

Notes on Propositional Attitudes and Hidden Indexical Theories

Jonathan Cohen*

Hidden indexical theories (HITs) are among the most popular proposals of the semantics of propositional attitudes. In this essay I'll begin by discussing HITs against the background of Fregean and Russellian theories. Next I'll consider some of the benefits of HITs, and then compare HITs against some other well-known accounts. Finally, I'll examine some of the most important outstanding challenges to the viability of HITs, and indicate how I think these might be met.

1 Russell, Frege, and HITs

A well-worn problem in the philosophy of language involves propositional attitude ascriptions such as (1) and (2).

- (1) Max believes that Mark Twain is an aardvark.
- (2) Max believes that Samuel Clemens is an aardvark.

The difficulty is that it seems impossible to accommodate the intuitive possibility that (1) and (2) can differ in truth value while respecting the following three principles:¹

Direct Reference: Singular terms contribute their referents to propositions expressed by sentences in which they occur.

Semantic Innocence: Embedding a term in a that-clause does not change its semantic value.

Semantic Compositionality: The semantic value of a semantically composite expression is a function of its structure and the semantic values of its constituents.

To see the difficulty posed by what has come to be called Frege's puzzle, first notice that the difference between (1) and (2) is just a difference in the singular terms inside their that-clauses; but, according to Direct Reference,

*Department of Philosophy, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0119, joncohen@aardvark.ucsd.edu

¹This presentation of the problem comes from [Bach, 1997].

this difference can't make for a difference in the propositions expressed by the that-clauses, because the two singular terms under consideration have the same referent. But to preserve Semantic Innocence, we must hold that the semantic values of the embedded expressions are just the semantic values they would have if they were unembedded. And finally, Semantic Compositionality requires that the semantic values of the whole expression (1) and the whole expression (2) are composed from their structure (which they share), and the semantic values of their parts (which they also share). This means that the semantic values of (1) and (2) should be identical, contrary to the intuition that they needn't be. It would seem that something must give.

Russell is one of the fathers of the now-popular thesis of Direct Reference, and did much to make plausible the conception of propositions on which this puzzle is based. It is no surprise, then, that Frege's puzzle has attracted much attention from contemporary neo-Russellians (e.g., [Salmon, 1986], [Soames, 1987]). To avoid giving up any of the three principles, neo-Russellians have solved the puzzle by rejecting the intuition that (1) and (2) can differ in truth value. However, this Russellian line has seemed implausible to many theorists (see §3.1), and has convinced some to embrace a Fregean solution.

For Frege, the puzzle gave good reason for thinking that, inside the context of epistemic verbs such as 'believes', expressions refer to their customary senses rather than their customary references. On this line, 'Mark Twain' in (1) can have a different semantic value (make a different propositional contribution) from 'Samuel Clemens' in (2), and therefore we can explain the potential difference in truth value between (1) and (2) by claiming that they express different propositions. However, this Fregean solution comes at a price. First, although it accommodates Semantic Compositionality, it violates Direct Reference. Moreover, in holding that expressions inside epistemic contexts (but not outside these contexts) can refer to their customary senses rather than their customary referents, the Fregean solution runs afoul of Semantic Innocence.² A further cost of this violation is the trouble the Fregean has in accounting for

(1') Max believes that Mark Twain is an aardvark, but he is not.

If the Fregean is correct, then the name 'Mark Twain' in (1') refers to a sense, rather than an individual. But presumably, the pronoun 'he' in (1') is anaphoric on the name 'Mark Twain', so the Fregean must insist that the pronoun refers to a sense as well. This seems unpalatable.

²There are further problems for Frege's proposal. For example, since senses vary intersubjectively, it's not obvious whether the sense named in a report such as (1) should be the sense associated with 'Mark Twain' by Max, or that associated with the name by the person making the ascription. As argued in ([Richard, 1990], 66–67), neither choice is acceptable: the first alternative would prevent us from ascribing beliefs to individuals who use different names for objects (e.g., presumably Hamurabi associates no sense at all with the name 'Hesperus', but it makes sense to ascribe beliefs to him by using that name), while the second would permit us to make true ascriptions only by sharing a sense with the believers to whom we ascribe beliefs. Richard also argues, convincingly, that the Fregean will have no acceptable answer to the question which sense is expressed by 'Anne is married' inside multiple embeddings such as 'Jane believes that Barbara believes that Anne is married' (69ff). I cannot examine the details of these criticisms here.

Thus, both Russellian and Fregean theories face serious difficulties. HITs, which were introduced by [Schiffer, 1977], and have been advocated recently in various forms by [Crimmins and Perry, 1989], [Richard, 1990], [Crimmins, 1992], [Schiffer, 1992], and [Recanati, 1993], among others, have tried to combine the best aspects of both Fregean and Russellian theories while eschewing their defects.³ The heart of the HIT solution to Frege’s puzzle is the claim that, while (1) and (2) agree in what they say explicitly about *what* Max believes, they differ in what they say implicitly about *how* Max believes what he believes: (1) and (2) represent Max as believing the same content (that a certain individual has a certain property), but they represent him as believing it under different modes of presentation, to which (1) and (2) make implicit reference.

To see what these claims come to, consider the following canonical version of a HIT:⁴

(S) ‘Ralph believes that Fido is a dog’ is true iff $(\exists m)(\Phi^*m \ \& \ B(\text{Ralph}, \langle \text{Fido}, \text{doghood} \rangle, m))$ ([Schiffer, 1992], 503).

On this account, Φ^* is a contextually determined type of mode of presentation, i.e., “that property that a propositional mode of presentation has when and only when it requires thinking of Fido as being the dog who appears in the morning and requires thinking of doghood as a property shared by such-and-such-similar-looking creatures” (508). The mode of presentation type introduced by ‘believes’ on this analysis is both phonologically null (whence “hidden”) and context-dependent (whence “indexical”).⁵ The claim, then, is that, in the context of utterance, the believer thinks of the properties and individuals represented in the singular proposition under a particular mode of presentation and that this mode of presentation is an unarticulated constituent of the belief relation. Applying this to our case, we can see that (1) will be true just in case

$(\exists m)(\Phi_1^*m \ \& \ B(\text{Max}, \langle \text{Mark Twain}, \text{aardvarkhood} \rangle, m))$

is true, while (2) will be true just in case

$(\exists m)(\Phi_2^*m \ \& \ B(\text{Max}, \langle \text{Mark Twain}, \text{aardvarkhood} \rangle, m))$.

Significantly, this account leaves available the possibility that (1) and (2) can differ in truth value, so long as the mode of presentation Max associates with the

³Schiffer occupies an odd stance with respect to HITs. He is generally credited with having proposed the account, and he has done much to argue that HITs represent the best semantics for propositional attitudes relative to the assumption that Semantic Compositionality is true. However, he does not believe this assumption (see especially [Schiffer, 1987], ch. 7), so has felt free to argue against these theories. Indeed, his are some of the more trenchant objections against HITs (we’ll consider some of them below).

⁴Of course, there are many different versions of HIT in the literature. I’ve tried to pick a relatively generic (i.e., uncommitted) version for the purposes of my exposition. I’ll comment on some of the differences between HITs in later sections.

⁵As many writers have noticed, the modes of presentation HITs claim are introduced by ‘believes’ lack the kind of stable, articulable meaning rule we find for indexical expressions like ‘I’ and ‘now’; this has led some to suggest that the hidden indexicals of HITs are in fact more similar to pure demonstratives than true indexicals.

embedded name in (1) is distinct from that Max associates with the embedded name in (2).

It should be clear how this HIT borrows elements from both the Fregean and Russellian proposals we have considered: it follows the Russellian in counting individuals and relations (rather than Fregean senses) as the constituents of propositions expressed by attitude sentences, but it follows the Fregean in using differences in modes of presentation to explain failures of substitutivity of coreferential names inside attitudinal contexts. Given its basis in these theories, we may ask whether a HIT will fare any better than its progenitors with respect to the desiderata considered above. First, as we have seen, a HIT (unlike a Russellian theory) can accommodate the data that (1) and (2) can differ in truth value. However, our HIT (unlike a Fregean theory) is compatible with Direct Reference; indeed, it is left open by this theory (and in fact all of the leading proponents of HITs believe) that the relationship between the name ‘Mark Twain’ and the individuals appearing in the propositions that are the objects of Max’s beliefs is causal, and operates independently of whatever beliefs Max may have about what the name refers to, as required by the picture advocated in [Kripke, 1980]. In addition, our HIT (unlike a Fregean theory) is Semantically Innocent, since it claims that the embedded sentence in (1) expresses in its embedded context just what it expresses in non-embedded contexts, viz., the proposition that Mark Twain is an aardvark.

2 Further Advantages of HITs

As we have seen, HITs can respect Frege’s data and cohere with the demands of Direct Reference and Semantic Innocence.⁶ But HITs have other advantages as well.

2.1 That-Clauses Are Referential Singular Terms

For example, as Schiffer has pointed out, HITs sustain and explain the intuition that that-clauses are referential singular terms. This intuition can be supported by two considerations. First, on its face, it seems plausible that a sentence such as (1) tells us one of the things that Max believes — viz., that Mark Twain is an aardvark — as it would if that-clauses were referential singular terms. And second, this intuition, if true, would explain the apparent validity of the following arguments (adapted from ([Schiffer, 1992], 505)):

(a1) Max believes that Mark Twain is an aardvark, and so does Mary.

(a2) Therefore, there is something that Max and Mary believe.

(b1) Max believes everything that Mary says.

⁶The question whether HITs respect compositionality is a bit more vexed, and I shall return to it in §4.1.

- (b2) Mary says that Mark Twain is an aardvark.
- (b3) Therefore, Max believes that Mark Twain is an aardvark.
- (c1) Max believes that Mark Twain is an aardvark.
- (c2) That Mark Twain is an aardvark is impossible (/ridiculous/funny/true/etc.).
- (c3) Therefore, Max believes something that is impossible (/ridiculous/funny/true/etc.).⁷

Schiffer claims that these inferences cannot be explained successfully without taking that-clauses to be referential singular terms, and therefore that it is incumbent on any semantics for propositional attitudes to accommodate this intuition.⁸ Whether or not this strong claim is correct, Schiffer is certainly right to point out that HIT accounts vindicate his intuition, and that accepting this intuition provides a very natural explanation of the validity of the inferences he considers.

2.2 Solving Puzzle Cases

A further motivation for HITs is that they resolve many of the puzzle cases that plague other theories. To take a prominent example, HITs can easily resolve Kripke’s puzzle about belief ([Kripke, 1979]). For a proponent of a HIT, ‘Pierre believes that London is pretty’ is true because

$$(\exists m)(\Phi_1^*m \ \& \ B(\text{Pierre}, \langle \text{London}, \text{pulchritude} \rangle, m))$$

is true. Nonetheless, ‘Pierre believes that London is not pretty’ is true because

$$(\exists m)(\Phi_2^*m \ \& \ B(\text{Pierre}, \langle \text{London}, \text{non-pulchritude} \rangle, m))$$

is true. But making these two ascriptions does not commit us to ascribing inconsistency to poor Pierre, as Kripke worries it will, because the following is *false*:

⁷To see the flavor of how the intuition facilitates the explanations of these inferences, notice that this intuition would allow us to understand the inference from (a1) to (a2) in terms of the (valid) inference from ‘B (Max, *t*) & B (Mary, *t*)’ to ‘(∃*p*) B (Max, *p*) & B (Mary, *p*)’. Of course, I am suppressing hidden indexicals in these representations because the intuition doesn’t require them (HITs commit to more than just the truth of the intuition). But notice that, if a HIT is right, hidden indexicals can derail the inference if they fail to match up correctly: the inference from ‘(∃*m*)(Φ₁^{*}*m* & B(Max, *t*, *m*)) & (∃*m*)(Φ₂^{*}*m* & B(Mary, *t*, *m*))’ to ‘(∃*m*)(∃*p*)(Φ^{*}*m* & B (Max, *p*, *m*) & B(Mary, *p*, *m*))’ is *invalid*. This result seems to vindicate intuitive predictions as well.

⁸Part of Schiffer’s reason for claiming that these inferences can only be explained by respecting this intuition is his confidence in his objections from ([Schiffer, 1987], ch. 5) against a Davidsonian-style paratactic analysis (one which extends the account of saying-that in [Davidson, 1968] to an account of belief and other propositional attitudes), which he regards as the most plausible alternative explanation. However, I don’t think Schiffer has said enough here to rule out other approaches; in particular, the measurement theoretic semantics for propositional attitudes defended in [Matthews, 1994] and [Davidson, 1989] provides a basis for explaining these inferences without accepting the truth of Schiffer’s intuition.

$(\exists m)(\Phi^*m \ \& \ B(\text{Pierre}, (\text{London}, \text{pulchritude}), m) \ \& \ B(\text{Pierre}, (\text{London}, \text{non-pulchritude}), m))$.

On the other hand, a HIT advocate can explain how the suspicion of inconsistency arises: since Kripke tells the story of Pierre in such a way that his two histories of acquaintance with the city are kept separate, and since the hidden indexicals used in the belief ascriptions are hidden (phonologically null), one might erroneously fail to distinguish between the distinct modes of presentation mediating Pierre’s beliefs about London. But this erroneous failure would be erroneous, and a failure.⁹

3 Some Competing Theories

3.1 Neo-Russellian Theories

We have already seen how Fregean theories of propositional attitudes can only preserve the intuition that (1) and (2) can differ in truth value by violating Semantic Innocence (they also violate Direct Reference).

Neo-Russellians, in contrast, preserve all the semantic principles given above, but reject the intuition that (1) and (2) can differ in truth value. On its face, the neo-Russellian strategy of resolving an apparent counterexample to the theory by rejecting the intransigent data seems suspect. But neo-Russellians defend their methodology by throwing doubt on how the data should be understood. A neo-Russellian claims that her account only addresses the (strict, literal) truth-conditional content of sentences like (1) and (2), and that the intuitions adduced against her view in fact concern the pragmatic implications of these sentences rather than their truth-conditional contents. Thus, a neo-Russellian diagnoses the anti-Russellian intuitions concerning (1) and (2) as a confusion between (i) common pragmatic implications about *how* we believe certain truth-conditional contents and (ii) the truth-conditional contents themselves. If this is right, and if the relevant anti-Russellian intuitions really can be understood as pragmatic rather than semantic, then the Russellian account can be saved.

However, there are reasons for doubting that we can understand anti-Russellian intuitions purely pragmatically, as the neo-Russellian insists we should. The strongest worry we can raise in this regard is that, unlike the situation with the non-truth-functional connotations of ‘and’ (for example), it is extremely difficult to convince ordinary speakers that their intuitions about (1) and (2) are semantically irrelevant. Richard puts the point this way:

...other than using bribery, threats, hypnosis, or the like, there is simply nothing you can do to get most people to say that Jones believes that Tully was an orator, once they know that Jones sincerely denies “Tully was an orator”, understands it, and acts on his denial

⁹By roughly similar moves, HITs can defuse most of the other common puzzles regarding propositional attitudes, e.g., Richard’s steamroller ([Richard, 1983]). I shall refrain from discussing these other cases for the sake of brevity.

in ways appropriate thereto. In particular, pointing out that Jones can express something he believes with “Cicero was an orator” seems simply irrelevant to most people ([Richard, 1990], 125).

A further problem for the neo-Russellian’s pragmatic defense, raised in ([Recanati, 1993], 342–344), is that it commits her to holding that negative reports such as (1-) are never literally true.

(1-) Max does not believe that Mark Twain is an aardvark.

This is because, for the neo-Russellian, there will be a way of taking the proposition that Mark Twain is an aardvark — viz., the one pragmatically conveyed by an utterance of (2) — under which Max *does* believe this proposition, contrary to (1-). Finally, ([Recanati, 1993], ch. 17) brings out several disanalogies between the kinds of pragmatic implicatures invoked in a neo-Russellian account and the characteristic features of pragmatic implicatures described in [Grice, 1975] (and subsequent literature). All these difficulties cast doubt on the plausibility of the neo-Russellian program.

3.2 Sententialist Theories

Another family of theories of the semantics of propositional attitude statements includes various forms of sententialism. On a particularly simple version of sententialism, popular among logical empiricists (e.g., [Carnap, 1956]), an embedded sentence refers to itself. Thus, on this theory, (1) expresses a relation between Max and the sentence ‘Mark Twain is an aardvark’. However, this simple form of the theory is subject to worries about learnability (see the criticisms of Carnap, Scheffler, and Quine in [Davidson, 1965]). This, together with complications concerning, e.g., pronouns and binding relations, structural ambiguity in both sentences and individual words, etc. (cf. [Larson and Segal, 1995], 419–422, for discussion) has led to some to advocate a modified sententialism according to which (1) expresses a relation between Max and some (perhaps quite elaborate) linguistic structure encoding an English sentence (this might include a syntactic representation, a phonological representation, a semantic representation, or any combination of the above).

However, even this elaborated form of sententialism can’t be right as it stands, because it demands too tight a relationship between believers and sentences of English: thus, on the version of sententialism contemplated so far, sentences such as (1) could never be used to ascribe beliefs to non-English speakers. Consequently, fans of sententialism have followed the analysis of saying-that in [Davidson, 1968] in building a notion of similarity or “samesaying” directly into their accounts. On this modified form of sententialism, then, (1) reports that Max stands in some relation to a sentence (or linguistic structure encoding a sentence) similar to the sentence ‘Mark Twain is an aardvark’. Advocates of these theories point out that the notion of similarity appearing in their semantics for propositional attitudes is deliberately left vague with respect to the features

of particular contexts. Through this means, sententialists hope to incorporate context-relativity into their semantics.¹⁰

I won't criticize these theories at length (see [Schiffer, 1987], ch. 5 for discussion), but I would like to register three brief worries about various forms of sententialism. First, although sententialists consider the vagueness of the samesaying relation a virtue for their theories in that it makes available explanations of how different reports can appropriately serve to report a given belief in different contexts, the explanations offered by sententialists are, I think, a bit thin on the ground. To be told that the objects of Max's beliefs stand in the samesaying relation to the sentence 'Mark Twain is an aardvark' in some but not all contexts remains rather empty until we are told something more about the samesaying relation.¹¹ In particular, it would be nice to know what are the different contextual parameters to which the relation is sensitive (or what are contexts, for that matter!); without this information, it's hard to begin to understand (hence to evaluate) sententialist views.¹²

A connected concern for sententialism is this. Appealing to sentential material to distinguish the truth conditions of (1) and (2) is an attractive and motivated way of resolving the puzzle Frege found. However, in many cases, these appeals to sentential material will result in excessively fine-grained individuation: in some cases, it seems that (1) and (2) might *agree* in truth value.¹³ The usual sententialist strategy here is to claim that, in such cases, the complements of (1) and (2) both stand in the contextually-sensitive samesaying relation to the object of Max's belief. Of course, sententialists have no semantic account of samesaying: "same-saying... is fundamentally a matter of *usage* and not content.... the correct account of these phenomena falls outside the

¹⁰One recently popular theory of roughly this form is the Interpreted Logical Form (ILF) proposal advocated in [Larson and Ludlow, 1993] and [Larson and Segal, 1995]. This theory is a hybrid between sententialist and Russellian treatments, however, since the objects to which it claims believers are related are complexes containing both linguistic material and Russellian-style propositional constituents. One problem that arises in force for the ILF proposal, but not for all other forms of sententialism, involves its (metalinguistic) appeal to a belief relation holding between persons and ILFs. Clearly, this metalinguistic notion cannot be identical with any pre-theoretical notion of belief, since naive intuition does not recognize ILFs or relations to them. Therefore, the proponent of ILFs owes us an explanation of her theoretical term 'believes'. (Of course, the same problem might be pressed against HITs, which have us believing propositions under modes of presentation; indeed, it is useful to understand the logical form problem discussed in §4.4 as a dispute over whether the metalinguistic notion of belief appealed to in a HIT coincides with the ordinary notion of belief.)

¹¹One reflection of the vacuousness of appeals to the samesaying relation is that different sententialists disagree about what serve as its relata, but have had no direct way of arguing against each other on this point.

¹²One might lodge the same complaint against HITs, whose modes of presentation are claimed to be contextually-sensitive as well. This complaint seems fair enough, but several HIT advocates have tried to explain how their theories could answer such questions (e.g., [Crimmins, 1992], ch. 2–3 on "normal" notions and ideas). The same cannot be said for sententialists.

¹³NB: In *some* cases, not *all* cases. Hence, the proposed solution in ([Larson and Ludlow, 1993], 322) and ([Larson and Segal, 1995], 442–445) — viz., that it's consistent with sententialism that distinct sentences (/sentence-encoding-structures) turn out to be logically equivalent — doesn't help.

domain of semantics proper and into pragmatics” ([Larson and Ludlow, 1993], 339). This move should remind us of the strategy employed by neo-Russellians: where the fine-grained semantics of sententialism conflicts with intuition, sententialists are claiming that those intuitions reflect pragmatic rather than semantic factors. But if the sententialist must ultimately explain away clashes between her theory and intuition in pragmatic terms, we may reasonably wonder what advantages sententialism has over the sparser neo-Russellian accounts considered in §3.1.

A final worry about sententialism is that, by holding that embedded expressions refer to sentences or linguistic structures, such theories often violate Semantic Innocence.¹⁴

3.3 Modal Theories

Modal theories of belief ([Lewis, 1979], [Lewis, 1986], [Stalnaker, 1981], [Stalnaker, 1984]) claim that (1) is true just in case Mark Twain is an aardvark in all the worlds doxastically accessible to Max (or that all Max’s doxastic alternatives live in worlds where Mark Twain is an aardvark). These views are radically different from the other views we have considered insofar as they fail to attribute structure to beliefs.

The classic problems for modal theories of belief involve overgeneration:¹⁵ since the class of worlds where Mark Twain is an aardvark is identical to the class of worlds where Samuel Clemens is an aardvark, it is impossible to believe (1) without believing (2). Similarly, if p entails q , then every world where p holds is a world where q holds, so modal theorists must admit that our beliefs are closed under entailment. This conclusion is implausible enough, but it leads to particular embarrassment in the case of contradictory beliefs: since anything follows from a contradiction, someone whose beliefs are inconsistent (probably all of us) will believe literally everything. Equally unpalatably, since mathematical truths are true in all worlds, we must all be said to believe all the *recherche* truths of mathematics. In addition to these counterintuitive consequences, a modal theory takes embedded sentences to have classes of worlds as their semantic values, and thereby sacrifices Semantic Innocence. For these reasons, it would seem that accepting a modal theory of belief is inadvisable.¹⁶

¹⁴This complaint will not be telling against a paratactic view though, since on such a view, the only metalinguistic expression in (1) is the word ‘that’, which is regarded by such theories as a demonstrative rather than a complementizer. Surely it is possible for a demonstrative to refer to a linguistic entity without violating Semantic Innocence.

¹⁵Modal theories also give rise to undergeneration problems, but I won’t discuss these here.

¹⁶There are various maneuvers advocates of modal theories have relied on in replying to the worries stated above, but they are complicated, and I can’t discuss them here. A nice presentation of the different strategies and the problems faced by each is given in [Robbins, 2003].

4 Worries about HITs

Now that we've considered problems facing other leading theories of propositional attitudes, I want to consider some of the most important objections against HITs, and indicate what I think are the most promising lines of response to these worries.

4.1 Hidden Indexicals and Compositionality

Although Schiffer defends HITs as the best semantics for propositional attitudes within a compositional semantics ([Schiffer, 1992], 519), several theorists have alleged that HITs do not respect Compositionality. For example, Bach complains that HITs violate Semantic Compositionality because, on these theories, the semantic value of (1) is composed from not only the semantic value of 'Mark Twain is an aardvark' and the semantic value of 'Max believes that' plus the syntactic arrangement of these two parts, but also a mode of presentation [Bach, 1997]. Now, the hidden indexical theorist claims that the mode of presentation to which she appeals is, in John Perry's phrase, an "unarticulated constituent" of the proposition expressed by (1). This can be understood in two ways.

For some hidden indexical theorists (e.g., [Recanati, 1993]), the presence of an unarticulated constituent in the proposition expressed by (1) is not a violation of Compositionality, since the unarticulated constituent is also part of the semantic value of what is expressed by the articulated constituent 'Mark Twain'. Unfortunately, I don't think this suggestion is promising. For although it would dissolve concerns about Compositionality, it would seem either to violate Direct Reference (by holding that expressions always have unarticulated constituents as part of their semantic values), or else give up Semantic Innocence (by holding that embedded sentences but not unembedded sentences refer to "quasi-singular propositions" containing both modes of presentation and Russellian constituents).

However, other hidden indexical theorists (e.g., [Crimmins and Perry, 1989]) have followed a different strategy. They have admitted that their unarticulated constituents "are not held to be parts of the content of any unit smaller than the entire ascription" ([Crimmins, 1995], 201), and therefore that their theory violates the principle of Semantic Compositionality as stated above. But they have held that this is unobjectionable because this principle can be shown to be false on other grounds. That said, these theorists typically claim that they can accommodate some (weaker) form of compositionality which doesn't suffer from the difficulties they find in the strong principle. For example, Crimmins accepts that "since the contents of complex expressions presumably are determined *somehow* and depend systematically on *some* features of the uses of the [constituent] expressions, . . . some principle of compositionality must be correct" ([Crimmins, 1992], 9). But he denies "the principle of articulated (universal) compositionality: the content, in a statement of any complex expression depends only on the contents, in the statement, of its component expressions"

([Crimmins, 1992], 10).

If Crimmins's defense is to save HITs from worries about compositionality, it must turn out that (i) whatever compositionality principle can be successfully motivated by arguments concerning the learnability, productivity, and systematicity of natural languages is weaker than the principle of Semantic Compositionality, and (ii) this weaker compositionality principle is compatible with the presence of unarticulated constituents as demanded by HITs. Of course, it remains to be seen whether these claims can be sustained.

4.2 The Candidate Problem

In [Schiffer, 1992], Schiffer raises the so-called candidate problem against HITs. Schiffer points out that the modes of presentation to which HITs appeal are defined functionally (as whatever mediates our cognitive relations to propositions in the right way). However, this definition leaves open the question what sorts of things play that functionally described role. As Schiffer notices, answers to this query have not been wanting: modes of presentation have been variously held to be individual concepts, general properties, percept tokens, stereotypes, prototypes, characters, public language expression types, Mentalese expression types, functional roles, causal chains between Mentalese names and individuals, and so on. Unfortunately, claims Schiffer, almost all of these proposals are subject to decisive objections. Therefore, he concludes, until the proponent of HITs can explain what a mode of presentation is without falling prey to the objections crippling existing proposals, we should refuse to endorse any HIT.

However, I don't find the candidate problem particularly pressing. First, it is unclear whether Schiffer is justified in demanding that HIT proponents explicate all the notions deployed in their theories. After all, explanations have to stop somewhere, and Schiffer has given no reason for thinking that the notion of a mode of presentation is inherently contradictory or otherwise ineluctably flawed. Second, even if we accept Schiffer's demand for explanation, we may reasonably wonder (if only because of the large number of theories mentioned) whether all of his objections against the whole range of proposals are apodictic. Moreover, even if all of Schiffer's arguments were devastating, we've been given no reason for believing future accounts will be likewise unsuccessful. For these reasons, I think a HIT advocate may hold out hope that her account can surmount the challenges posed by the candidate problem.

4.3 The Meaning-Intention Problem

The meaning-intention problem is that "one may reasonably doubt that belief ascribers mean what the hidden indexical theory requires them to mean when they ascribe beliefs" ([Schiffer, 1992], 518). This is because, if a HIT is right, (1) tells us that there is a mode of presentation of the right type such that Max believes the reported proposition under it. But surely most reporters of propositional attitudes are not aware of making reference to mode of proposition types. After presenting this objection, Schiffer recognizes what I take to be the

best reply to it, viz., an answer appealing to tacit intentions of the reporter. On this line of response, reporters have the intentions a HIT says they have (mean what the theory says they mean, refer to what the theory says they refer to) even though they lack conscious awareness of these intentions (meanings, referings).

But Schiffer is unhappy with this response for two reasons. First, it entails that ordinary belief ascribers typically have no conscious knowledge of what they are asserting and don't have the conscious thoughts they think they are having. Consequently, a hidden indexical theory "riddles the propositional attitude ascriptions of ordinary speakers with error; it also forces us to qualify our views about first-person authority in an important way" ([Schiffer, 1992], 515). Second, if a roughly Gricean conception of nonnatural meaning is right, then part of referring to a thing is intending one's audience to recognize that reference. Thus, if a hidden indexical theory is right and reporters do refer to a mode of presentation, then they intend their audience to recognize that mode of presentation. But this seems optimistic in the extreme:

If a proposition is believed under one mode of presentation, then it will typically be believed under many modes of presentation. Further, each of those mode of presentation will instantiate infinitely many *types* of modes of presentation, many of which will be equally salient in the communicative context. This makes it extremely implausible that of all the equally salient type of ways that [the believer] has of believing the proposition. . . [the reporter] should mean — and intend to be taken to mean — a proposition about one definite one of them ([Schiffer, 1992], 516).

I think these objections can be answered. First, the consequence that ordinary belief ascribers typically have no conscious knowledge of what they are asserting is not as heterodox as Schiffer makes it seem. Indeed, the history of technical semantics is replete with analyses of expressions departing quite radically from their surface forms. If any of these analyses correctly tells us what speakers say when they employ the relevant locutions, then the reasonable assumption that these analyses are not consciously known to ordinary speakers would entail that ordinary speakers lack conscious knowledge of what they are saying. If we followed Schiffer's lead in denying the plausibility of such analyses merely because they "riddle with error" the conscious knowledge of ordinary speakers, thereby leading us to "qualify our views about first-person authority," we would have to forego Davidsonian event analyses of various constructions, relational analyses of determiners, and on and on.¹⁷ These theories may, of

¹⁷Moreover, as Schiffer himself points out, positing tacit intentions is even more plausible with respect to syntactic knowledge. It is uncontroversial that speakers lack conscious knowledge of the syntactic representation of the expressions they utter. But if, as has been claimed, we nonetheless can be said to have knowledge of syntactic properties, relations, etc., then this by itself would "riddle with error" the conscious knowledge of ordinary speakers and lead us to "qualify our views about first-person authority." By and large, philosophers and linguists have accepted these consequences for the domain of syntax; it is puzzling then why they should be thought unacceptable for the domain of semantics.

course, be false; but presumably we cannot realize their falsity simply by noticing that their truth would require modification of our views about first-person authority.

Second, I think we should reject Schiffer’s claim that the reporter does not mean — and intend to be taken to mean — a proposition about just one of the equally salient types of ways the believer has of believing the proposition. Of course Schiffer is correct to point out that there are typically multiple candidates for the mode of presentation type, and that many of these can be equally salient. But it doesn’t follow that reporters can’t succeed in intending a token of just one of these, and in intending to be taken to intend a token of just one of these. Sometimes audiences fail to identify what was intended by the reporter, and sometimes these failures are irrelevant to the communicative purpose at hand since the audience makes do with an equally salient mode of presentation not intended by the reporter (e.g., you don’t know that that guy is Mark Twain, but Max and I do, and I believe that you do too; I report to you that the Max thinks that that guy is an aardvark, intending the Mark Twain mode of presentation, but you think I intended the Samuel Clemens mode of presentation; nonetheless, you make do perfectly well in the communicative context). But these situations provide no reason for thinking that there *can’t* be a fact of the matter about which of the multiple modes of presentation was intended. They simply show that such intentions are underdetermined by our reporting behavior in ordinary cases; and this is something we already knew.

Thus, it looks as if the objections Schiffer raises against the tacit intention solution to his meaning-intention problem are unsuccessful. If this is right, then we may make confident use of tacit intentions in answering the meaning-intention problem.

4.4 HITs and Logical Form

According to Schiffer, there are syntactic grounds for doubting that ‘believes’ has a hidden mode of presentation argument, but HITs can only succeed if it does. At best, claims Schiffer, a mode of presentation can act as an adjunctive modifier of a two place relation, but not a third argument of a three place relation.

To see his objection, consider

- (3) To whom did you wonder whether John gave the book?
- (4) Under what influence did you wonder whether Louise hit Ralph?
- (5) Under what mode of presentation did you wonder whether Max believes Mark Twain is an aardvark?

Schiffer proposes the following test to distinguish arguments from adjuncts in the above constructions:¹⁸

¹⁸Schiffer refers to (SC) as a “revised Chomskian criterion” [Schiffer, 1996], but [Ludlow, 1996] retorts that this criterion is unlike anything Chomsky has proposed.

(SC) When there is no ambiguity we have an argument, otherwise an adjunct.

So for example, claims Schiffer, the argument ‘to whom’ in (3) can only be understood as modifying something downstairs (the giving), while the adjunct ‘under what influence’ in (4) can be understood as modifying either something upstairs (the wondering) or downstairs (the hitting). Schiffer thinks the expression ‘under what mode of presentation’ gives rise to the ambiguity characteristic of adjuncts, since it can modify either the wondering or the believing in (5). Thus, he claims, the syntax classifies the modes of presentation deployed in HITs as adjuncts rather than arguments, but HITs must understand these modes as arguments rather than adjuncts. Consequently, the logical form demanded by HITs clashes with our best syntactic theories.

There are several ways of answering this argument. First of all, it’s not clear why Schiffer thinks semantic theories must surrender to syntactic demands when there is a conflict between the two. Surely the ultimate syntactic theory and the ultimate semantic theory should be compatible, but neither the syntactic nor the semantic issues in this area are settled, so there’s plenty of room for maneuvering.¹⁹ However, we may put this objection aside, since [Ludlow, 1996] suggests two other ways of answering Schiffer’s logical form problem: he complains that (i) (SC) is the wrong criterion for testing whether expressions are adjuncts or arguments, and that modes of presentation come out as arguments on more appropriate criteria, and (ii) HITs needn’t demand that modes of presentation are arguments, so none of this matters to the viability of HITs.

Why does Ludlow think (SC) is an inappropriate criterion? He suggests that the ambiguities to which (SC) is sensitive should be explained by appeal to the meanings of verbs, rather than to their argument structures. For example, the univocality of (3) can be explained by noticing that acts of giving involve recipients, while acts of wondering do not, hence that ‘to whom’ can only modify the act of giving in (3). But this explanation leaves it open whether ‘to whom’ functions as an adjunct or an argument.

On the other hand, Ludlow points to other tests for distinguishing arguments from adjuncts, and suggests that these favor an argument encoding of modes of presentation in attitude contexts. One such other criterion, advocated by [Bresnan, 1982] (among others), is that adjuncts can be iterated, but that arguments cannot. Thus, compare the following:

- (6) John buttered his toast in the restaurant, in the dark, on Tuesday.
- (7) *John buttered the toast with a knife, with a spoon, with a fork.
- (8) *Max believed that Mark Twain is an aardvark under mode of presentation m , under mode of presentation m' , under mode of presentation m'' .

On this criterion, (7) is unacceptable (except on a conjunctive, i.e., non-adjunctive reading); advocates of the iteration criterion have taken this as evidence for an implicit argument place for the instrument of the verb ‘butters’.

¹⁹Moreover, even if all the issues were settled, it’s possible that the ultimate syntax and the ultimate semantics might cross-classify some expressions.

Ludlow points out ([Ludlow, 1995], [Ludlow, 1996]) that, on this iteration test, (8) and (7) fall together, which suggests that modes of presentation are encoded as arguments of ‘believes’.²⁰

Ludlow’s second form of response to the logical form problem is more radical. Here he argues that, even if he’s wrong in his contention that ‘under mode of presentation m ’ is an argument of ‘believes’, at least some version of a HIT can live with this outcome. Thus, if we assume that adjuncts are modifiers of events, we can defend a modified HIT given by the following:

Adjunct-HIT ‘Max believes that Mark Twain is an aardvark’ is true iff $(\exists e)(\exists m)(\Phi^*(e, m) \ \& \ B(e, \text{Max}, \langle \text{Mark Twain}, \text{aardvarkhood} \rangle) \ \& \ \text{mode}(e, m))$.²¹

Assuming Ludlow’s strategy is successful, we may conclude that HIT proponents can sidestep the logical form problem even if they cannot resolve it.

5 Conclusion

Although HITs offer many theoretical advantages, there remain outstanding problems for the approach. However, as we have seen, all of the prominent competitors to HITs are subject to serious doubts as well, and there is reason for believing that HITs might be defended against some of the more serious threats to their survival. To quote Cicero, who has loomed so large in the literature under discussion, where there’s life, there’s hope.²²

References

[Bach, 1997] Bach, K. (1997). Do belief reports report beliefs? *Pacific Philosophical Quarterly*, 78(3):215–241.

²⁰Schiffer responds to this argument in [Schiffer, 1996] by claiming that some paradigmatic cases of adjunction fail the iteration test, and that therefore the iteration test is an unacceptable criterion. Schiffer’s case is:

(9) ?John hit Bill under the influence of crack, under the influence of rock music.

However, at least to my ear, (9) is fine: (9) might describe a situation in which John listens to a rock song that instructs him to take crack and then hit his friends, for example. (Cf. [Ludlow, 1996]).

²¹Schiffer rejects this theory, demanding that HITs must construe modes of presentation as arguments rather than adjuncts. This is because, he thinks, the adjunctive analysis represents ‘believes’ lexically as a two-place verb (ignoring the event place). Consequently, in a compositional semantics, the verb will be held to express a two-place relation; and this, he says, “with the rest of the machinery of the compositional semantics, *fixes* the logical form of any sentence that results from filling the verb’s argument places with singular terms as being of the form $V(a, b)$ ” ([Schiffer, 1996]). But [Ludlow, 1996] shows how a compositional semantics can allow full sentences containing ‘believes’ (and other verbs with sentential complements) to contain an adjunct with an argument place for a mode of presentation without requiring that ‘believes’ itself contains an argument place for a mode of presentation.

²²*Ad Atticum*, IX, 10.

- [Bresnan, 1982] Bresnan, J., editor (1982). *The Mental Representation of Grammatical Relations*. MIT Press, Cambridge, Massachusetts.
- [Carnap, 1956] Carnap, R. (1956). *Meaning and Necessity*. University of Chicago Press, Chicago.
- [Crimmins, 1992] Crimmins, M. (1992). *Talk About Beliefs*. MIT Press, Cambridge, Massachusetts.
- [Crimmins, 1995] Crimmins, M. (1995). Quasi-singular propositions: The semantics of belief reports. *The Aristotelian Society Supplement*, 69:195–209.
- [Crimmins and Perry, 1989] Crimmins, M. and Perry, J. (1989). The prince and the phone booth. *The Journal of Philosophy*, 86:685–711.
- [Davidson, 1965] Davidson, D. (1965). Theories of meaning and learnable languages. In Bar-Hillel, Y., editor, *Logic, Methodology, and Philosophy of Science*, pages 383–394. North-Holland Publishing Company, Amsterdam. Reprinted in [Davidson, 1984], 3–15.
- [Davidson, 1968] Davidson, D. (1968). On saying that. *Synthese*, 19:130–146. Reprinted in [Davidson, 1984].
- [Davidson, 1984] Davidson, D. (1984). *Inquiries into Truth and Interpretation*. Clarendon Press, Oxford.
- [Davidson, 1989] Davidson, D. (1989). What is present to the mind. In Brandl, J. and Gombocz, W., editors, *The Mind of Donald Davidson*, pages 3–18. Rodopi, Amsterdam.
- [Grice, 1975] Grice, H. P. (1975). Logic and conversation. In Cole, P. and Morgan, J. L., editors, *Syntax and Semantics*, volume 3, pages 41–58. Academic Press, New York. Reprinted in [Martinich, 1990].
- [Kripke, 1979] Kripke, S. (1979). A puzzle about belief. In Margalit, A., editor, *Meaning and Use*, pages 239–283. Reidel, Dordrecht. Reprinted in [Salmon and Soames, 1988].
- [Kripke, 1980] Kripke, S. (1980). *Naming and Necessity*. Harvard University Press, Cambridge, Massachusetts.
- [Larson and Segal, 1995] Larson, R. and Segal, G. (1995). *Knowledge of Meaning: An Introduction to Semantic Theory*. MIT Press, Cambridge, Massachusetts.
- [Larson and Ludlow, 1993] Larson, R. K. and Ludlow, P. (1993). Interpreted logical forms. *Synthese*, 95:305–355.
- [Lewis, 1979] Lewis, D. (1979). Attitudes *de dicto* and *de se*. *The Philosophical Review*, 88:513–543.

- [Lewis, 1986] Lewis, D. (1986). *On the Plurality of Worlds*. Basil Blackwell, Oxford.
- [Ludlow, 1995] Ludlow, P. (1995). Logical form and the hidden indexical theory. *The Journal of Philosophy*, 92:102–107.
- [Ludlow, 1996] Ludlow, P. (1996). The adicity of ‘believes’ and the hidden indexical theory. *Analysis*, 56(2):97–101.
- [Martinich, 1990] Martinich, A. P. (1990). *The Philosophy of Language, second edition*. Oxford University Press, New York.
- [Matthews, 1994] Matthews, R. J. (1994). The measure of mind. *Mind*, 103(410):131–146.
- [Recanati, 1993] Recanati, F. (1993). *Direct Reference: From Language to Thought*. Blackwell, Cambridge, Massachusetts.
- [Richard, 1983] Richard, M. (1983). Direct reference and ascriptions of belief. *Journal of Philosophical Logic*, 12:425–452. Reprinted in [Salmon and Soames, 1988].
- [Richard, 1990] Richard, M. (1990). *Propositional Attitudes: An Essay on Thoughts and How We Ascribe Them*. Cambridge University Press, New York.
- [Robbins, 2003] Robbins, P. (2003). To structure, or not to structure. *Synthese*. In press.
- [Salmon, 1986] Salmon, N. (1986). *Frege’s Puzzle*. MIT Press, Cambridge, Massachusetts.
- [Salmon and Soames, 1988] Salmon, N. and Soames, S. (1988). *Propositions and Attitudes*. Oxford University Press, New York.
- [Schiffer, 1977] Schiffer, S. (1977). Naming and knowing. *Midwest Studies in Philosophy*, 2:28–41.
- [Schiffer, 1987] Schiffer, S. (1987). *Remnants of Meaning*. MIT Press, Cambridge, Massachusetts.
- [Schiffer, 1992] Schiffer, S. (1992). Belief ascription. *The Journal of Philosophy*, 89:490–521.
- [Schiffer, 1996] Schiffer, S. (1996). The hidden-indexical theory’s logical form problem: A rejoinder. *Analysis*, 56(2):92–97.
- [Soames, 1987] Soames, S. (1987). Direct reference, propositional attitudes, and semantic content. *Philosophical Topics*, 15:47–87. Reprinted in [Salmon and Soames, 1988].

- [Stalnaker, 1981] Stalnaker, R. C. (1981). Indexical belief. *Synthese*, 49:129–151.
- [Stalnaker, 1984] Stalnaker, R. C. (1984). *Inquiry*. MIT Press, Cambridge, Massachusetts.