that the posited extraordinary entities be individuated finely enough to solve all instances of the problem of opacity. The pull in this direction forces the posited entities to be identified with very finely individuated, esoteric, mental entities. The pull in the other direction arises from the requirement that an analysis of attitude ascriptions preserve the legitimacy of our ordinary attitude ascribing practices. Ordinary people who lack detailed knowledge, or even beliefs, concerning one another's mental states are able to utter true attitude ascriptions about each other, and thereby give true explanations of one another's behavior. But this ability of ordinary speakers requires that the posited extraordinary entities be more coarsely individuated and publicly accessible. As nothing could satisfy these competing requirements, nothing could play the role of the posited extraordinary entities. I will first explicate this general difficulty in terms of something like Frege's theory of sense and reference, though the difficulty clearly arises for any analysis of attitude ascriptions that utilizes Frege's strategy.

Frege claims that occurrences of ascriptions such as (1) and (2) differ in truth conditions because the embedded occurrences of 'Twain' and 'Clemens' refer to distinct senses. But, granting the existence of senses, if the embedded occurrences of 'Twain' and 'Clemens' refer to the *same* sense, then Frege's theory fails to solve the problem of opacity. That is, if the occurrences of (1) and (2) have *distinct* truth conditions, yet the embedded occurrences of 'Twain' and 'Clemens' refer to the *same* sense, then the truth conditions of the occurrences of (1) and (2) are *not* a function of (i) their syntactic structure and (ii) the invoked semantic values, and thus Frege's theory of sense and reference would fail to preserve sentential compositionality. So merely positing senses to serve as "secondary referents" is not in itself sufficient to solve the problem of opacity. In order to solve the problem of opacity the posited senses must be individuated finely enough so that for any possible pair of ascription occurrences such as (1) and (2) that differ in truth conditions, there are expressions in the complement clauses of these ascriptions that refer to distinct senses. Let us make this slightly more precise. Consider all actual and possible attitude ascription occurrences that are of the general form, $\ulcorner nVs \text{ that } \Phi(\alpha) \urcorner$. A pair of such occurrences constitutes an *instance of opacity* if and only if

- (i) the occurrences have the same syntactic structure;
- (ii) the occurrences invoke the same *ordinary* semantic values, in the same order; and
- (iii) the occurrences are judged by competent speakers to have distinct truth conditions.

In order for senses to be *appropriately individuated*, the following constraint must be satisfied:

The Individuation Constraint: If two occurrences of attitude ascriptions O_1 and O_2 constitute an instance of opacity, then there must be some sense referred to by some expression in the complement clause of O_1 that is not referred to by any expression in the complement clause of O_2 .⁶

If (something like) Frege's theory of sense and reference is to solve the problem of opacity, then it must individuate senses finely enough so that the individuation constraint is satisfied.

Satisfying the individuation constraint is hardly a trivial matter, and I am skeptical that an *independent* account of senses (sentences of mentalese, cognitive particulars, whatever) that satisfies this constraint can be formulated.⁷ My central objection against theories that utilize Frege's strategy, however, does not presume that such a theory cannot provide an independent account of senses that satisfies this constraint. My central objection depends only upon the much weaker thesis, for which I argue below, that such a theory can satisfy the individuation constraint only if it identifies its posited entities, senses or whatever, with "ways of thinking," where these "ways of thinking" are individuated by appeal to particular details concerning the subject's mental life. If this is correct, then Frege's theory cannot maintain that 'Twain' has only one secondary referent; i.e. Frege's theory cannot maintain that all embedded occurrences of 'Twain' refer to the same *sense*, where a sense is something like a *linguistic meaning* known by all competent speakers. Rather it must allow distinct embedded occurrences of 'Twain' to refer to distinct "ways of thinking," where details concerning the subject's mental life determine which "way of thinking" ought to be referred to by a given ascription. But this result is incompatible with the pull in the other direction: ordinary speakers are

⁶ This statement of the *individuation constraint* is not quite right, as occurrences of 'John believes that Twain likes Clemens' and 'John believes that Clemens likes Twain' constitute instances of opacity, yet there need not be a sense referred to by some expression in one complement clause that is not referred to by an expression in the other. It is relatively easy to amend the individuation constraint so that it accounts for such special instances of opacity, though for the sake of simplicity I will not do so.

⁷ It would of course be trivial to provide a *dependent* account of senses that satisfies the individuation constraint, i.e. an account that individuates senses by appeal to opacity. For example, one could build into one's account of senses that if there is an instance of opacity involving occurrences of two embedded terms α and β , then these occurrences of α and β must, by fiat, express distinct senses. Such a *dependent* account would no doubt satisfy the individuation constraint, but it would also render the Fregean theorist's explanation of opacity circular, and therefore vacuous.

able to use and understand attitude ascriptions despite their ignorance of, and lack of intentions to refer to, such esoteric entities.

As examples such as Kripke's (1979) story involving Paderewski illustrate, there are *extreme* instances of opacity in which distinct occurrences of the same attitude ascription (type) have different truth conditions; i.e. there are many cases in which competent speakers would judge that one occurrence of an attitude ascription is true, while a different occurrence of the very same sentence, seemingly referring to the very same entities in the same order, is false. The phenomenon of extreme opacity does not require a setting as complex as Kripke's story involving Paderewski. Such cases are easy to construct using ascriptions containing indexicals and demonstratives. For example, competent speakers might judge one occurrence of 'John believes that *he* is an author' to be true, and another occurrence of this same sentence to be false, where the two occurrences of 'he' refer to the same person. Moreover, as is pointed out by Bach (1997, and this volume), "any 'that'-clause could be used, given the right circumstances, to describe something that someone believes and to describe something he disbelieves, and do so without imputing any incoherence to him" (1997, p. 233). So for example, even if John's mental state is held fixed, in the right circumstances one occurrence of 'John believes that Twain wrote *Huckleberry Finn*' can be true, and another occurrence of this very same ascription can be false, and yet no *irrationality* is thereby imputed to John. Consideration of the phenomenon of extreme opacity makes it clear that if (something like) Frege's theory is to satisfy the individuation constraint, it must individuate senses more finely than the *meanings* of expressions of natural language, and even more finely than such expressions themselves. So an advocate of Frege's analysis cannot maintain that senses are something like meanings or expressions of natural language. Rather he must identify the posited extraordinary entities with "ways of thinking" of some kind, and he must deny that every word (type) of natural language expresses only one such "way of thinking."⁸

But how are such "ways of thinking" to be individuated? At the heart of the problem of opacity lies the phenomenon of recognition failure. Where there is an instance of recognition failure, instances of opacity are not far to seek. The phenomenon of recognition failure occurs when a rational subject is twice cognizant of the same entity, and yet is not aware that both cognitions are *about* the same entity. For example, one might read about Chelsea Clinton in the newspaper, and then see her on Martha's Vineyard, and not be aware at the time of the seeing that the person one is seeing is the person one read about. At the core of Frege's strategy there lies this familiar model of recognition failure: If on a given occasion a subject fails to recognize an entity, then that subject must on that occasion instantiate

⁸ Even Frege (1893) allowed names to express different senses for different people. Frege famously suggested that for some the sense of 'Aristotle' is given by the description 'the pupil of Plato and the teacher of Alexander the Great', while for others the sense is given by the description 'the teacher of Alexander the Great who was born in Stagira'. Frege also suggested that such variance in sense was a defect of natural language.

two “ways of thinking” (senses, modes of apprehension, whatever) of the entity, and fail to “connect” these “ways of thinking.” The upshot is that “ways of thinking” must be individuated finely enough to account for every possible instance of recognition failure. For given any two cognitive acts in which a person thinks about the same entity, later circumstances *could* conspire to provide him with overwhelming, though misleading, evidence that during the two cognitive acts he was not thinking about the same entity. So given almost any two acts of cognizing in which a subject thinks about the same entity, the subject *might* not recognize the entity he thinks about in one act as the entity he thinks about in the other. But to allow for the this *possibility* of recognition failure, the advocate of Frege’s strategy must posit two actual “ways of thinking” that the subject might fail to “connect” (or might later “disconnect”).⁹ But this entails that for almost every cognizant act in which a subject thinks about an entity there must correspond a unique “way of thinking.” Consequently, if the individuation constraint is to be satisfied, “ways of thinking” must be individuated as finely as particular cognitive acts of thinking of entities. So, for example, John might have one (or several) “way of thinking” of Twain associated with his reading *Huckleberry Finn*, another with his reading *Tom Sawyer*, another with a certain visual experience of Twain, and yet another with a distinct visual experience of Twain, and so on.

If Frege’s theory is to satisfy the individuation constraint, and thereby preserve sentential compositionality, it must identify the posited extraordinary referents with such finely individuated “ways of thinking.” That is, it must analyze an occurrence of an ascription of the form “ n Vs that $\Phi(\alpha)$ ” as asserting that the subject referred to by n holds the attitude referred to by V toward a thought (sentence of mentalese, whatever) referred to by the embedded $\Phi(\alpha)$. The ascription is true if and only if this thought (sentence of mentalese, whatever) is composed of particular “ways” in which *the subject* thinks of the ordinary referents of the embedded terms.¹⁰ (I ignore ascriptions in which one or more of the embedded terms lacks an ordinary referent.) Consequently the referents of an embedded occurrence will be *particular to the occurrence*. Different embedded occurrences of α will have to refer to different “ways of thinking,” depending upon who the subject is, how they

⁹ Richard (1990, p. 184), Crimmins (1992, p. 78), and even Salmon (1986, p. 107) note that their posited entities must be individuated finely enough to account for recognition failure.

¹⁰ The amended version of Frege’s analysis we have arrived at seems to be what is suggested in Fodor (1978). It is a version of what Richard (1990) and Forbes (1990) refer to as the “subject oriented” analysis of attitude ascriptions. On this analysis the sense referred to by a referring term appearing in the complement clause of a true ascription is a “way” in which the *subject* thinks of the referent of the term. (Problems of course arise for “empty” embedded terms, and for ascriptions that have more than one person as subject, but nevermind.) A potential alternative to the subject oriented analysis is the “speaker oriented” analysis, which maintains that embedded terms refer to a way in which the *speaker* thinks of the referent of the embedded term. I here consider only the subject oriented proposal, as the speaker oriented proposal clearly gets the truth conditions of attitude ascriptions wrong, and thus does not solve the problem of opacity.

think of the ordinary referent of α , and what the particular circumstances of the ascription are.

But now Frege's analysis has been pulled too far in one direction. If the posited extraordinary referents are identified with such finely individuated "ways of thinking," then the legitimacy of our ordinary attitude ascribing practices is undermined.¹¹ In uttering attitude ascriptions ordinary speakers are not referring to one another's finely individuated "ways of thinking." Given how little we know about the details of our own and each other's mental states, we *could not* be referring to such esoteric entities.¹² Indeed, the value of folk psychology lies in the fact that it produces relatively accurate predictions and explanations in the *absence* of such specialized knowledge.

Consider an occurrence of

- (3) Dostoevsky believed that Shakespeare was a great author.¹³

Frege's analysis maintains that a speaker who utters (3) is attempting to use 'Shakespeare' as it appears in the complement clause to refer to a "way" in which the subject, Dostoevsky, thought of Shakespeare, and that an occurrence of (3) is true (or false) only if the speaker succeeds in so referring to a particular "way" in which Dostoevsky thought of Shakespeare. But ordinary speakers who are perfectly competent in uttering and interpreting (3) have little knowledge, or even belief, concerning the "ways" in which Dostoevsky thought of Shakespeare. Thus Frege's analysis requires too much for the truth of attitude ascriptions.

That Frege's analysis requires too much for the truth of attitude ascriptions is made evident when one considers the *felicity* conditions of attitude ascriptions. According to Frege's analysis, a speaker who utters (3) is, among other things, intending (perhaps "tacitly intending") to refer to a particular "way" in which Dostoevsky thought of Shakespeare, and thus a hearer knows the truth conditions of such an utterance only if she is able to identify which one of Dostoevsky's "ways of thinking of Shakespeare" is being referred to. These claims are extremely implausible, but let us grant them for the time being, so that we might consider what

¹¹ Frege himself was well aware of the problems that would arise if the sense expressed by a term was permitted to vary intersubjectively. In a letter to Jourdain Frege wrote,

Now if the sense of a name was something subjective, then the sense of the [sentence] in which the name occurs, and hence the thought, would also be something subjective, and the thought that one man connects with this [sentence] would be different from the thought another man connects with it; a common store of thoughts, a common science would be impossible. It would be impossible for something one man said to contradict what another man said, because the two would not express the same thought at all, but each his own. (Frege, 1980.)

¹² Schiffer (1992) presents this problem as "the meaning intention problem".

¹³ The objection presented briefly here is developed in more detail against Crimmins and Perry's (1989, 1992) "hidden indexical" analysis in Clapp (1995). Similar objections against Crimmins and Perry's "hidden indexical" analysis, and Richard's "overt indexical" analysis are presented in Saul (1997).

will happen when the speaker *fails* to refer to a relevant “way of thinking.” So suppose that the utterer of (3) fails to refer to a “way” in which Dostoevsky thought of Shakespeare. This might occur for a number of reasons: perhaps Dostoevsky did not think about Shakespeare at all, and thus there is no relevant “way of thinking.” Or perhaps, as seems likely given how little speakers know concerning one another’s “ways of thinking,” the speaker’s intentions fail to determine a *unique* “way” in which Dostoevsky thought of Shakespeare. To clarify the case, let us assume that Dostoevsky did not think about Shakespeare at all. Under this assumption, our judgments dictate that an occurrence of (3) would be *false*; if Dostoevsky did not think of Shakespeare at all, then he certainly did not believe that Shakespeare was a great author. But Frege’s analysis predicts that an occurrence of (3) would be *infelicitous*, as it would suffer from reference failure. (If one maintains, incorrectly in my view, that occurrences that suffer from reference failure are *false*, then our version of Frege’s theory would incorrectly predict that an occurrence of the *negation* of (3) would be false.) The general problem is that our version of Frege’s analysis requires too much for the truth of an occurrence of (3). If Dostoevsky believed that Shakespeare was a great author, if, say, he was disposed to declare sincerely that Shakespeare was a great author, and so on, then an occurrence of (3) is *true*. The speaker’s (alleged) intentions to refer to this or that particular “way of thinking of Shakespeare” are *irrelevant* to the truth conditions of her utterance.¹⁴

An advocate of Frege’s strategy might attempt to avoid the problem of requiring too much for the truth of ordinary attitude ascriptions by rejecting the idea that in making ordinary attitude ascriptions a speaker is intending to refer to a specific “way” in which the subject thinks of an entity. Instead it might be claimed that in making ordinary attitude ascriptions speakers are merely *describing* the subject’s “ways of thinking.”¹⁵ For example, an occurrence of (3) might be analyzed as asserting something like this:

(3*) $\exists w(P(w) \ \& \ \text{Believes}(\text{Dostoevsky}, w \wedge S(\text{'was a great author'})))$.

Here ‘*w*’ ranges over “ways of thinking of Shakespeare,” and ‘*S*()’ designates a function from expression types to their senses – for convenience I assume that predicates are associated with only one “way of thinking.” And ‘ \wedge ’ designates a concatenation device for “ways of thinking”; ‘ \wedge ’ designates a partial function from atomic “ways of thinking” to molecular “ways of thinking” (thoughts, sentences of mentalese, whatever). The predicate ‘*P*()’ designates a property, or constraint, on “ways of thinking.” For example ‘*P*()’ might designate something like, “is a way

¹⁴ Similar problems are raised by general ascriptions such as ‘Nobody believes that Twain wrote’ and ‘Everybody believes that Twain wrote’. The problem posed by such sentences is that there is no particular subject whose “way of thinking of Twain” could serve as the extraordinary referent of ‘Twain’.

¹⁵ Crimmins and Perry (1989) utilize versions of the descriptive analysis. Crimmins (1992), however, eschews the descriptive analysis in favor of a referential analysis.

of thinking of Shakespeare as being an author." If ' $P()$ ' is so interpreted, then (3*) is appropriately *false*, instead of infelicitous, in the situation where Dostoevsky has no thoughts about Shakespeare at all, since no "way of thinking" satisfying ' $P()$ ' exists. Moreover, the above descriptive analysis does not require the speaker to have intentions sufficient for determining a unique "way" in which Dostoevsky thought of Shakespeare. The above analysis requires only that the speaker have intentions sufficient for determining a (potential) property of Dostoevsky's "ways of thinking of Shakespeare."

An advocate of Frege's strategy is pulled toward the above sketched *descriptive* analysis because of the problem of preserving the legitimacy of our ordinary attitude ascribing practices. The problem is that if "ways of thinking" are individuated as finely as they must be to satisfy the individuation constraint, then, because ordinary speakers lack detailed knowledge concerning one another's "ways of thinking," Frege's analysis will incorrectly predict that many true (or false) occurrences of attitude ascriptions suffer from reference failure. The descriptive analysis is invoked to make up for the fact that ordinary speakers lack the knowledge, beliefs, or even abilities, required for *referring* to one another's particular "ways of thinking." In effect, the descriptive analysis individuates the posited "ways of thinking" more coarsely in an attempt to preclude reference failure and/or to ensure that the subjects of attitude ascriptions will have a requisite "way of thinking." Understood in this way then it is not surprising that the problem of opacity re-emerges for the *descriptive* analysis. That is, if Frege's theory adopts the descriptive analysis, then it cannot solve all instances of the problem of opacity. Suppose that Dostoevsky in fact thought that Shakespeare was a lousy author. Suppose that he regularly asserted that Shakespeare was a lousy author, and so on. Under this assumption our semantic intuitions dictate that (3) is false. Further suppose, however, that Dostoevsky was once shown a picture of someone, and told that the person in the picture was a great author. Dostoevsky had no reason to doubt what he was told, and thus he came to believe that the author depicted, whoever it was, was a great author. Unknown to Dostoevsky the author depicted in the picture was, of course, Shakespeare. Thus there is some "way of thinking of Shakespeare" w that satisfies ' $P()$ '; i.e. Dostoevsky instantiates some "way of thinking of Shakespeare" w that is a "way of thinking of Shakespeare as being an author." Moreover, w is a constituent in a thought (sentence of mentalese, whatever) expressing that Shakespeare was a great author, and this thought is believed by Dostoevsky. So in this situation (3*) is true, though (3) is false. Hence (3*) is not an adequate analysis of (3). The *descriptive* analysis attempts to preclude the problem of reference failure and thereby takes steps toward preserving the legitimacy of our ordinary attitude ascribing practices, but in so doing it becomes susceptible to the problem of opacity.

The advocate of Frege's strategy might respond by pointing out that the above objection depends upon an unfortunate interpretation of ' $P()$ ', and thus such problems can be avoided so long as ' $P()$ ' is always appropriately interpreted to include

certain "ways of thinking," yet exclude others. It is probably correct that if one *assumes* our semantic intuitions concerning an occurrence of an attitude ascription, one can, based upon these intuitions, construct an appropriate interpretation for ' $P()$ '. But if this is how the interpretation of ' $P()$ ' is to be fixed, our version of Frege's theory fails to provide any sort of *explanation* of the phenomenon of opacity. The advocate of Frege's theory cannot explain and/or predict that occurrences of (1) and (2) have different truth conditions *because* they invoke different properties of "ways of thinking," and then justify the claim that they invoke these different properties by appealing to the fact that (1) and (2) have different truth conditions.¹⁶ If the advocate of Frege's strategy is to justify his claim that in uttering attitude ascriptions ordinary speakers are somehow designating *just the right* properties of "ways of thinking," he must show that *something* about the communicative intentions of the speaker serves to determine *which* property of "ways of thinking" is designated. But the advocate of Frege's strategy cannot do this, for ordinary speakers simply lack the requisite intentions. It is no more plausible to suggest that in uttering attitude ascriptions ordinary speakers intend to designate specific properties of "ways of thinking" than it is to suggest they intend to refer to individual "ways of thinking." Saul puts the point exactly right: "The only plausible intention which could yield the right readings... would be an intention to get the truth conditions right. While this surely is an intention of any speaker who is concerned with truth conditions, allowing this intention to determine semantic content merely evades, rather than solves, the problem of propositional attitude semantics" (1997, p. 435).

A subject can think of an entity in myriad different "ways"; the more a subject knows, or believes, about an entity, the more "ways of thinking about the entity" he will employ. According to analyses that utilize Frege's strategy, if a speaker is to make a true attitude ascription concerning what a subject believes (etc.) about an entity, he must be able to either refer to an individual "way" in which the subject thinks of the entity, or to designate a specific property that is possessed by just the right individual "ways" in which the subject thinks of the entity. But given how little ordinary speakers know, or even believe, concerning one another's "ways of thinking" ordinary speakers will rarely if ever possess such ability.

3. The general difficulty and Richard's overt indexical analysis

The general difficulty described above will plague *any* analysis that utilizes Frege's strategy, but I cannot establish this by examining every analysis that utilizes Frege's strategy. So in this section I will provide evidence in support of this general claim by examining Richard's (1990) ingenious indexical analysis and illustrating

¹⁶ This response on behalf of Frege's strategy in effect violates the *independence* condition on the posited extraordinary entities. See note 7 above.

how the general difficulty arises for it. I have chosen Richard's analysis because Richard is sensitive to the general difficulty with Frege's strategy. In fact, Richard (1990) argues that his overt indexical analysis is superior to Frege's analysis because it does not fall prey to these sorts of difficulties.¹⁷ And in his more recent (1997) paper, Richard amends his earlier analysis in response to difficulties similar to those described above. Thus my purpose in this section is, first, to demonstrate that despite what is claimed for it Richard's (1990) analysis falls to the general difficulty with Frege's strategy, and second, to demonstrate that the amendments proposed in Richard (1997) fail solve the problem.

Let us consider Richard's (1990) proposal in terms of the following example, which I borrow from Richard (1997). Suppose that late at night you and Mary are observing a man whom you both know to be Ortcutt. Ortcutt is dressed in black, and he is sneaking around Ralph's house. Ralph, who is inside his house, looks out his window and seems to catch a fleeting glimpse of Ortcutt, who immediately ducks behind a bush. You then observe Ralph calling the CIA. Mary now utters

(4) Ralph believes that Ortcutt is a spy.

On Richard's overt indexical analysis Mary's utterance is true if and only if the complement clause of the utterance *appropriately translates* some sentence of mentalese on Ralph's "belief blackboard." And whether or not the complement clause does *appropriately translate* some such sentence of mentalese varies with context. According to Richard "context supplies some 'instructions for translation' to get from the mental idiom of [Ralph] to the spoken idiom of [Mary]" (1997, p. 106). So the context of Mary's assertion might determine that, for Ralph, 'Ortcutt' is to be used to translate or represent ω , where ω is a "way of thinking" of Ortcutt consisting of Ralph's recent fleeting glimpse of a man in his yard. Thus Mary's utterance of (4) in this context is true if and only if, roughly, Ralph has a sentence of mentalese on his belief blackboard that expresses the Russellian proposition that Ortcutt is a spy, and in that sentence of mentalese ω plays the role of 'Ortcutt'. (And a similar account applies to the other terms in the complement clause of (4).)

Let us be slightly more precise, so that it is apparent how Richard's (1990) account can preserve sentential compositionality. Richard maintains that 'believes' is an indexical – it designates different relations in different contexts. Glossing some irrelevant details, Richard can be interpreted as claiming that, relative to a context, 'believes' designates a relation between believers and ordered pairs of sentences and Russellian propositions, where for each such pair the first member – the sentence type – expresses (in the context) the second member – the Russellian proposition.¹⁸ Let 'believes_c' stand for the particular relation designated

¹⁷ See Richard (1990), Chapters 2 and 3.

¹⁸ Richard actually states (1990, p. 142) that occurrences of 'believes', etc., designate relations between agents and RAMs (Russellian Annotated Matrices). RAMs are structures composed of pairs of

by 'believes' in a context c . According to Richard context c determines a set t of translation instructions for every relevant subject (believer) b . A set t of translation instructions for a believer b will specify which of b 's "ways of thinking" are represented by various natural language expressions. For example, a set of translation instructions for John determined in a context c might specify that in c 'Twain' represents, or translates, w_1 , w_2 , and w_3 , where w_1 , w_2 , and w_3 are expressions of mentalese referring to Twain, or "ways in which John thinks of Twain." Say that set of instructions t for a believer b translates b 's mentalese sentence m just in case the instructions in t specify expressions of natural language that can be used in c to represent, or translate, m . Now we can state that a believer b stands in believes $_c$ to a sentence-proposition pair $\langle s, r \rangle$ if and only if there is some sentence of mentalese m on b 's belief blackboard such that m expresses r and the instructions t for b translate m . So, in terms of our example, Mary's utterance of (4) is true just in case Ralph holds the particular believes relation designated in the context toward the Russellian proposition expressed by 'Ortcutt is a spy'. Suppose, as was assumed above, that the operant translation instructions for Ralph require, among other things, that 'Ortcutt' be used to translate w , where w is a "way of thinking" of Ortcutt consisting of Ralph's recent fleeting glimpse of a man in his yard. Under these assumptions, Mary's utterance of (4) is true if and only if, roughly, Ralph has a sentence of mentalese on his belief blackboard that expresses the Russellian proposition that Ortcutt is a spy, and w plays the role of 'Ortcutt' in this sentence of mentalese.

In Richard's analysis the relations designated by 'believes' ('doubts', etc.) vary across contexts because the operant sets of *translation instructions* are essential to the designated relations, and what translation instructions are operant varies from context to context. Consequently, in order for one to understand Mary's utterance of (4), one must know the semantic value of the relevant occurrence of 'believes'; i.e. one must know what relation is designated by this occurrence. But to know what relation is designated by this occurrence, one must know what translation instructions are operant in the context. And in order to know (perhaps tacitly) that a particular set of translation instructions are operant for a subject, one must know that the subject utilizes the "ways of thinking" that are specified in these instructions. For example, if one is to know that in a context c , for a believer b , 'Twain' translates any of w_1 , w_2 , or w_3 , then one must know that w_1 , w_2 , and w_3 are "ways in which b thinks of Twain." But the claim that such knowledge is required for competently uttering and interpreting attitude ascriptions is extremely implausible, for competent speakers typically lack any sort of detailed knowledge (or belief) concerning one another's "ways of thinking." Moreover, it cannot be that such knowledge is required, for the truth conditions of attitude ascriptions simply do not depend upon such translation instructions.

term types and associated semantic values, and thus are more complex than the sentence-proposition pairs I use in explicating Richard's view. But the additional complexity is not relevant to my purposes here, so I ignore it for sake of simplicity.

Suppose that the impression you and Mary had of the situation involving Ralph and Ortcutt was somewhat mistaken. Perhaps Ralph did not catch a fleeting glimpse of Ortcutt; perhaps Ralph's wife, unseen by you and Mary, saw Ortcutt, and she subsequently alerted Ralph of the trespasser's presence. So no such visual "way of thinking" as w exists. Consequently either (i) Mary's utterance of (4) is predicted to be false, as Ralph does not have an appropriate sentence of mentalese on his belief blackboard, or, what I think is more plausible, (ii) Mary's utterance is infelicitous, as Mary's intentions fail to determine an appropriate set of translation instructions, and as a result the relevant occurrence of 'believes' fails to designate an appropriate relation. But both predictions are incorrect. If at the time of Mary's utterance Ralph is prone to utter sincerely things like 'That guy out there is a spy!' and so on, then Mary's utterance of (4) is true. Granted, Mary is slightly confused since she incorrectly thinks that Ralph caught a fleeting glimpse of Ortcutt when he actually did not. But in uttering (4) Mary is not asserting a belief to the effect that Ortcutt recently caught a glimpse of a man in his yard; in uttering (4) Mary is asserting her belief that Ralph believes Ortcutt to be a spy.

Or suppose Ralph actually caught several discontinuous glimpses of shadowy figures flitting about his yard, so unbeknownst to you and Mary, there is not a unique "visual way of thinking" of Ortcutt w instantiated by Ralph, but rather several such "visual ways of thinking." If this, very likely, circumstance obtains, then Mary's intentions again do not determine an appropriate translation manual for the context, as there is no *unique* visual "way of thinking" w . And consequently Mary again fails to designate an appropriate relation by her utterance of 'believes' in (4). So again Mary's utterance would be predicted to lack a truth value. The problem, which should by now be familiar, is that Richard's (1990) overt indexical proposal requires too much for the truth of attitude ascriptions. As a result it requires ordinary speakers to know far more about the details of one another's mental lives than they typically do know.¹⁹

If Richard is to avoid this sort of problem, he must somehow preclude such misconceived *translation instructions* from ever being determined by a context. That is, he must not require ordinary speakers to have such detailed knowledge of one another's individual "ways of thinking." One way to do this would be, again, to make the translation instructions *descriptive*; instead of specifying how individual "ways of thinking" can be represented, translation instructions might describe how

¹⁹ Richard's (1990) proposal requires ordinary speakers to know what natural language expressions can be used to translate *representations*, where a *representation* is, roughly, a class of *unified* individual "ways of thinking." Two "ways of thinking" are *unified* for a subject just in case he "connects" them, i.e. just in case he takes them to be representations of the same entity. (See Richard, 1990, pp. 184–190.) Thus *representations* are even more esoteric than are "ways of thinking." If translation instructions tell a speaker/listener how such *representations* can be translated with natural language expressions, then knowing what relation is designated by an occurrence of 'believes' will require the speaker/listener to know about *all* of a subject's "ways of thinking," and to know which one's are *unified* with which other ones. Requiring this much knowledge is extremely implausible.

kinds or *types* of “ways of thinking” can be represented. For example, perhaps the operant translation instructions for Mary’s utterance of (4) do not require that ‘Ortcutt’ be used to represent an individual “way of thinking” such as *w*, but instead require merely that ‘Ortcutt’ be used to represent, say, any one of Ralph’s “ways of thinking” of Ortcutt that was somehow caused by a perception of Ortcutt. As our story goes, Ralph does utilize a “way of thinking” of Ortcutt of this kind: Ralph’s wife did see Ortcutt in the yard, and this perception caused her to say something to Ralph, and this saying in turn caused Ralph to instantiate a particular “way of thinking” of Ortcutt. Hence, if at the time of Mary’s utterance Ralph is prone to say things like, ‘That guy out there is a spy!’ and so on, then this revised account correctly predicts that Mary’s utterance of (4) is true.

What the revised *descriptive* account essentially does, again, is to make the “ways of thinking” invoked by Richard’s analysis more coarsely individuated, and therefore more accessible to ordinary speakers who have very limited knowledge of one another’s mental states. But we know from the above discussion of Frege’s analysis that by more coarsely individuating “ways of thinking,” the proposed solution to the problem of opacity is undermined. The same point applies, *mutatis mutandis*, to the descriptive version of Richard’s analysis. Suppose that the translation instructions determined by the context are descriptive, as suggested above: suppose the operant translation instruction require, among other things, that ‘Ortcutt’ be used to represent a “way of thinking” of Ortcutt that was somehow caused by a perception of Ortcutt. Then Mary’s utterance of (4) is predicted to be true if and only if, roughly, there is a sentence of mentalese on Ralph’s belief blackboard that expresses the Russellian proposition that Ortcutt is a spy and is such that a “way of thinking” of Ortcutt that was somehow caused by a perception of Ortcutt plays the role of ‘Ortcutt’ in this mentalese sentence. But now suppose that Mary’s utterance is false; suppose that at the time of Mary’s utterance Ralph is *not* prone to say things like, ‘That guy out there is a spy!’ and so on. Suppose he thinks the guy out there is an escaped criminal, and he calls the CIA because, in his anxious state, he confuses the CIA and FBI. Also suppose, however, that five years ago Ralph’s friend Smiley who worked for MI5, pointed Ortcutt out to Ralph and said, ‘that guy is the best spy we’ve got’. (Ortcutt was of course wearing a clever disguise at the time.) Ralph believed his friend Smiley, and thus there is (still) a sentence of mentalese on Ralph’s belief blackboard that expresses the Russellian proposition that Ortcutt is a spy and is such that a “way of thinking” of Ortcutt that was somehow caused by a perception of Ortcutt plays the role of ‘Ortcutt’ in this mentalese sentence. Hence relative to this situation the proposed *descriptive* version of Richard’s analysis incorrectly predicts that Mary’s utterance of (4) is true. By making the translation instructions more general and therefore more accessible to ordinary speakers, Richard’s solution to the problem of opacity is undermined.

One might respond on behalf of the descriptive version of Richard’s (1990) analysis by pointing out that the above objection succeeds only because of an unfortunate choice of property, or “way of thinking” kind. If Mary’s utterance is

interpreted as adverting to a property of "ways of thinking" that is more discriminating than *being a way of thinking of Ortcutt somehow caused by a perception of him*, then Mary's utterance of (4) can be assigned appropriate truth conditions. But, again, if Richard's defender is to justify his claim that in uttering attitude ascriptions ordinary speakers are adverting to *just the right* properties of "ways of thinking," he must show that *something* about the communicative intentions of the speaker serves to determine which property of "ways of thinking" features in the operant translation instructions. But, again, ordinary speakers simply lack the requisite intentions for determining such a property. In uttering (4) Mary does not intend to advert to any particular property of Ralph's "ways of thinking" of Ortcutt, and thus she does not advert to any such "way of thinking."

In a recent paper Mark Richard presents and responds to an objection similar to the one presented above.²⁰ The objection Richard considers is this:

(a) There is practically *no* use of an attitude ascription where it is determinate that a particular representation or property thereof is the intended object of reference (or adversion). (b) A meaningful expression cannot be mired in this sort of indeterminacy. In particular, if uses of a device generally fail to achieve determinate reference, then the device isn't referential. But (c) if contextualism and (a) are correct, then predicates such as 'believes that Ortcutt is a spy' are mired in such indeterminacy. So contextualism is wrong. (Richard, 1997, p. 105.)

(Richard's term 'contextualism' broadly refers to analyses of attitude ascriptions according to which different finely individuated "ways of thinking" or properties thereof are "adverted to" in different contexts. Hence our amended version of Frege's analysis, Crimmins and Perry's hidden indexical analysis, and Richard's overt indexical analysis are all *contextualist*.)

Richard presents a number of objections against this argument,²¹ but I will focus on his rejection of (c). Richard rejects (c) on the grounds that even if something

²⁰ Richard's (1997) paper presents a number of criticisms of the broadly Fregean analysis of propositional attitudes proposed in Devitt (1996). Devitt responds to Richard's criticisms in Devitt (1997). This exchange between Richard and Devitt perfectly illustrates the general difficulty with Frege's strategy. Richard argues, quite correctly, that if the posited extraordinary entities are to do what they are posited to do, viz. solve the problem of opacity, they must be individuated very finely, more finely than Devitt's "*d*-senses." And Devitt argues, also quite correctly, that Richard's analysis, which posits very finely individuated extraordinary entities, is unable "to explain how hearers use linguistic conventions together with accessible context to understand ascriptions" (Devitt, 1996, p. 199; 1997, p. 124).

²¹ Richard rejects (a) on the grounds that there are many occurrences of attitude ascriptions where "a particular property of [ways of thinking] is distinguished enough that it may... be identified as an object of reference" (p. 105). Richard does not, however, adequately support this claim. Richard states, "If I say 'After she heard the lecture on Cicero and read a poem signed 'Tully' she believed that Cicero orated, but not Tully,' it is clear to [what sort of "way of thinking"] I mean to advert..." (p. 105). But it is not clear. Even if we assume that Richard's utterance "advert[s]" to some sort of "way of thinking" of Cicero that she acquired during the lecture, it is not clear, for she no doubt acquired many sorts of "ways of thinking of Cicero" during the lecture. Which acquired during the lecture sort of "way of thinking"

like Crimmins and Perry's hidden indexical theory fails because ordinary speakers lack the knowledge and intentions required for referring to "ways of thinking," or properties thereof, still there will be ways of utilizing Frege's strategy that do not suffer from this problem. In what follows I briefly explicate Richard's more recent (1997) proposal, and demonstrate that it fares no better than his earlier (1990) proposal.

In his more recent (1997) paper Richard seems to concede that the above considered descriptive version of the overt indexical analysis would require the speaker, Mary, to have intentions she simply does not have. He suggests, quite rightly, that Mary "need not have a particular way of identifying Ralph's representations in mind; she means only to be relaying something about Ralph's attitudes..." (p. 106). In response to the problem posed by Mary's lack of the requisite intentions, Richard proposes what I will call the "unintended majority" version of his overt indexical analysis. Richard claims that despite Mary's lack of the requisite intentions, nonetheless "there can still be a collection *C* of 'best candidates' for the property of representations which, in Mary's context, one of Ralph's Orcutt representations needs to have, in order to be acceptably translated using 'Orcutt'. Roughly, *C* is the collection of properties expressed by what Mary would offer as ways to fill in the ellipsis in *This belief of Ralph's is one which involves a way of thinking of Orcutt ...*" (p. 106). Richard goes on to suggest that *being a representation caused by perceiving him*, and *being a representation of Orcutt involved in the belief that caused Ralph to call the police* are plausible candidates to be in *C*. On this latest proposal the translation instructions do not specify how *individual* "ways of thinking" are to be translated. Nor do they even specify how *kinds* of "ways of thinking" are to be translated. Richard concedes that such translation instructions could not be determined by the context, as Mary lacks the requisite knowledge and intentions. Rather the translation instructions appeal to sets of "best candidate" properties of "ways of thinking," where unmanifested dispositions of the speaker determine such sets of properties. Richard proposes the following translation instruction as an example:

So far as Ralph is concerned, use 'Orcutt' to translate a representation of Orcutt that has most of the properties in *C* (p. 107).

Thus, under the unintended majority version of Richard's analysis, a context determines two things: First, it determines – via the speaker's unmanifested dispo-

is adverted to? Aurally acquired "ways of thinking" of Cicero she acquired during the lecture? Aurally or visually acquired "ways of thinking" of Cicero she acquired during the lecture? Suppose the lecturer uttered the name 'Tully', with an ancient Roman pronunciation, several times during the lecture. Are the "ways of thinking" she acquired as a result of such utterances to be included, or excluded from the type adverted to? If they are excluded, how are they excluded? It is not at all clear what sort of "way of thinking" Richard would be "adverting" to with his utterance of 'After she heard the lecture on Cicero and read a poem signed 'Tully' she believed that Cicero orated, but not Tully'. The truth conditions of such an utterance, however, would be relatively clear. And so much the worse for contextualism.

sitions – sets of best candidate properties of “ways of thinking.” Second, it determines translation instructions that constrain how natural language expressions can be used to represent “ways of thinking” that have *most* of the properties in this set. On this proposal Mary’s utterance is true if and only if, roughly, there is some sentence of mentalese on Ralph’s belief blackboard that expresses the Russellian proposition that Ortcutt is a spy, where the individual “way of thinking” playing the role of ‘Ortcutt’ in this sentence of mentalese has *most* of the properties in the determined candidate set.

The unintended majority version of the overt indexical analysis is well motivated in that it attempts to avoid saddling ordinary speakers with intentions that they do not, and cannot, possess. It is unlikely, however, that ordinary speakers will *always* have the sorts of dispositions that Richard supposes would determine the sets of best candidate properties. If Mary is a typical speaker, she may not be disposed to fill in the relevant ellipsis in any way at all. Moreover, it is even more unlikely that ordinary speakers, or *hearers* for that matter, have *knowledge* (even tacit knowledge) of these dispositions that would enable them to determine the relevant set of best candidate properties. And since such knowledge would be required for determining the truth conditions of occurrences of attitude ascriptions, Richard’s unintended majority proposal is incompatible with the fact that ordinary language users typically *do* know the truth conditions of attitude ascriptions. But beyond these problems, or rather *because* of these problems, Richard’s unintended majority proposal simply fails to make correct predictions concerning the truth conditions of attitude ascriptions.

First, let us suppose, as seems likely, that the number of best candidate properties in *C* is relatively large, say around fifty or one hundred. The larger the number of properties in *C*, the more likely it is that *none* of the subject’s relevant “ways of thinking” will have *most* of the properties in *C*. In terms of Mary utterance of (4), suppose that the context determines a set *C* that contains exactly one hundred properties. But suppose that none of the individual “ways of thinking” of Ortcutt utilized by Ralph has more than fifty of these properties. In this case the unintended majority version of Richard’s analysis predicts that Mary’s utterance is false; Ralph does not hold the attitude designated by Mary’s utterance of ‘believes’ toward the pair (‘Ortcutt is a spy’, *p*), where *p* is the Russellian proposition that Ortcutt is a spy. But Mary’s utterance could well be true; if at the time of Mary’s utterance Ralph is prone to say things such as “That guy out there is a spy!” and so on, then Mary’s utterance is true, regardless of how many properties in *C* are possessed by Ralph’s “ways of thinking” of Ortcutt. Putting many properties in *C* is in effect to individuate “ways of thinking” very finely, and, as we have seen, when this is done the legitimacy of our ordinary attitude ascribing practices is undermined.

If, on the other hand, it is claimed that *C* contains relatively few best candidate properties, then the proposal will be unable to account for all instances of the problem of opacity. To make *C* contain fewer properties is in effect to individuate “ways of thinking” more coarsely. And, as we have seen, when an analysis utilizing

Frege's strategy coarsely individuates the posited "ways of thinking," the ability of the analysis to solve the problem of opacity is undermined. Suppose that John does not know that Twain is Clemens. He thinks, as he would put it, that both men were authors, but he thinks that Twain wrote *Huckleberry Finn*, while Clemens wrote *Tom Sawyer*. Moreover, John believes that Clemens smoked, but that Twain did not smoke; he is prone to utter sincerely things such as 'Clemens smoked', and 'Twain did not smoke', and so on. And further suppose that Ralph and Mary know that John does not believe that Twain and Clemens are the same person, but, as is typical, they do not know, nor even believe, much more than that; they assume that John believes that Twain was a famous author, but they really have no idea what John thinks of Clemens. Now, for whatever reason, Mary utters

(5) John believes that Twain smoked.

Assuming the above description of John's mental states, Mary's utterance of (5) is false. Given Mary's limited knowledge and assumptions concerning John's mental states, it is as plausible as any other proposal that *C* contains the following best candidate properties of John's "ways of thinking" of Twain:

- C*:
- (a) Being a "way of thinking" of Twain involved in a belief toward the Russellian proposition that Twain was an author.
 - (b) Being a "way of thinking" of Twain involved in a belief toward the Russellian proposition that Twain wrote *Huckleberry Finn*.
 - (c) Being a "way of thinking" of Twain involved in a belief toward the Russellian proposition that Twain wrote *Tom Sawyer*.

According to the unintended majority version of Richard's analysis, Mary's utterance of (5) is true if and only if, roughly, there is a sentence of mentalese on John's belief blackboard that expresses the Russellian proposition that Twain smoked, where the individual "way of thinking" that plays the role of 'Twain' in this sentence of mentalese has *most* of the properties in *C*. So in this situation the unintended majority proposal makes the wrong prediction concerning Mary's utterance of (5): because John believes that Clemens smoked, and that Clemens was an author who wrote *Tom Sawyer*, there is a sentence of mentalese on John's belief blackboard that expresses the Russellian proposition that Twain smoked where the individual "way of thinking" that plays the role of 'Twain' in this sentence of mentalese has *most* of the properties in *C*. More specifically, because John believes that Clemens smokes and that Clemens was an author who wrote *Tom Sawyer*, John utilizes a relevant "way of thinking" of Twain that has properties (a) and (c) in *C*, and thus he does utilize a relevant "way of thinking" of Twain that has *most* of the properties in *C*. And thus the unintended majority proposal incorrectly predicts that Mary's utterance of (5) is true.

Richard might respond to these last two objections against his unintended majority proposal by noting that they presuppose an unfortunate set *C* of best candidate properties, and thus such problems could be avoided so long as appropriate sets *C* of best candidate properties are always somehow determined by the context. But, again, this response is not really to the point. It is probably correct that if one assumes our semantic intuitions concerning an occurrence of an attitude ascription, one can, based upon these intuitions, construct an appropriate set *C* of best candidate properties. But, again, if this is how an appropriate set *C* of best candidate properties is to be determined, Richard's analysis fails to provide any sort of *explanation* of the phenomenon of opacity. Richard cannot explain and/or predict that occurrences of (1) and (2) have different truth conditions *because* the contexts of the occurrences determine different appropriate sets *C* of best candidate properties of "ways of thinking", and then justify the claim that they invoke these appropriate sets by appealing to the fact that (1) and (2) have different truth conditions. Moreover, while it is probably true that an appropriate set *C* *can always be constructed*, there is no reason to believe that an appropriate set *C* *will always be determined*. That is, there is no reason to believe that the set *C* of best candidate properties actually determined in a context, together with the detailed facts concerning the subject's mental state, will *always* yield appropriate truth conditions. Given how little ordinary speakers know about the details of one another's mental lives, it seems likely that *often* their dispositions would determine *inappropriate* sets of best candidate properties.

I have not established that every analysis that utilizes Frege's strategy will fall to the general difficulty presented above. But the above discussion provides very strong evidence in support of this conclusion. The general difficulty with Frege's strategy is that the identity conditions for the posited extraordinary entities are pulled in incompatible directions. In order to account for all instances of the problem of opacity, the posited entities must be identified with very finely individuated "ways of thinking". But if this is done, then the legitimacy of our ordinary attitude ascribing practices is undermined. Even if it is granted that such "ways of thinking" exist, ordinary language users simply do not have the knowledge required for identifying, describing, or referring to such entities. But, conversely, if the posited entities are more coarsely individuated, then an analysis that utilizes Frege's strategy cannot account for all instances of opacity. This general difficulty cannot be avoided by identifying "ways of thinking" with different sorts of entities. Nor can it be avoided by positing another new semantic mechanism – in addition to "reference shifting", "hidden indexicals", or analyzing attitude verbs as overt indexicals – to account for how such "ways of thinking" (or properties thereof) are invoked by attitude ascriptions. This strongly suggests that Frege's strategy is inadequate, and an alternative strategy for responding to the problem of opacity ought to be formulated.

4. An alternative strategy for responding to the problem of opacity

The proper response to the problem of opacity is not to posit extraordinary entities and special referential mechanisms to somehow achieve the result that occurrences of (1) and (2) invoke different semantic values. Rather the proper response is to deny that the truth conditions of occurrences of ascriptions such as (1) and (2) are a function of only (i) their syntactic structures, and (ii) the semantic values they invoke. If sentential compositionality is rejected, then it is not *especially* problematic that occurrences of (1) and (2) have the same syntactic structure and invoke the same semantic values in the same order, yet have distinct truth conditions.

Two kinds of evidence are needed to support this response to the problem of opacity. First, it ought to be demonstrated that semantic phenomena independent of attitude ascriptions are incompatible with sentential compositionality. For if this cannot be demonstrated, the rejection of sentential compositionality can be legitimately objected to on the grounds that it is *ad hoc*. Second, and more importantly, an alternative analysis of attitude ascriptions that does not presuppose sentential compositionality must be formulated and shown to be superior to its rivals. I will not here attempt to provide evidence of the first kind.²² I will, however, take some preliminary steps toward providing evidence of the second kind by sketching an analysis of attitude ascriptions that does not presuppose sentential compositionality.

Attitude ascriptions are very context sensitive. Even if the facts concerning a subject's mental states are held fixed, there are some contexts of utterance in which an attitude ascription concerning the subject would be true, and other contexts in which an occurrence of the same ascription, referring to the same subject, would be false. Creating these *extreme* instances of opacity, however, requires a significant amount of what Devitt (1997) calls "stage setting". In order for us to assign distinct truth conditions to different occurrences of the same ascription, the occurrences must be embedded within carefully constructed discourses of the sort that philosophers excel at constructing. Moreover, creating even a *standard* instance of the phenomenon of opacity requires a bit of clever story telling. Even getting firm judgments that occurrences of sentence pairs such as (1) and (2) have distinct truth conditions requires a significant amount of "stage setting". I propose that the stage setting essential to the context sensitivity of attitude ascriptions be taken much more seriously. I propose that the preceding discourse environment of an occurrence of an attitude ascription (and other sorts of sentences as well) is an essential semantic feature of the occurrence: The occurrence cannot be given an adequate semantic analysis independently of this previous discourse environment, *even after all the semantic values invoked by the occurrence are determined*. On this proposal the previous discourse environment does not merely serve to facilitate

²² Such evidence against sentential compositionality can be found in Travis (1985), Carston (1988), Recanati (1989), Sperber and Wilson (1986), Bach (1994), Stainton (1994), Culicover and Jackendoff (1995), and Saul (1993).

assignment of semantic values to the relevant syntactic elements; rather the previous discourse environment itself is essential to determining the truth conditions of a sentence occurring later in a discourse. Thus this proposal is incompatible with sentential compositionality: the truth conditions of an occurrence need *not* be a function of (i) the syntactic structure of the sentence, and (ii) the semantic values assigned to the semantically relevant syntactic elements in the sentence. Rather on this proposal, which I will call *discourse holism*, the truth conditions of an occurrence need only be a function (i) the syntactic structure of the occurrence, (ii) the semantic values assigned to the words and other semantically relevant syntactic elements in the sentence, and (iii) relevant features of the discourse environment.

The fundamental idea of *discourse holism* has been developed in some detail by theorists working within *Discourse Representation Theory* (DRT) and associated programs in *dynamic semantics*.²³ In what follows I will introduce and use, albeit in a simplistic and informal way, the machinery of DRT to illustrate how a discourse holistic approach can distinguish the truth conditions of occurrences of sentences very similar to (1) and (2).²⁴

The underlying idea of DRT is this: In the process of natural language discourse competent speakers construct a *context*, where a context is a set of beliefs that are mutually shared by the discourse participants. A context is a mutually constructed cognitive model of reality. The semantic effect of an utterance within a discourse is to amend the context being constructed, but exactly how an utterance affects the context relative to which it occurs is partly a function of what is already in the context at the time of utterance. The machinery of DRT is a means of representing formally, with "Discourse Representation Structures" (DRSs), how this recursive process of context construction occurs. At the heart of DRT is the "discourse construction algorithm" which constitutes a function from previous DRSs and current utterances to subsequent DRSs. This process is described by Kamp and Ryle as follows:

²³ The general idea of discourse holism was, I believe, first proposed by Stalnaker (1978). The formal machinery of "Discourse Representation Theory" is introduced in Kamp (1984), and developed in some detail in Kamp and Reyle (1993). A somewhat similar approach is developed in Fauconnier (1994). The DRT analysis of attitude ascriptions I sketch below owes much to Kamp (1990) and Asher (1986), though it also differs in a number of important respects. In particular, in the analysis sketched here DRSs represent (partial) mutually believed contexts, whereas Kamp and Asher take DRSs to represent (partial) cognitive states of individual discourse participants.

²⁴ The discourse holistic proposal has some affinity with the "descriptivist" account proposed by Bach (1997 and this volume). Bach claims that attitude ascriptions such as (1) and (2) "though semantically equivalent, are also *semantically incomplete*. That is, they do not express complete propositions, and to that extent they are like such sentences as ['Fred is ready']" (1997, p. 228). Yet Bach, unlike Crimmins and Perry, denies that the requisite completion involves reference to anything like "ways of thinking." According to Bach, the complement clause of a belief report does not *specify* the content believed, but merely *describes*, or *characterizes*, one of the subject's beliefs, where precisely how a given complement clause characterizes a belief varies from context to context.

... when the algorithm is applied to a sequence of sentences S_1, \dots, S_n it deals with sentences in order of appearance. It first incorporates S_1 into starting DRS K_0 , then it incorporates S_2 into DRS K_1 resulting from the first incorporation, etc. The first step of the process by which S_i gets incorporated into K_{i-1} consists in adding the syntactic analysis $[S_i]$ of S_i to the set of conditions of $K_{i-1} : \dots : K_{i-1}$ acts as a context of interpretation for S_i . (Kamp and Reyle, 1993, p. 85.)

Thus according to DRT semantic competence is knowledge of how to construct contexts in the process of natural language discourse. Or in Kamp and Reyle's terminology, what competent speakers know is how to construct DRSs from utterances of sentences and previous DRSs via the discourse construction algorithm.

In DRT truth conditions are directly assigned to DRSs, and thereby indirectly assigned to the utterances from which the DRSs resulted. Thus, since DRS K_{n+1} of a discourse is a function of the utterance of a sentence S_{n+1} together with the previous DRS K_n , the truth conditions of the utterance of S_{n+1} are *not* a function of only (i) the syntactic structure of S_{n+1} and (ii) the semantic values of the words and relevant syntactic features of S_{n+1} . Rather it is more accurate to say that in DRT the truth conditions of S_{n+1} are determined by the discourse construction algorithm as a function of (i) the syntactic structure of S_{n+1} and (ii) the semantic values of the words and relevant syntactic features of S_{n+1} , and (iii) the location of S_{n+1} in a discourse, or more specifically, the relations that obtain between S_{n+1} and the previous sentences in the discourse of which S_{n+1} is a part. According to DRT a discourse creates "a semantic 'web' that cannot in general be equated with a simple conjunction of propositions expressed by the individual sentences" (Kamp, 1990, p. 34).

I cannot here develop the proposal in detail, but discussion of a simple case will suffice to demonstrate that discourse holism constitutes a plausible alternative to Frege's strategy. Suppose that Ralph and Mary are engaged in a discourse which concerns their friend John and his confusion concerning the identity of Twain and Clemens. The following much simplified DRS (see DRS 1) represents the original context of the discourse. It represents some of the relevant mutual beliefs of Ralph and Mary.

DRS 1

	u	v	w
(c1) is named 'John'	(u)		
(c2) is named 'Twain'		(v)	
(c3) is named 'Clemens'			(w)
(c4) is a great author		(v)	
(c5) wrote <i>Huck Finn</i>		(v)	
(c6) (v) is the same person as (w)			(w)

The letters 'u', and 'v' and 'w' at the top of DRS 1 are *reference markers*, and the set $\{u, u, v\}$ is the "universe" of DRS 1. A DRS can be thought of as a sort of cognitive model, and reference markers are the things in the model. Thus reference markers are not referents, rather they are representations that may or may not correspond to actual referents. The reference markers also function in the process of DRS construction as "pegs" on which the discourse participants can "hang" property ascriptions. These property ascriptions are represented by the *conditions* (c1)–(c6). The conditions represent some of the things that the discourse participants believe about the entities that, at least allegedly, correspond to the discourse referents. If the mutual beliefs represented by the DRS are accurate, the reference markers in the universe of the DRS will correspond with actual individuals who have the properties represented by the conditions. Thus DRS 1 represents the mutual beliefs of Ralph and Mary that somebody named 'John' exists, that somebody named 'Twain' exists, that somebody named 'Clemens' exists, and the person named 'Twain' was a great author, and the person named 'Twain' wrote *Huck Finn*. Condition (c6) represents that Ralph and Mary mutual belief that Twain is Clemens, or more precisely, that the person named 'Clemens' is the person named 'Twain'.²⁵

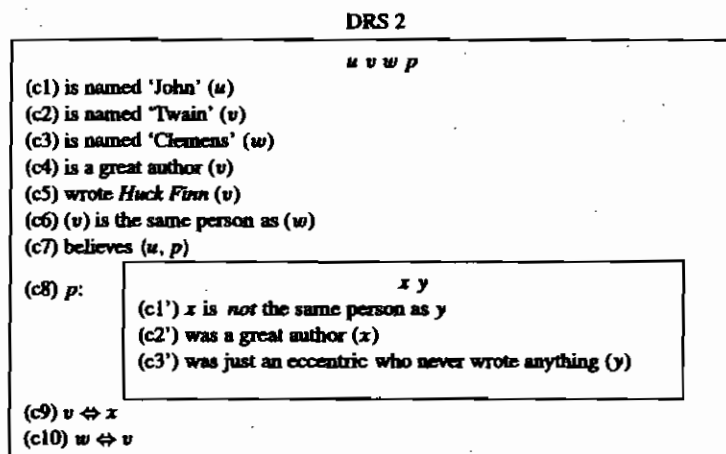
Now suppose that relative to the initial context represented by DRS 1, Ralph utters the following:

John thinks that Twain and Clemens are different people. He believes that Twain was a great author, but he thinks that Clemens was just an eccentric who never wrote anything.

Ralph's utterances change the context of the discourse between Ralph and Mary. The new context which is brought about by Ralph's utterance, as directed by the discourse construction algorithm, is represented by the following DRS (see DRS'2), which contains all the information contained in DRS 1 plus the information added by Ralph's utterances.

The discourse construction algorithm constructs DRS 2 from DRS 1 together with Ralph's utterances roughly as follows. From the first sentence Ralph utters the algorithm introduces the new reference marker *p* into the universe of the DRS, and creates (c7) and (a preliminary version of) (c8). Reference marker *p* represents an (alleged) attitudinal state, or at least a component of a subject's attitudinal state. Condition (c7) in effect states that the (alleged) attitudinal state represented by '*p*' is a belief state of John's. Condition (c7), which introduces a *subDRS*, states a condition on this attitudinal state; it states the "content" of John's attitudinal state. A *subDRS* is required in (c7) because Ralph's first utterance is an attitude

²⁵ There are some complicated issues raised here, as believing that somebody named 'Twain' exists is not the same thing as believing that Twain exists. One might say that the former is *de dicto*, while the latter is *de re*. These difficulties can be resolved in the framework of DRT (see Kamp, 1990), but they are not relevant to the fundamental point I am making here. So for the sake of simplicity I ignore them.



ascription, and hence what it expresses is Ralph's beliefs about John's beliefs, and thus what results is a subDRS – a cognitive model of a cognitive model.

Ralph's second utterance further develops the subDRS in (c7). That the occurrences of the pronoun 'he' in Ralph's second utterance is interpreted as coreferring with 'John', and not 'Twain' or 'Clemens', is implicitly represented in DRS 2 by the conditions (c1')–(c3') being introduced into the subDRS describing u 's (i.e. John's) attitudinal state. In DRT every occurrence of a definite noun must be associated with a reference marker already present in the previous DRS, otherwise the utterance containing the definite noun is infelicitous. This is known as *the familiarity constraint*. The conditions (c1')–(c3') represent that Ralph and Mary mutually believe that John believes that there are two people, one of whom wrote *Huck Finn*, the other of whom was just an eccentric who never wrote anything. Note that the subDRS does *not* represent that John is familiar with the names 'Twain' and 'Clemens'. This is as it should be: Ralph's utterances could be true, for example, even if John has never heard the names 'Twain' and 'Clemens', and is familiar only with the first names 'Mark' and 'Samuel'.

The embedded occurrences of 'Twain' and 'Clemens' in Ralph's utterances introduce *new* reference markers x and y into the subDRS of (c8'). Sentences containing *unembedded* occurrences of 'Twain' and 'Clemens' would have merely introduced new conditions on v and w . But because the occurrences of 'Twain' and 'Clemens' in Ralph's utterances are embedded in attitudinal complement clauses, they cannot merely be associated with already present reference markers u and v . For this would represent that John shares Ralph and Mary's mutual beliefs concerning Twain; but Ralph's utterances *contrast* John's beliefs about Twain with those of himself and Mary. Ralph's utterances, however, are nonetheless about John's beliefs concerning *Twain*, the guy believed by Ralph and Mary to be the

same person as Clemens. Moreover, the familiarity constraint requires that the embedded occurrences of the definite nouns 'Twain' and 'Clemens' in Ralph's utterances be associated with reference markers already present in DRS 1. For these reasons the reference markers x and y of the subDRS are *linked* via (c9) and (c10) to reference markers v and w of the main DRS, which are in turn associated with the names 'Twain' and 'Clemens'. That x and y are *linked* to v and w represents that the embedded occurrences of the names 'Twain' and 'Clemens' in Ralph's ascriptions refer to *Twain*, and because reference markers x and y are linked to already present reference markers u and v , Ralph's utterances satisfy the *familiarity constraint*.²⁶

Ralph's utterances are true just in case the DRS constructed by his utterances accurately models reality, i.e. just in case *there are* real individuals a , b , c and attitudinal state d , where a plays the role of reference marker u , b plays the role of reference marker v , c plays the role of reference marker w , and d plays the role of reference marker p .²⁷

Slightly more precisely, Ralph's utterances are true if and only if

$\exists a \exists b \exists c \exists d$ [is named 'John'(a) & is named 'Twain'(b) & is named 'Clemens'(c) & was a great author(b) & wrote *Huck Finn*(b) & b is the same person as c & believes(a , d) & (d is true iff $\exists e \exists f$ [e is not the same person as f & was a great author(e) & was just an eccentric who never wrote anything(f)])]

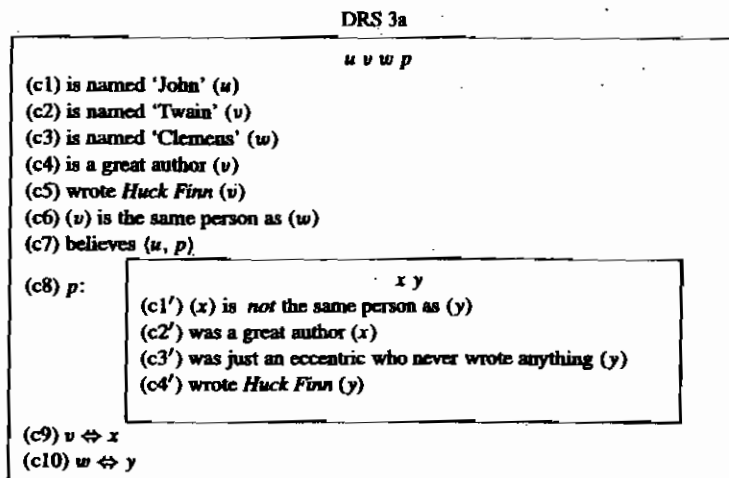
Now let us consider what happens to the context when, in response to Ralph's utterances, Mary utters sentences relevantly similar to ascriptions (1) and (2). Let us first suppose that, in response to Ralph's utterances, Mary utters an ascription relevantly similar to (2):

He believes that Clemens wrote *Huck Finn*.

Mary's utterance of this sentence evaluated relative to DRS 2 brings about DRS 3a. Mary's utterance amends the subDRS that represents Ralph and Mary's mutual beliefs concerning John's belief state. Thus the subDRS, which is a condition on p , is amended to incorporate the information expressed by Mary's

²⁶ Clearly much more needs to be said concerning linking conditions. In Fauconnier (1994) the idea is developed in much more detail, though in Fauconnier linking is accomplished via "connectors" that map between "triggers" and "targets." The same sorts of issues that arise in attempting to state what exactly it is for reference markers to be linked arise in possible world semantics under the rubric of "transworld identity conditions for objects."

²⁷ Kamp and Asher define truth for DRSs relative to models in terms of assignment functions, where an assignment function is a (perhaps partial) function from the discourse referents in a DRS to individuals in the universe U of the model. Thus, roughly, a DRS K is true in a model $\langle U, I \rangle$ iff there is some assignment function f such that f is a *proper embedding* for K in $\langle U, I \rangle$. And f is a proper embedding for K in $\langle U, I \rangle$ iff, for every condition in K of the form $C(x)$, where $C()$ is a unary condition and x is a reference marker, $f(x) \in I(C)$. (A recursive clause is required for the case in which C itself is a DRS.)



utterance. More specifically, condition (c4') is added to the subDRS. Reference marker y , rather than x , is associated with Mary's embedded utterance of 'Clemens' because 'Clemens' is familiar to the discourse, and is associated with reference marker w , which is internally linked via (c10) to reference marker y . That Mary's utterance is so linked is required by the familiarity constraint. Mary's utterance of 'He believes that Clemens wrote *Huck Finn*', relative to DRS 2, is thus true if and only if

$\exists a \exists b \exists c \exists d$ [is named 'John'(a) & is named 'Twain'(b) & is named 'Clemens'(c) & was a great author(b) & wrote *Huck Finn*(b) & b is the same person as c & believes(a, d) & (d is true iff $\exists e \exists f$ [e is not the same person as f & was a great author(e) & was just an eccentric who never wrote anything(f) & wrote *Huck Finn*(f)])]²⁸

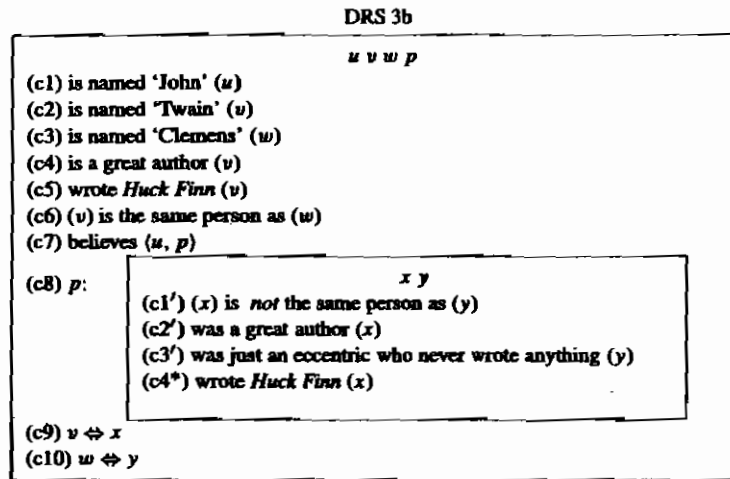
The central point here is that relative to the context created by Ralph's previous utterances, Mary's utterance is true only if John is in a belief state that accurately represents the world only if somebody who never wrote anything wrote *Huck Finn*. Thus, unless John is seriously confused, Mary's utterance of 'He believes that Clemens wrote *Huck Finn*' in the above discourse is false.

Now let us suppose that instead of uttering 'He believes that Clemens wrote *Huck Finn*', in response to Ralph's utterances, Mary were to utter an ascription that is relevantly similar to (1):

He believes that Twain wrote *Huck Finn*.

²⁸ For the sake of simplicity I ignore the contribution that (c9) and (c10) make to the truth conditions of Mary's utterance. These conditions require, roughly, that the possible persons that could make John's belief state true be counterparts to the actual individual Twain.

When an utterance of this sentence is evaluated relative to DRS 2 a significantly different DRS results (see DRS 3b).



Again, Mary's utterance amends the subDRS that represents Ralph and Mary's mutual belief concerning John's belief state. More specifically, condition (c4*) is added to the subDRS. This time reference marker x , rather than y , is associated with Mary's utterance of 'Twain' because, again, 'Twain' is familiar to the discourse, and is associated with reference marker v , which is in turn linked via (c9) to reference marker x . That Mary's utterance of 'Twain' be so linked is, again, required by the familiarity constraint. Mary's utterance of 'He believes that Twain wrote *Huck Finn*' relative to DRS 2, is thus true if and only if

$\exists a \exists b \exists c \exists d$ [is named 'John'(a) & is named 'Twain'(b) & is named 'Clemens'(c) & was a great author(b) & wrote *Huck Finn*(b) & b is the same person as c & believes(a, d) & (d is true iff $\exists e \exists f$ [e is not the same person as f & was a great author(e) & was just an eccentric who never wrote anything(f) & wrote *Huck Finn*(e)]]]

Thus in this simplistic DRT analysis 'He believes that Clemens wrote *Huck Finn*' uttered in response to Ralph's utterances does not have the same truth conditions as does an utterance of 'He believes that Twain wrote *Huck Finn*'. In particular, such an occurrence of 'He believes that Clemens wrote *Huck Finn*' is true only if John is in a belief state that accurately represents the world only if somebody who never wrote anything wrote *Huck Finn*, while the truth of such an utterance of 'He believes that Twain wrote *Huck Finn*' does not require that John have such an absurd belief.

The above sketched DRT analysis is inadequate because it is incomplete. It merely assumes the results of the discourse construction algorithm, and it pro-

vides only a rough picture of how the truth conditions of DRSs are determined. An adequate DRT analysis of attitude ascriptions would have to overcome these deficiencies. But despite its incompleteness the above sketched DRT analysis suffices to illustrate an important advantage that a discourse holistic analysis has, or would have, over analyses that utilize Frege's strategy: A discourse holistic analysis will not face the general difficulty with Frege's strategy. Because a discourse holistic analysis rejects sentential compositionality, it has no need to posit finely individuated extraordinary entities – "ways of thinking" of whatever sort – to account for the truth conditions of attitude ascriptions. The discourse holist, because he rejects sentential compositionality, need not claim that occurrences of (1) and (2) that have distinct truth conditions somehow invoke distinct extraordinary entities. Rather he can maintain that the difference in truth conditions is due to *overt* differences in the discourse environments of the occurrences. In particular, as is illustrated by the DRT analysis sketched above, the discourse holist can appeal to anaphora-like intersentential connections between occurrences of proper names to explain why the truth conditions of ascriptions such as (1) and (2) differ. Thus, as is proper, a discourse holistic analysis of ordinary attitude ascriptions need not presuppose the existence of any sort of "ways of thinking", whose existence may or may not be supported by future empirical evidence. Moreover, a discourse holistic analysis need not suppose that in using attitude ascriptions ordinary speakers are referring to, or otherwise "adverting to", such esoteric entities, and therefore it need not suppose that ordinary speakers have intentions they do not, and cannot, have.

One might object that the proposal of discourse holism confuses *semantic* and *pragmatic* considerations. It might be objected that any effect that the surrounding discourse environment has on what is communicated by an occurrence of an attitude ascription can be only a *pragmatic*, and not a *semantic*, matter. According to this objection the domain of semantics is limited to properties and factors *internal* to a sentence occurrence that are relevant to determining *what is said*.²⁹ And thus if there is information communicated by an occurrence that is communicated only because of features of the discourse environment *external* to the sentence, then this information is not within the domain of semantics, but is instead "merely pragmatic." Therefore, if our judgment that occurrences of (1) and (2) can differ in truth value is due in part to features of the discourse environment and is not determined by only (i) the syntactic structure of the occurrences, and (ii) the semantic values of the words and other semantically relevant syntactic elements in the occur-

²⁹ Some, such as Bach (1987), define semantics so that "the semantics of an expression give the information that a competent speaker can glean from it independently of any context of utterance" (p. 5). On this construal, reference and truth are not within the domain of semantics. Another common usage, probably stemming from Grice (1975), includes within semantics everything relevant to determining *what is said* by an occurrence, as opposed to what is (merely) *implicated* by the occurrence. This latter usage would include reference and truth in the domain of semantics. The imagined objector uses 'semantics' in this latter, Grice inspired, way.

rences, then occurrences of (1) and (2) cannot literally *say* different things – they cannot have distinct truth conditions. And judgments to the contrary are based on a confusion of *semantic* and merely *pragmatic* matters. And so, the objection concludes, the discourse holistic proposal simply leads us back to *Salmon's strategy*, briefly mentioned at the outset.

It should now be clear what is wrong with this sort of objection. The objection not only commits us to an analysis of attitude ascriptions that disregards our semantic judgments and is thus incompatible with the central task of a semantic theory, but it *presupposes* sentential compositionality and thereby begs the question against discourse holism. On the discourse holistic proposal sketched above *what is said* by an occurrence – its truth conditions – need *not* be determined by (i) its syntactic structure, and (ii) the semantic values assigned to its semantically relevant syntactic elements. There can be no significant objection to reserving the term 'semantic' for properties determined by such *intrinsic* features of an occurrence. But if that is how the term is to be used, then all the *semantic* properties of an occurrence need not determine *what it says* – need not determine the truth conditions of the occurrence. To merely assert otherwise is to beg the question against the discourse holistic proposal.

References

- Asher, N. (1986), *Belief in discourse representation theory*, *Journal of Philosophical Logic* 15, 137–189.
- Bach, K. (1987), *Thought and Reference*, Oxford University Press, Oxford.
- Bach, K. (1994), *Conversational implicature*, *Mind and Language* 9 (2).
- Bach, K. (1997), *Do belief reports report beliefs?*, *The Philosophical Quarterly* 78, 215–241.
- Carston, R. (1988), *Implicature, explicature, and truth-theoretic semantics*, *Mental Representations: The Interface Between Language and Reality*, R. Kempson, ed., Cambridge University Press, Cambridge.
- Clapp, L. (1995), *How to be direct and innocent: A criticism of crimmins and perry's theory of attitude ascriptions*, *Linguistics and Philosophy*.
- Crimmins, M. (1992), *Talk About Beliefs*, MIT Press, Cambridge.
- Crimmins, M. and Perry, J. (1989), *The Prince and the Phone Booth*, *Journal of Philosophy*.
- Culicover, P. and Jackendoff, R. (1995), 'Something else' for the binding theory, *Linguistic Inquiry* 26 (2), 249–275.
- Davidson, D. (1968), *On saying that*, *Synthese* 19, 130–136.
- Devitt, M. (1996), *Coming to Our Senses: A Naturalistic Program of Semantic Localism*, Cambridge University Press, Cambridge.
- Devitt, M. (1997), *Meanings and psychology: A response to Mark Richard*, *Nous* 31 (1), 115–131.
- Fauconnier, G. (1994), *Mental Spaces: Aspects of Meaning Construction in Natural Language*, Cambridge University Press, Cambridge.
- Fodor, J. (1978), *Propositional attitudes*, *The Monist* 61 (4). Also in J. Fodor (1981), *Representations: Philosophical Essays on the Foundations of Cognitive Science*, MIT Press, Cambridge.
- Forbes, G. (1990), *The indispensability of Sinn*, *Philosophical Review* 99, 535–564.
- Frege, G. (1893), *Über Sinn und Bedeutung*, *Zeitschrift für Philosophie und Philosophische Kritik* 100, 25–50.

- Frege, G. (1980), *Philosophical and Mathematical Correspondence* (Gabriel, Hermes, Kambartel, Theil and Veraart, eds, McGuinness, abr., and Kaal, tr.), University of Chicago Press, Chicago.
- Grice, P. (1975), *Logic and conversation*, *Speech Acts, Syntax and Semantics 3*, P. Cole and J. Morgan, eds, Academic Press, New York.
- Haim, I. and Kratzer, A. (1998), *Semantics in Generative Grammar*, Blackwell Publishers, Oxford.
- Kamp, H. (1984), *A theory of truth and semantic representation*. Truth, Representation and Information, J. Groenendijk, T. Jaassen and M. Stokhof, eds, Foris, Dordrecht.
- Kamp, H. (1988), *Comments on Stalnaker*, *Contents of Thought*, Grimm and Merrill, eds, University of Arizona Press, Tucson, 156–181.
- Kamp, H. (1990), *Prolegomena to a structural theory of belief and other attitudes*, *Propositional Attitudes*, A. Anderson and J. Owens, eds, CSLI Lecture Notes 20, pp. 27–90.
- 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.
- Kripke, S. (1979), *A puzzle about belief*, *Meaning and Use*, A. Margalit, ed., Reidel, Dordrecht.
- 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*, *Contemporary Perspectives in the Philosophy of Language II*, P. French, T. Uehling and H. Wettstein, eds, *Midwest Studies in Philosophy 14*, University of Notre Dame Press, Notre Dame, 338–356.
- Recanati, F. (1989), *The pragmatics of what is said*, *Mind and Language 4*.
- Richard, M. (1990), *Propositional Attitudes: An Essay on Thoughts and How We Ascribe Them*, Cambridge University Press, Cambridge.
- Richard, M. (1997), *What does commonsense psychology tell us about meaning?*, *Nous 31* (1), 87–114.
- Salmon, N. (1986), *Frege's Puzzle*, MIT Press, Cambridge.
- Saul, J. (1993), *Still an attitude problem*, *Linguistics and Philosophy 16*, 423–435.
- Saul, J. (1997), *Substitution and simple sentences*, *Analysis 57* (2), 102–108.
- Schiffer, S. (1992), *Belief ascription*, *Journal of Philosophy 89*, 499–521.
- Sperber, D. and Wilson, D. (1986), *Relevance: Communication and Cognition*, Blackwell, Cambridge.
- Stainton, R. (1994), *Using non-sentences: An application of relevance theory*, *Pragmatics and Cognition 22*, 269–284.
- Stalnaker, R. (1978), *Assertion*, *Syntax and Semantics, Vol. 9: Pragmatics*, Cole, ed., Academic Press, New York.
- Travis, C. (1985), *On what is strictly speaking true*, *Canadian Journal of Philosophy 15*, 187–229.