A Plea for Monsters

Philippe Schlenker

A PLEA FOR MONSTERS[1]

Philippe Schlenker

0 Introduction

According to an influential theory (Kaplan 1989), indexicals are directly referential: they pick out their referents directly from the context of utterance, without the mediation of a Fregean Sense. A corollary of this theory is that the value of an indexical is fixed once and for all by the context of utterance, and cannot be affected by the logical operators in whose scope it may appear. This is summarized in the following thesis, which makes indexicals 'scopeless' (to borrow Kaplan's recent terminology):

(1) Fixity Thesis (a corollary of Direct Reference): The semantic value of an indexical is fixed solely by the context of the actual speech act, and cannot be affected by any logical operators.

The term 'context' is taken to apply to any speech situation[2], not just to the speech act in which a sentence is actually produced. Thus operators could be defined that shift the context of evaluation of an indexical[3]. Although Kaplan grants that such operators are conceivable, he calls them 'monsters' because they contradict his thesis. He goes on to claim (optimistically) that monsters do not and *could not* exist in natural language.

On the face of it, this conclusion appears to be correct. This is best seen by contrasting the behavior of definite descriptions with that of indexicals. I may utter truly: 'At some point, the person talking was John'. But I couldn't do the same with: 'At some point, I was John' (unless I am John, of course). Why should that be? The definite description can take scope below the temporal operator, with a reading ('de dicto') rendered in (2a), equivalent to: at some point in the past, the person talking *at that time* was John. And of course there is no reason the current speaker (me) should have been talking at all past moments as well. An alternative, 'de re' reading is represented in (2b), where the description is given wide scope:

(2) a. P [ιx: S(x)] (x=John)
b. [ιx: S(x)] P (x=John)
a'. [∃t: t<t*][ιx: speaker(x, t)] (x=John)
b'. [∃t: t<t*][ιx: speaker(x, t*)] (x=John)

The same contrast can be represented in an extensional system without scopal difference, as in (2a') and (2b'). The distinction there is whether the time variable of the predicate 'speaker' is bound by the time quantifier, or is left free, with a contextually supplied value (here 't*' is taken to denote the time of utterance).

Interestingly, the indexical 'I' appears only to have the equivalent of the 'wide scope' reading represented in (2)a and (2)a'. In this case one might well argue that this is simply because indexicals are *context*-dependent rather than time-dependent. If we could find a quantifier over *contexts* rather than over moments, maybe it *could* affect the semantic value of 'I'. But this is precisely what Kaplan denies. Since indexicals are directly referential, their semantic value *cannot* be manipulated in this fashion. The failure of our attempt to get 'I' to refer to someone other than me, P.S., is not the result

[1] I wish to thank the following for comments on various stages of this research, which originated in my MIT dissertation: J. Almog, C. Barker, R. Bhatt, N. Chomsky, P. Dekker, A. Fahrhan, K. von Fintel, K. Fine, M. Hackl, K. Hale, I. Heim, J. Higginbotham, S. Iatridou, R. Kayne, E. Keenan, P. Ludwig, O. Matushansky, J.C. Pariente, O. Percus, D. Pesetsky, D. Petros, F. Recanati, K. Safir, U. Sauerland, B. Schein, Y. Sharvit, D. Sportiche, R. Stalnaker, A. von Stechow, I. Stojanovic, T. Stowell, A. Szabolcsi, H. Zeevat. Special thanks to J. Almog, I. Heim and A. von Stechow for extended discussions and criticisms, and to two anonymous reviewers for extremely constructive comments.

[2] This notion will be extended below to cover contexts of thought as well as contexts of speech.

[3] If the term 'context' were defined as 'a situation on which the semantic value of certain expressions (indexicals) depends rigidly', the Fixity Thesis would become true by definition, and there would be nothing left to discuss. I do not think, however, that this is how Kaplan understands the term 'context'.

of choosing the wrong quantifier. The attempt was doomed on principled grounds. This result is clearly positive *if* it is assumed (as is commonly done) that indexicals do indeed display the behavior predicted by the Fixity Thesis. In fact, a theory that posited that indexicals have the same kind of Sense as descriptions would be at a loss to account for such a behavior. In particular, since (according to Frege) an expression in indirect discourse can refer to its ordinary Sense, a sentence such as 'John thinks that I am a hero' should have a reading that attributes to John a thought of the form: 'I am a hero'. But in that case 'I' would *fail* to refer to me, the speaker of the actual speech act. If a theory based on Sense is correct, why does this case not occur?

Why not, indeed? I will claim that this case does in fact occur, that it falsifies the Fixity Thesis, and that it argues for a version (albeit a somewhat elaborate one, which involves variables and presuppositions) of a Sense-based theory. Specifically, I will display several indexicals across languages which display a 'monstrous' behavior when they appear in the scope of an attitude operator. Here is as schematic example (from Amharic):

(3) Situation to be reported: John says: 'I am a hero'
 a. Amharic (lit.): $John_i$ says that I_i am a hero
 b. English: $John_i$ says that he_i is a hero / *$John_i$ says that I_i am a hero

Although English 'I' is well-behaved, its Amharic counterpart isn't, and may in (3a) refer to the speaker of the *reported* speech act rather than of the actual speech act[4]. The value of the Amharic 1st person pronoun can in this case be affected by an attitude operator, which contradicts the Fixity Thesis. (Amharic is the key here, but even if you distrust such exotic data, or are solely interested in English, do not rejoice too quickly. Similar examples can be constructed for English, albeit in a different domain. Details are given below.)

At this point one might be tempted to plead terminological ambiguity. In what sense is the Amharic expression an 'indexical' in Kaplan's sense? Let us use context-dependency as a Definition: *an expression qualifies as indexical if its semantic value is determined by some feature of a context of utterance*. We may further distinguish expressions that are lexically specified as indexical ('strict indexicals', which can only be interpreted indexically) from those that have indexical *uses*. For instance 'he' can be used indexically (e.g. with a pointing gesture); but it can also be a bound variable, and thus it is not lexically specified as indexical. According to this criterion Amharic 'I' *is* a strict indexical: it always refers to the speaker of *some* context. As it happens, this doesn't have to be the context of the *actual* speech act, and for this reason the expression fails to satisfy the Fixity Thesis. Obviously if the Thesis were taken as a definition, the purported counterexamples would go away. However the interesting claim (the one I want to challenge) is that the class of expressions singled out by the above Definition happen to satisfy the Fixity Thesis[5].

The reason there are systematic counterexamples to the Thesis, I will claim, is that quite generally an attitude report manipulates a context variable, whose value may fix the reference of indexicals that appear in its scope. In traditional model-theoretic accounts, attitude verbs are essentially construed as quantifiers over possible worlds. Thus 'John believes that it is raining' is true just in case it is raining in every world compatible with John's belief. I will argue for a minimal modification of this analysis. What shifted indexicals of the Amharic variety show, I'll suggest, is that attitude verbs are *quantifiers over contexts of thought- or of speech*. Since a context determines a single world of thought or utterance, and other things in addition (a speaker, a time of utterance,

[4]It could be suggested that (3)a is an instance of quotation rather than of indirect discourse. This would make the example theoretically uninteresting, since it is uncontroversial that an indexical that appears in a quotation does not have to be evaluated with respect to the context of the actual speech act. It is shown below that some instances of (3)a are *not* quotations in anything like the usual sense.

[5]The claim is substantive only so long as one refrains from pleading systematic ambiguity in the face of counterexamples. 'You say that item i falsifies the Thesis because its semantic value may be fixed *either* by the context of the actual speech act *or* by some other context. But this only goes to show that i is ambiguous between i_1 and i_2. i_1 is a true indexical, and satisfies the Thesis. i_2 is no indexical at all'. The Thesis has been saved, but at the price of vacuity.

and sometimes a hearer), the theory to be presented is strictly more fine-grained than the standard one, in a way which is systematically related to (though conceptually different from) Chierchia's theory of Attitudes De Se (Chierchia 1987). In the logical syntax attitude verbs will be represented roughly as follows[6] (which is the logical form of an Amharic sentence):

(4) $\text{SAY}_{<\text{John,now,actually}>}c_i$ be-a-hero (agent(c_i), time(c_i), world(c_i))

↑

context of the
reported speech act

To put it somewhat sloppily, the context of the reported speech act is represented in the logical form as a context variable (c_i), bound by the attitude operator (analyzed here extensionally, as a universal quantifier). This will allow an indexical to depend either on the context of the actual speech act, as is generally the case in English, or on the context of the reported speech act, as can happen in Amharic (hence *agent(c_i)*, which picks out the agent of the context c_i). If correct, this theory will vindicate an insight formulated by Israel & Perry 1996, who suggested (without empirical argument) that one should *expect* attitude verbs to systematically violate the Fixity Thesis stated above.

The rest of this paper is organized as follows. I first lay out the problem posed by indexicals for a Fregean theory of Sense (Section 1), and then go on to discuss Kaplan's solution as presented in *Demonstratives* (Section 2). In order to eliminate inessential sources of inadequacy, an extensional system is introduced in which Kaplan's theory can be restated (Section 3). It is then shown that Kaplan's theory is empirically inadequate because his semantics for attitude reports is insufficiently fine-grained; and it is further demonstrated that the source of the problem lies in the prohibition against monsters (Section 4). A solution is then offered, which follows (in simplified form) the theory of Schlenker 1999 (Section 5). Two major stipulations of that theory are then criticized and partially eliminated in the final parts of the paper (Sections 6 and 7).

1 The Problem: Fregean 'Senses' and Indexicality

1.1 Frege's notion of 'Sense'

Frege's notion of 'Sense' was designed to kill two birds with one stone. It was supposed to account *both* for the cognitive significance of unembedded sentences and for the truth-conditional contribution of clauses embedded under an attitude operator. Two sentences might thus have the same value but differ in cognitive significance because they have different Senses. And by the same token sentences with the same truth-value may still fail to be substitutable *salva veritate* under an attitude operator if they do not share the same Sense. On Frege's view, then, it is for the same reason that 'The Morning Star is the Morning Star' and 'The Morning Star is the Evening Star' don't have the same cognitive status even though they have the same truth-value; and that 'John believes that the Morning Star is the Morning Star' can be true even when 'John believes that the Morning Star is the Evening Star' is false. In both cases 'Sense' is the key, which allows the cognitive significance of a sentence *and* the truth-conditional contribution of an embedded clause to be strictly more fine-grained than their mere 'Reference'. (In Fregean parlance, this is because an element embedded under an attitude verb has its 'indirect' rather than its normal reference. And an element's 'indirect reference' is nothing but its normal Sense.)

[6] For simplicity I have picked a syntax that is closer to First-Order Logic than to natural language. This should be seen as an idealization. In principle, the goal is to find rules of semantic interpretation for the syntactician's 'Logical Forms'. Nothing essential hinges on the choice of the syntax in this paper, since the action is in the morphological and in the semantic components.

1.2 The Problem with Indexicals: Cognitive Significance and Truth-Conditional Contribution

1.2.1 Cognitive Significance

Things become more complicated when indexicals are brought into the picture. The first observation is that the cognitive significance of a sentence containing indexicals may crucially depend on its context-sensitivity. The following example, due to Kaplan, should suffice to convince the reader:

(5) 'If I see, reflected in a window, the image of a man whose pants appear to be on fire, my behavior is sensitive to whether I think, 'His pants are on fire', or 'My pants are on fire', though the object of thought may be the same (i.e. the proposition expressed may be the same, P.S.)'. (Kaplan 1990)

What Kaplan calls the 'object of thought' is what the sentence tells us about the world, or what I will call its 'objective Sense'. In a possible worlds framework, the object of thought is just a proposition (a set of possible worlds). Thus in Kaplan's scenario 'his pants are on fire' and 'my pants are on fire' say the same thing about the world, characterized as a world in which Kaplan's pants are on fire. The difference, however, is that in the 1st person case it is asserted that the context of speech is one in which the *speaker*'s pants are on fire; nothing similar is implied in the 3rd person case, which simply indicates that *the person who is pointed at* is in that unfortunate situation. Thus the two sentences convey the same information about the world, but they say different things about the context of speech. Kaplan's example shows that the latter sort of information is crucial to the cognitive significance of a sentence: I might stand to watch regretfully (or sadistically) in the 3rd person case, but I am bound to take immediate action if my thought is in the 1st person. In other words, a sentence doesn't just tell us in which world we are situated, but also in which context of speech or thought we are located. Although Kaplan doesn't use the terminology[7], this is just the *De Se* problem, as described for instance by Lewis 1979; this, in turn, is identical to the 'problem of the essential indexical' as discussed by Perry 1979. The conclusion is that indexicals are cognitively irreducible, and that the notion of 'Sense' responsible for the cognitive significance of a sentence must itself be indexical.

1.2.2 The Dilemma: abandon the unity of Frege's Sense or abandon the Fixity Thesis

At this point two possibilities are open. We may preserve Frege's insight, and conclude that the Sense responsible for the truth-conditional contribution of an embedded clause is also irreducibly indexical, contrary to common wisdom. Alternatively, we may grant that the cognitive significance of a sentence is given by an indexical Sense, but still maintain that the truth-conditional contribution of an embedded clause is given by an objective Sense. This means abandoning the unity of Frege's notion of 'Sense', a move advocated with considerable success in Kaplan's *Demonstratives.* Frege's notion is then ramified into two components: an objective Sense (Kaplan's 'Content'), identified with a set of possible worlds (a proposition), is responsible for the truth-conditional contribution of embedded clauses; while an indexical Sense, which is strictly more fine-grained and is roughly equivalent to Kaplan's 'Character', is supposed to account for the cognitive significance of unembedded sentences (in the above example, 'My pants are on fire' and 'His pants are on fire' have the same Content, but different Characters).

[7] Presumably this is because (as an anonymous reviewer points out) Kaplan's manuscript was already finished when Lewis published his 'Attitudes De Dicto and De Se' in 1979.

1.3 Two Theories

1.3.1 Kaplan's Theory

On superficial inspection, there appears to be overwhelming support for Kaplan's position. First, since objective Senses are strictly less fine-grained than indexical Senses, Kaplan's approach predicts that the indexical nature of a direct discourse should systematically be lost in a report. Specifically, attitude reports should not have the ability to distinguish between thoughts that have the same objective Sense but different indexical Senses (in Kaplan's terminology: between thoughts that have different Characters but the same Content). The prediction *appears* to be borne out. In the foregoing scenario, the agent may have thought 'My pants are on fire' or 'His pants are on fire' (where 'his' refers to the agent himself), but both thoughts are reported in the same way: 'He_i thought that his_i pants were on fire'.

As was mentioned in the introduction, there is a second apparent argument in favor of Kaplan's position. If an element embedded under an attitude verb could refer to its indexical rather than to its objective Sense, a sentence such as 'John thinks that I am a hero' should have a reading on which John's belief is that *he, John,* is a hero, contradicting both the Thesis and the facts of English. The reasoning is that on Frege's theory the Sense of the embedded clause is just the thought we attribute to John. Thus if the Sense is indexical, the thought could be indexical as well, and have the form: 'I am a hero'. But this means that we would attribute to John the thought that *John* is a hero - incorrectly, as it appears. In fact, the empirical content of the Thesis is precisely that such readings should never occur, since they would allow an indexical to be evaluated with respect to something other than the context of the actual speech act. Kaplan grants that operators that shift the context of evaluation of an indexical are perfectly well-formed from a logical standpoint, but he claims that such monstrous operators do not exist in English and *could not even be added to it*.

1.3.2 An alternative

An alternative is possible, however, which seeks to restore the unity of Frege's notion of Sense, and thus to challenge Kaplan's theory. In what follows I grant that the cognitive significance of a sentence is given by an indexical Sense, but unlike Kaplan I posit that the truth-conditional contribution of a clause embedded under an attitude verb is given by an indexical Sense as well[8]. (This idea is discussed in von Stechow 1982, Lewis 1981 ('Postscript to General Semantics'[9]), and implemented in Chierchia 1987 and Heim 1991. It is conceptually related to -but technically very different from- Kuno 1972's 'direct discourse' analysis of attitude reports). In order to reach such a conclusions, I deny Kaplan's empirical premise, along the following lines.

First, I review arguments that show that *the indexical nature of a thought or discourse can in some cases be retained in a report.* Castañeda 1968 devised an artificial pronoun, 'he*', to represent in reported speech the use of a 1st person in direct discourse. Thus 'John_i says that he_i* is a hero' is Castañeda's way of reporting (in an artificial language) 'I am a hero' as uttered by John. In case John thought 'He is a hero', the report in Castañeda's system simply becomes 'John_i says that he_i is a hero', without 'he*'. Castañeda called elements such as 'he*' 'quasi-indicators', to emphasize their connection with 'indicators', or, in current terminology, 'indexicals' (in order to make the connection completely explicit, I'll henceforth use the term 'quasi-indexicals' rather than of 'quasi-indicators'). Kaplan's empirical prediction is that quasi-indexicals could not exist in natural language. But this is incorrect. Castañeda had already conjectured that 'he himself' can only be read as 'he*', and never as 'he'; and recent linguistic research has convincingly established that 'he*' does indeed exist in

---

[8] The 'indexical Sense' we will end up defining is simpler than Kaplan's Characters. It will turn out to be the *diagonal* of a Kaplanian Character. See below for discussion.

[9] In his 'Postscript' Lewis discussed explicitly a De Se analysis of some infinitives. His example (which is similar to those of Morgan 1970 and Chierchia 1987) was: 'Lothario strives to find x'. Lewis sketched a De Se analysis based on the theory of Lewis 1979. But while the latter dealt with attitudes, the 1981 postscript discussed attitude *reports.*

natural language, under various guises (Morgan 1970; Mitchell 1986; Chierchia 1987; Kratzer 1997, Kusumoto 1998).

The second part of the refutation lies in the demonstration that *some indexicals can systematically be shifted in attitude reports*. As was mentioned in the introduction, the reading which is excluded in English is systematically available in Amharic, where (literally) 'John says that I am a hero' *can* mean that John thinks that he (John) is a hero[10]. Furthermore, it will be shown that similar facts can be replicated within English, although in a different domain (that of temporal adverbials). Shifted indexicals make the same semantic point as quasi-indexicals, since they show that the indexical nature of a thought or discourse can in fact be retained in reported speech. But unlike quasi-indexicals they wear their indexicality on their sleeves, so to speak. And this provides a direct rebuttal to Kaplan's claim concerning the non-existence of monsters, and an important hint to the correct theory. As I will argue, in truth *every attitude verb is a Kaplanian monster*.

A particularly simple and elegant version of the present theory was developed within a modal framework in von Stechow 2001 (cf. also Israel & Perry 1996). I do not adopt it in the general case because it is insufficiently expressive and requires numerous operations of covert movement for which there is no syntactic evidence. However it is particularly well suited for a comparison with Kaplan's logic of demonstratives, of which it is a minimal (but monstrous) variation. In a nutshell, the suggestion is that attitude operators manipulate a context parameter rather than a world parameter. *John says that* ϕ is thus analyzed as: every *context* compatible with John's claim is one in which ϕ could be uttered truly. Von Stechow's monstrous operator can be defined as follows, where $\text{Assert}^{m}_{J}(t, w)$ is the set of *contexts* compatible with John's assertion at t in w[11]:

(6) $|=_{c, f, t, w} \text{Say}^{m}_{J}\ \phi$ iff for all $c' \in C$ such that $c' \in \text{Assert}^{m}_{J}(t, w)$: $|=_{c', f, \text{time}(c'), \text{world}(c')} \phi$
(*John says that* ϕ is true in context c at time t in world w just in case every context compatible with what John asserts at t in w is one in which ϕ could be uttered truly)

By contrast, on a standard modal analysis the attitude operator only manipulates a *world* parameter. As a result, *John says that* ϕ is taken to be true just in case ϕ is true in every world compatible with John's assertion (here $\text{Assert}_{J}(t, w)$ is the set of *worlds* compatible with John's assertion at t in w):

(7) $|=_{c, f, t, w} \text{Say}_{J}\ \phi$ iff for all $w' \in W$ s.t. $w' \in \text{Assert}_{J}(t, w)$: $|=_{c, f, t, w'} \phi$
(*John says that* ϕ is true in context c at time t in world w just in case every world compatible with what John asserts at t in w is one in which ϕ is true)

No context parameter is manipulated in the standard analysis, and thus no indexical may fail to be evaluated with respect to the actual context. Such isn't the case on the monstrous account. With a standard definition of 'I' (*'I' evaluated in a context c denotes the author of c*), it is now predicted that a first-person pronoun embedded under an attitude verb may *fail* to refer to the actual speaker. For the Amharic first person (abbreviated as 'I^{Amh}'), this is the right result:

(8) $|=_{c, f, t, w} \text{Say}^{m}_{J}\ \text{hero}(I^{Amh})$ iff for all $c' \in C$ s.t. $c' \in \text{Assert}_{J}(t, w)$: $|=_{c', f, \text{time}(c'), \text{world}(c')} \text{hero}(I^{Amh})$
(*John says that* I^{Amh} *am a hero* is true in context c at time t in world w just in case every context c' compatible with what John asserts at t in w is such that: the author of c' is a hero in the world of c' at the time of c')[12]

Castañeda's 'he*' (which corresponds to logophoric pronouns in natural language) can be given the same semantic treatment, with the additional *syntactic* stipulation that it may only appear within the scope of an attitude operator. English 'I' is the mirror-image of 'he*': it shares the same semantics but can only appear *outside* the scope of an attitude operator. 'John says that I am a hero' would thus

[10] These facts have long been known in the typological literature (e.g. Anderson & Keenan 1985), and are occasionally hinted at in formal semantics, e.g. in Partee 1989 (fn. 2), who credits Emmon Bach for this observation.
[11] Note that in Kaplan's and Stalnaker's terms, Say^{m}_{J} operates on the diagonal of a Character.
[12] 'f' is an assignment function. It is necessary in order to treat '$I_{x}\ \text{Say}^{m}_{J}\ \text{hero}(x)$' below.

be analyzed in this system as 'I_x Say^m_J hero(x)', where 'I' has had to leap out of the scope of the attitude operator by an operation of invisible movement; this allows it to be evaluated with respect to the actual context, as it should be (such a movement won't be necessary in the final version of the system developed here).

2 Kaplan's Theory: indexicals and indirect discourse

Let us now turn to a more systematic discussion of the conceptual situation. In this section I briefly summarize Kaplan's well-known theory of indexicals, and then focus on his (less well-known) theory of indirect discourse (from Kaplan 1989). I then modify the theory to rid it of an inessential inadequacy. Both versions of Kaplan's theory are criticized in the latter parts of the paper.

2.1 Kaplan's Analysis of Indexicals

2.1.1 Context and World Parameters in an Intensional Setting

In *Demonstratives* Kaplan seeks to offer an account of the logic of indexicals, one that can both explain why 'I am here now' is in some sense *a priori* true, even though 'I am necessarily here now' is so obviously false. Kaplan's endeavor is framed within a standard modal logic, a point we'll modify for empirical reasons in Section 3.2. Suffice it to say for the moment that Kaplan adds to the standard world (or time) parameter of a modal logic a context parameter, which allows him to derive the following results:

(9) a. 'I am here now' is *a priori* true because in all possible contexts c, 'I am here now' uttered in c is true in c. Formally:
for all contexts c, for all assignments f, $|=_{c, f, time(c), world(c)}$ I am here now
b. 'Necessarily I am here now' is not true. Formally:
$|=_{c, f, time(c), world(c)} \Box$I am here now iff for all w' $\in$ W s.t. world(c)Rw' $|=_{c, f, time(c), w'}$ I am here now. But there are worlds w' such that $|\neq_{c, f, time(c), w'}$ I am here now.

Once Kaplan's system is set up in this (modal) fashion, a natural question is whether there are any operators that can manipulate the context parameter c. Formally, nothing prevents this, as is demonstrated by von Stechow's and Israel & Perry's monstrous operator, defined in (6) above. Still, Kaplan claims that such operators do not and could not exist in English. Here is his reasoning:

> Are there such operators as 'In some contexts it is true that', which when prefixed to a sentence yields a truth if an only if in some context the contained *sentence* (not the content expressed by it) expresses a content that is true in the circumstances of that context? Let us try it:
>
> (9) In some contexts it is true that I am not tired now.
>
> For (9) to be true in the present context it suffices that some agent of some context not be tired at the time of that context. (9), so interpreted, has nothing to do with me or the present moment. But this violates Principle 2! *[NB. Principle 2 states that 'Indexicals, pure and demonstrative alike, are directly referential'. P.S.]*. Principle 2 can also be expressed in a more theory laden way by saying that indexicals always take primary scope. If this is true - and it is - then no operator can control the character of the indexicals within its scope, because they will simply leap out of its scope to the front of the operator. I am not saying we could not construct a language with such operators, just that English is not one. And such operators *could not be added to it.* (Kaplan 1989 p. 510)

Kaplan's only empirical argument appears to be a generalization based on a single (but representative) example, namely that the 'operator' 'in some contexts it is true that' cannot shift the context of an indexical. To anticipate somewhat, here is what I will conclude about this argument: Kaplan picked the wrong example. If he had selected an attitude operator (e.g. 'John believes that...') instead of 'in some contexts ...', and if he had looked at Amharic instead of English, or alternatively if he had stuck to English but looked at a different indexical ('two days ago' instead of 'I'), he would

have seen that *monsters do exist.* For the moment, however, let me review the consequences that Kaplan draws from his observation.

2.1.2 Character and Content

In modal re-interpretations of Frege's theory of Sense, attitude verbs are analyzed as modal operators, while Senses are re-interpreted as intensions. Kaplan inherits this system, which he alters to account for context-dependency. In fact, the revision he advocates is rather minimal. True, the logic must be enriched so as to encompass contexts. But since the Principle of Direct Reference prevents context parameters from being manipulated, once the context of the actual speech act is fixed, everything remains as it used to be in a standard modal logic. As a consequence, the only notion that is relevant for truth-conditional purposes *given a fixed context* is the 'Content' or proposition expressed by a clause.

Of course, to account for the cognitive significance of a sentence context-dependency must be taken into account. This is why Kaplan introduces the notion of a Character, which is just a function from Contexts to Contents. The cognitive significance of an utterance (its 'subjective meaning', to use the terminology of Haas-Spohn 1995) lies in what it asserts about the *context* in which the speech act occurred. Although Kaplan does not elaborate on this point, the cognitive significance is thus given by the *diagonal* of a Character χ, i.e. the set of contexts c such that χ uttered at c is true in the world c_w of c (the diagonal of χ is defined as: $\Delta(\chi)=\lambda c\ \chi(c)(c_w)$; see Stalnaker 1978 and Haas-Spohn 1995). In the end, then, the account is twofold: Characters (or rather: the diagonals of Characters), which are indexical Senses, account for the cognitive significance of a sentence. Contents (objective Senses) are used whenever a world parameter is manipulated. In particular, Contents are supposed to account for the truth-conditional contribution of a clause embedded under an attitude operator.

2.2 Kaplan's Analysis of Indirect Discourse

Because of the Prohibition Against Monsters, context parameters can never be shifted. As a result, attitude reports cannot be construed as relations between individuals and Characters, but only as relations between individuals and Contents. This is but the formal counterpart of the empirical observation we made earlier (and will challenge shortly): two thoughts or discourses that differ in Character but not in Content appear to be reported in the same way in indirect discourse. Whether the agent thinks 'My pants are on fire' or 'His pants are on fire', the report will be of the form: 'He_i thinks that his_i pants are on fire'. And since Kaplan also maintains that attitudes (though not attitude reports) are -irreducibly- relations between individuals and Characters rather than Contents, he must conclude that the indexical nature of a thought is systematically lost in a report.

From this perspective, the challenge is to give an account of indirect discourse that explains how an attitude report can be construed in terms of Contents even though the attitude itself is analyzed in terms of Characters. The solution is simply to provide a theory of indirect discourse that makes attitude operators sensitive only to the Content, and never to the Character of the thought they seek to report. Kaplan sketches such a proposal in *Demonstratives*. This attempt is important for two reasons. First, it shows that Kaplan's theory of Characters and Contents can give a coherent account of indirect discourse. Although the theory will be refuted, this will be on empirical, not conceptual grounds. Second, Kaplan's theory of indirect discourse has the advantage of displaying with great clarity the logical independence of two issues that are often confused in the linguistic literature. Often one (correctly) observes that (i) a thought or discourse is essentially indexical or 'De Se', and goes on to conclude without further argument that (ii) the report must be De Se as well. But the conclusion simply *does not* follow. Proof: in Kaplan's theory, the De Se nature of the thought is granted, but the report can never be unambiguously De Se, although it is of course *compatible* with a De Se situation.

Here, then, is Kaplan's account of indirect discourse, which is buried in Paragraph 20 of *Demonstratives:*

> What is special and different about the present approach is the attempt to use the distinction between direct and indirect discourse to match the distinction between character and content. Thus when you wonder, 'Is that me?', it is correct to report you as having wondered whether you are yourself. These transformations are traced to the indexical form of your inner direct discourse rather than to any particular referential intentions. The idea is that the full analysis of indirect discourse includes mention of the suppressed character of the direct discourse event which the indirect discourse reports, thus:
> (i) ∃c, C [c is a context & C is a character & x is the agent of c & x direct-discourse-verb C at the time t of c & the content of C in c is that...]
> approximates a full analysis of
> (ii) x indirect-discourse-verb that ... at t." (Kaplan 1989 p. 554)

Consider again Kaplan's own example: 'My pants are on fire' vs. 'His pants are on fire'. In the relevant context (with K. as the agent of both thoughts), these utterances (or thoughts) have the same content, namely that K.'s pants are on fire. And they get reported in the same way: 'K. says that his pants are on fire.' To simplify things, let us assume for a second that only two Characters with the same Content could be asserted by the agent (the 1st person and the 3rd person ones). In this special case, Kaplan's analysis gives the following truth-conditions:

(10) 'K.$_i$ says that his$_i$ pants are on fire' is true (at t* in w*)
iff K. asserts (at t* in w*) the Character of 'My pants are on fire' or K. asserts (at t* in w*) the Character of 'His pants are on fire' (with 'his' referring to K.).

But of course there could in principle be any number of Characters that yield the same Content in the context of K.'s utterance. In the general case, then, all we require for the report to be true is that there be *some* Character satisfying the relevant conditions which is asserted by K. This gives the following truth-conditions, which involve an existential quantification over Characters that share the same Content:

(11) 'K.$_i$ says that his$_i$ pants are on fire' is true (at t* in w*)
iff there is a Character χ such that:
(i) the content of χ given the context of K.'s speech act (call it 'c') is that K's pants are on fire (i.e. $\chi(c)=\lambda w$ K.'s pants are on fire in w), and
(ii) K. asserts χ (at t* in w*)

Kaplan's theory is satisfactory given the facts we have seen so far. It is compatible both with what we know about (real-world) propositional attitudes (they are irreducibly indexical), and with the linguistic form of attitude reports (indexical distinctions get collapsed in indirect discourse, and indexicals can never be shifted).[13] We will see below that the linguistic assumptions are empirically

---

[13] It is instructive to look at a more subtle application of Kaplan's theory. Consider the following case:

(i) a. Heimson believes: 'I am David Hume'
b. Heimson believes that he is David Hume

(ia) was one of the many arguments for an analysis of attitudes in terms of objects that are strictly more fine-grained than propositions. While Heimson is certainly wrong, he need not be irrational. But if the objects of attitudes were propositions, Heimson would have to believe the empty proposition (i.e. λw Heimson=Hume in w), and thus be irrational - not a welcome result. The argument seems cogent when applied to thoughts or utterances. But does it also go through when applied to indirect discourse? Consider (ib) again. Using Kaplan's analysis of Indirect Discourse, we obtain the following result [c* is the context such that agent(c*)=Heimson, time(c*)=t*, world(c*)=w*]:

(ii) $\exists C$ [C is a character & Heimson believes C at t* in w* & $C(c^*)=\lambda w$ [Heimson=Hume w]]

Now it is indeed the case that this does *not* attribute irrationality to Heimson. For although λw [Heimson=Hume w]=Ø, there are non-empty Characters that yield precisely this when applied to the context of Heimson's thought act. In fact the Character of 'I am Hume' has exactly this property:

(iii) a. Character of 'I am Hume': $\lambda c\ \lambda w$ author(c)=Hume in w
b. $[\lambda c\ \lambda w$ author(c)=Hume in w$](c^*)=[\lambda w$ Heimson=Hume in w] = Ø

Now surely it is not irrational to believe 'I am Hume', and it is clear that the corresponding Character is non-empty.

Still, there is an important problem with this solution (the problem was pointed out to me by Henk Zeevat; it is discussed in Zimmermann 1991, pp. 206-207). The analysis makes the incorrect prediction that 'Heimson believes that

incorrect: attitude reports can in fact preserve the indexical distinctions that are found in a direct discourse, and furthermore indexicals can in some cases be shifted. But before we can tackle these issues, we have to remove an independent source of inadequacy in Kaplan's proposal.

3 An Extensional Framework

3.1 Intensional vs. Extensional Analyses

The independent problem is that Kaplan's system is framed within a standard modal logic. This has the advantage of making his proposal conceptually and technically simple. But unfortunately standard modal logic does not have the expressive means to model natural language quantification over times and possible worlds, even independently of the issue of indexicality (this is laid out in great detail in Cresswell 1990). Furthermore, the same point carries over to the representation of attitude reports, at least on the assumption that these should be analyzed in modal terms to begin with, that is, as a sort of (implicit) quantification over possible worlds (rather than in quotational terms, for instance; I do not discuss the latter possibility in this paper). Kaplan's system inherits the expressive inadequacy of standard modal logic, and since issues of expressive power do interact with the treatment of indexicality, the inadequacy must be removed before we can fully analyze the relation between indexicality and attitude reports. This is particularly important because our final theory would be extremely cumbersome to express in an intensional setting.

3.1.1 The trouble with a standard modal system (1): time and modality

Cresswell 1990 showed systematically that the full power of overt quantification over times and possible worlds is needed to handle modal and temporal talk in natural language. He considered examples of increasing complexity, showing at each step how a modal logic had to be enriched to handle them. Here are some examples, in which 'F', 'P' and '◊' stand respectively for the past, future and possibility operators in the purported (but incorrect) modal translations (b). In the extensional translations, the values of t_0 and w_0 are contextually supplied as the time of utterance and the world of utterance respectively:

(12) a. One day all persons now alive will be dead
b. # $F\forall x$ (alive x => dead x)
c. $[\exists t_1: t_1 > t_0] (\forall x (\text{alive}(x, t_0) \Rightarrow \text{dead}(x, t_1)))$

(13) a. Once everyone then alive would be dead
b. # $PF\forall x$ (alive x => dead x)
c. $[\exists t_1: t_1 < t_0] [\exists t_2: t_2 > t_1] (\forall x (\text{alive}(x, t_1) \Rightarrow \text{dead}(x, t_2)))$

---

he is Hume' should be true if and only if 'Heimson believes that he is Napoleon' is true, since in both cases the corresponding content is empty. In order to solve this problem within the framework of Kaplan's 'Adding 'Says'', we have to resort to the machinery of Kaplan's'Quantifying In' and give the following logical form (c* stands for the context of Heimson's thought act.):

(iii) $\exists\alpha_1 \exists\alpha_2 \exists\chi$ [R(α_1, Heimson, c*) & R(α_2, Hume, c*) & χ is a character & χ(c*)= [λw Heimson=Hume in w]=Ø & $\chi = \lambda c \lambda w$ [ιx: $\alpha_1(x, c)$]=[ιx: $\alpha_2(x, c)$] in w & Heimson believes χ at time(c*) in world(c*)]

where R(α_1, Heimson, c*) means that α_1 is a vivid description of Heimson for the agent of c* at the time of c* in the world of c*.

The results of this analysis are as follows:

(a) It is possible to distinguish between 'Heimson believes he is Hume' (true) and 'Heimson believes he is Napoleon' (false). The latter is false because Heimson does *not* believe any character of the form $\lambda c \lambda w$ [ιx: $\alpha_1(x, c)$]=[ιx: $\alpha_2(x, c)$], where [ιx: $\alpha_1(x, c)$] happens to pick out Heimson in the actual world while [ιx: $\alpha_2(x, c)$] picks out Napoleon.

(b) On the other hand, Kaplan's basic insight is preserved: it is impossible to distinguish between a situation in which Heimson believes: 'I am Hume', and one in which he believes 'He is Hume', even if 'he' happens to refer to Heimson himself.

(14) a. It might have been that every person actually rich was poor
b. # $\Diamond\forall$x (rich x => poor x)
c. [$\exists w_1$: w_0R w_1] ($\forall$x (rich(x, w_0) => poor(x, w_1)))

As can be seen, the purported modal translations are semantically inadequate, in that in each case the innermost formula ends up contradictory, contrary to speakers' judgments. The problem is that a modal operator shifts the point of evaluation of every element that appears in its scope. But there are numerous examples in natural language in which an element that is within the syntactic scope of a modal operator is not semantically dependent upon it, as is seen in the preceding examples. Furthermore, even if one were to postulate that the relevant items can leap out of the scope of the operator by an operation of covert syntactic movement, a standard modal solution would still be inadequate on semantic grounds. This is because a standard modal system is strictly less expressive than a logic with full quantification over times and worlds. But since the entire expressive power of the latter is needed to handle modal and temporal talk, there can be no motivation for using a standard modal system.

Because Cresswell wanted to stick to the syntax of modal logic, he refrained from resorting to a system with world and time variables. Rather, he enriched his initial modal logic with new operators (somewhat similar to those used in Quine's 'Variables Explained Away'). The final result was expressively equivalent to an extensional logic with full quantification over time and world variables. Although Cresswell's discussion is convincing, a simpler proof of the same expressive results can be provided. The technique was first used by Boolos 1984 in his study of plurals in relation to 2nd order logic. Boolos wanted to show that, due to the existence of plurals, English has at least the expressive power of (monadic) 2nd order logic. To administer the proof, he provided a recursive translation procedure from 2nd order logic into English. In this fashion, every 2nd order sentence was given an English counterpart which was true in exactly the same models – which showed that English had at least the expressive power of 2nd order logic. Applying the same technique, we can show that a subpart of English that includes only temporal or modal talk (what I call 'Temporal' and 'Modal' English, respectively) has at least the expressive power of 1st order (polyadic) logic. The interested reader is referred to Appendix I for details.

A word might be in order to explain *why* such a proof can be given at all. The reason is that Temporal and Modal English both have a counterpart of the expressive tools of 1st order logic, namely:
(i) Unary and binary connectives ('It is not the case that' or simply 'not'; 'and', 'or')
(ii) 1st order quantifiers: 'necessarily' and 'possibly' for Modal English, 'always' and 'sometimes' for Temporal English
(iii) Variables
The translation of variables (iii) is the only point of some subtlety: it is because natural language has time and world variables that world and time predicates of any arity can be constructed, and thus that *any* formula of 1st order logic can be translated. This, in turn, is necessary to construct a systematic translation from *polyadic* 1st order logic to Temporal or Modal English. Without the device of variables, we might be stuck with a system that is strictly less expressive[14] (this is the case of standard modal logic, which is expressively equivalent to a proper subpart of 1st order logic that contains a single dyadic predicate of worlds -the 'accessibility relation' between worlds- and any number of *monadic* world predicates[15]). But Partee 1973 noted that tenses share a

[14] As an anonymous reviewer points out, the expressive superiority of explicit quantification over systems with implicit parameters does not hold for higher-order logics.
[15] This is the case only when modal logic is interpreted with satisfaction/validity in a model (Definitions: Let M=<W, R, V> be a model, where W is a set of worlds, R an accessibility relation, and V a valuation. φ is valid in the model M iff for every world w of W, $V_{M, w}(\phi)=1$. φ is valid in the frame F iff for every world w of W *and for any valuation V,* $V_{M, w}(\phi)=1$. With satisfaction/validity in a frame, modal logic is equivalent to a fragment of 2nd order logic). A systematic translation procedure can be found in van Benthem 1983 or in Blackburn et al. 2001.

number of the interpretive properties of pronouns; in particular, tenses, just like pronouns, can behave as bound variables[16]. Stone 1997 further suggested that, in similar fashion, moods can be construed as modal pronouns; and they too can be treated as bound variables. The results mentioned above (temporal and modal English have at least the expressive power of 1st order logic) follow directly from these linguistic observations, since they give a way to provide a systematic translation for variables. Note, however, that considerably less than Partee's and Stone's full thesis is needed to get this result. If you don't believe that tenses and moods can be bound (long-distance), just consider the word 'then', in its temporal and modal uses[17]:

(15) a. When John comes, *then* Mary is always happy
b. [∀t: John comes at t] (Mary is happy at t)

(16) a. If John comes, *then* Mary will necessarily be happy
b. [∀w: John comes at w] (Mary is happy in w)

What is crucial, of course, is that there should be no locality constraints on the binding of 'then'. For simplicity the Appendix relies on 'then' rather than just tense or mood to define the recursive translations between 1st order logic and Temporal or Modal English[18].

A final point is that the translation procedures discussed here (and implemented in the Appendix) lend themselves to an extension to Generalized Quantifiers. Thus I would suggest that the same proofs can be used to show that Temporal and Modal English have at least the power of a 1st order logic with generalized quantifiers. Since the matter has been investigated at great length by others (see Kratzer 1991 for references), and is of no immediate consequence to the study of attitude reports, I leave it aside in what follows.

3.1.2 The trouble with a standard modal system (2): attitude reports

A final point is that the results of the preceding section generalize to attitude verbs, which in the modal tradition are analyzed as operators that manipulate a world parameter. Since in the modal case the full power of quantification over possible worlds is needed, we can expect the same to hold of attitude verbs as well. And this does appear to be the case, although the argument is a little harder to make. My goal in this section is to show that *if* attitude verbs are construed as manipulating world parameters (or anything else, for that matter), then this should be done within a system that has at least the power of full quantification over possible worlds[19] (or whatever it is that is actually quantified over; I will show later that the elements that are quantified over are contexts rather than possible worlds, but this is conceptually and technically a different point.)

Things are more complicated than in the modal and in the temporal case, both because the variables are implicit rather than explicit (cf. 'then' in the previous examples), and also because it is

---

[16] Partee partly retracted her claim in Partee 1984.

[17] For reasons that I do not understand, 'then' is not licensed unless it is preceded by a restrictive 'when'- or 'if'-clause. Since such clauses don't hurt, I include them in the recursive translation procedure, but I make them semantically harmless by having them express a trivial truth, e.g. '0=0'. Note also that 'then' makes a pragmatic contribution which I disregard in this study. See Iatridou 1994 for an analysis.

[18] This procedure *justifies* the use of overt time and world variables, since the full expressive power of a variable-full system is needed to formalize temporal and modal talk in natural language. What we have not shown, however, is that this is the *only* way such expressive power could be achieved. And for good reason, since the expressive power of 1st order logic can be achieved in a variable-free system as well (cf. Quine 1960 and Kuhn 1980). Since the point I wish to make is semantic rather than syntactic in nature, I am happy to use overt variables in my Logical Forms, although the reader may remain agnostic as to the syntactic reality of variables in natural language (see Jacobson 1999 for a recent defense of variable-free semantics.)

[19] An independent argument for the same conclusion was given in Heim 1991a. There the observation was that restrictors of quantifiers must be allowed to scope out of modal contexts - a fact which was already noted in Farkas's work (e.g. Farkas 1997). The claim was that these restrictors do not move syntactically, but are indexed with a time or a world quantifier that may appear arbitrarily far up in the tree. An anonymous reviewer informs me that similar considerations are found in Bäuerle 1983.

not easy to find an attitude verb that does not involve an accessibility relation (e.g. 'believe' does not quantify over *all* worlds (or contexts) but only over those that are *compatible with the agent's beliefs*). As in the previous case, I have relegated the discussion of the translation procedure to Appendix I. For present purposes the important point is that Attitudinal English (English with only attitude verbs as quantifiers) does appear to have the resources we need for the translation, namely:
(i) Unary and Binary connectives ('it is not the case that', 'and')
(ii) Quantifiers - for instance 'he believes that', analyzed as a universal quantifier over possible worlds (*John believes that* ϕ is true iff *every* world compatible with John's belief satisfies ϕ. The difficulty discussed in Appendix I is that this is an instance of *restricted* rather than unrestricted quantification).
(iii) Variables, which unfortunately are not explicitly represented. That they are semantically present can be seen in the standard analysis of sentences such as the following:

(17) a. John believes that it rains more than it actually does
b. [$\forall$w: B(w)(w*)] ([ɪd: rain(d, w)] > [ɪd: rain(d, w*)])

(17)a. receives the analysis in (17)b., where B(w)(w*) abbreviates: 'w is compatible with what John believes in w*'. (17)b is of the form '[$\forall$w: B(w)(w*)] Rww*', which clearly involves a polyadic world predicate. It so happens that in this case w* is a free variable - on standard analyses, it simply denotes the actual world. But this is an inessential feature of this example. It suffices to embed the entire sentence under another attitude verb to obtain a polyadic world predicate whose two variables are both bound by an attitude quantifier:

(18) a. Mary claims that John believes that it rains more than it does.
b. [$\forall$w: C(w)(w*)][$\forall$w': B(w')(w)] ([ɪd: rain(d, w')] > [ɪd: rain(d, w)])

['C(w)(w*)' abbreviates: 'w is compatible with what Mary claims in w*']

3.2 An Extensional System.

3.2.1 Adding time and world variables

In order to address the expressive power problem, I lay out an enriched version of Kaplan's logic, one that involves overt quantification over times and worlds (as well as over individuals, of course). 'Always' and 'necessarily' are treated as time and world quantifiers [restricted and generalized quantification could be added easily if one so wished]:

(19) $[[\forall x\ \phi]]^{s} = 1$ iff for every object X, $[[\phi]]^{s[x->X]} = 1$

(20) $[[\text{Nec w } \phi]]^{s} = 1$ iff for every world W, $[[\phi]]^{s[w->W]} = 1$

(21) $[[\text{Alw t } \phi]]^{s} = 1$ iff for every moment T, $[[\phi]]^{s[t->T]} = 1$

The logic is sorted, with variables of the form x_1..., (individual variables), t_1, ..., (time variables), and w_1, ... (world variables). Given Kaplan's Prohibition Against Monsters, context-dependency can be treated entirely in the meta-language. Indexicals such as 'I', 'now', 'actually' are given the following lexical entries, where c_A, c_T and c_w refer to the agent, time and world of c respectively (see Appendix II, A. for details):

(22) a. $[[I]]^{s, c} = c_A$
b. $[[now]]^{s, c} = c_T$
c. $[[actually]]^{s, c} = c_W$

The Logical Form of 'I actually exist now' is represented as follows:

(23) $[[\text{exist(I, now, actually)}]]^{c, s} = 1$ iff $<c_A, c_T, c_W> \in \boldsymbol{I}$(exist) which is true for every context c.

By contrast, 'I necessarily exist now' involves universal quantification over possible worlds, and is not true in every context:

(24) $[[\text{Nec } w_i \text{ exist}(I, \text{now}, w_i)]]^{c, s}=1$ iff for all $w \in W$, $[[\text{exist}(I, \text{now}, w_i)]]^{c, s[wi->w]}=1$, iff for all $w \in W <c_A, c_T, w> \in I(\text{exist})$, which does not have to be true.

3.2.2 Restatement of Kaplan's Theory

This is essentially enough to re-state Kaplan's theory within an extensional system. All we still need is a semantic rule for 'say', which is provided below:

(25) $[[\text{SAY}_{<\alpha', \beta', \gamma'>} \gamma\phi]]^{c, s}=1$ iff there exists a Character k in $K(<[[\alpha']]^{c, s}, [[\beta']]^{c, s}, [[\gamma']]^{c, s}>)$ [=the set of Characters asserted by α' at time β' in world γ'] such that
(i) $k(<[[\alpha']]^{c, s}, [[\beta']]^{c, s}, [[\gamma']]^{c, s}>)= \lambda w[[\phi]]^{c, s[\gamma->w]}$
(ii) $[[\alpha']]^{c, s}$ asserts k at $[[\beta']]^{c, s}$ in $[[\gamma']]^{c, s}$

A simple logic for this extensional version of Kaplan's system is presented in Appendix II A. For future reference, we now develop a more complete treatment of pronouns, tense and mood, which will be crucial for our final theory.

3.2.3 The treatment of pronouns

In the case of 'I', the reference of the pronoun is unambiguously determined by a given coordinate of the context (in this sense 'I' is a pure indexical). But such is not always the case. If there are several hearers, 'you' (in the singular) may be used in the same sentence to refer to different individuals. In such cases a demonstration typically completes the sentence:

(26) You [pointing] are elected, but you [pointing] are not.

If a 3rd person pronoun is used, anyone can be referred to who is not known to be either the speaker or the hearer of the current speech act:

(27) I've already met him [pointing] and him [pointing], though not her [pointing].

Clearly, in both types of examples the reference of the pronouns cannot be recovered from a given coordinate of the context, at least if the context is simply construed as a tuple of the form <author, (hearer), time of utterance, world of utterance>. Still, the context does *constrain* (although it does not ambiguously *determine*) the reference of these pronouns.

Two questions should be resolved:
(a) How should we handle pronouns whose reference cannot be recovered from a coordinate of the context?
(b) How should we analyze the constraints on reference that are introduced by context-sensitive features in these cases?

(a) In his 'Advice on Modal Logic', Dana Scott suggested that pronouns used demonstratively should be treated as free variables, whose value is given by an assignment function. Several contemporary researchers (e.g. Larson & Segal 1995, following Burge 1974 and Weinstein 1974; see also Higginbotham 2000) have observed that this isn't quite enough, since the theory must specify that the assignment function in question correctly captures the referential intentions of the speaker. This may be done by simply requiring that the original assignment function be properly determined by the context. Thus if the speaker meant to refer to John by 'he_2', s(2) should be John.

We may stipulate that the assignment function with respect to which a sentence is initially evaluated must be 'determined' by the context, in the sense that it properly represents the referential intensions of the speaker at the time and in the world of his utterance:

(28) Compatibility Condition
s is an assignment determined by c only if s properly represents the referential intentions of c_A at c_T in c_W

(b) How should the constraints on the reference of 'you' or 'I' be analyzed? As it turns out, almost nothing needs to be added beyond what is already necessary for the treatment of other grammatical features, such as gender. Cooper 1983 suggested that gender features contribute a presupposition that constrains the value of a pronoun under the relevant assignments. The same mechanism carries over to indexical features such as 1st and 2nd person.

Consider the case of gender features first. Within the present framework, Cooper's suggestion would be that 'She is clever' (used demonstratively) should receive the following analysis (I leave out time and world variables, and use in a'. a notation where presuppositions are written between curly brackets, as in Appendix II. C):

(29) a. She$_i$ is elected
a'. *LF:* be-elected (x_i{+feminine(x_i)})
b. *Assertion:* s(x_i) is clever
c. *Presupposition:* s(x_i) is female
d. *Compatibility:* s is determined by the context of utterance c.

The advantage of Cooper's system is that it extends straightforwardly to bound pronouns, and predicts interesting facts of presupposition projection. Using a standard rule of presupposition projection in universally quantified structures, Cooper's system (also applied in Heim & Kratzer 1998) explains why in the following sentence all the directors in the domain of discourse must be women:

(30) a. [Every director]$_i$ [t_i likes her$_i$ mother].
b. *Assertion:* For all x such that [[director]]^s(x)=1, [[t_i likes her$_i$ mother]]$^{s[i->x]}$=1
c. *Presupposition:* For all x such that [[director]]^s(x)=1, x is female.
d. *Compatibility:* s is determined by the context of utterance c.

Each assignment that satisfies the restrictor must satisfy the presuppositions of the nuclear scope, which directly derives the result[20].

Let us now extend this system to person features. A simple sentence with two occurrences of 'you' is represented as follows, where {+hearer*(x_i)} renders the presupposition that x_i is a hearer of the speech act (again, I disregard time and modality; a complete treatment is offered in Appendix II, C.):

(31) a. You$_i$ are elected, but you$_m$ are not.
a'. *LF:* elected(x_i{+hearer*(x_i)}) $\wedge$ $\neg$elected(x_m{+hearer*(x_m)})
b. *Assertion:* [[(b)]]$^{s,\,c}$=1 iff s(x_i) is elected and s(x_m) is not elected, which may both be true if s(x_i)$\neq$ s(x_m)
c. *Presupposition:* s(x_i), s(x_m) are hearers of the actual speech act
d. *Compatibility:* s is determined by the context of utterance c.

On this view there is no reason indexical pronouns couldn't be used as bound variables, since all the 1st or 2nd person features do is constrain the value of the variables under particular assignments. As expected, such a case does in fact arise, as was observed in Heim 1991 (with different examples, which are discussed later in this paper). Consider the following plural sentence:

(32) You (all) respect your wives

The most plausible reading is distributive, and thus crucially involves a bound possessive pronoun. The presupposition of the 2nd person pronoun does not lead to ungrammaticality, since the variable

[20] See Heim 1988 and Kadmon 2001, chapter 10, for an analysis of presupposition projection in quantified structures. Note that Appendix II C. should be extended to account for 'Every director likes her mother'. This is because for simplicity I have made quantification over individuals unrestricted. The extension to restricted quantification is relatively straightforward.

does in fact ranger over hearers.[21] Similarly it can be shown that a 1st person pronoun can be used as a bound variable. The examples in Heim 1991 involved the focus particle 'only' ('Only I did my homework', which has a reading which entails that Peter didn't do *his* homework). Ellipsis can also be used to make the point: 'I did my homework. Peter did too' has a reading on which Peter did *his* homework. Given standard assumptions concerning ellipsis, this suggests that the first conjunct has a bound variable reading as well (this reading is analyzed in detail in Appendix II C).[22]

Finally, let us consider in greater detail how gender and indexical features apply to plural pronouns. The preceding remarks can be extended to this case by (i) treating plural individuals as sums of singular individuals, and (ii) generalizing the constraints introduced above to the plural case. I start with the French pronominal system, which has the advantage of displaying more morphological distinctions in the plural than its English counterpart. First, consider 'ils', the masculine plural pronoun. It may denote any sum of individuals *as long as* one of its members is non-female. In the simplest case, where 'ils' is used deictically, this gives rise to the following conditions (where X_i ranges over sums of individuals):

(33) a. Ils_i sont élus.
They$_i$-M. Pl. are elected
a'. *LF:* be-elected (X_i{-feminine(X_i)})
b. *Assertion:* the members of s(X_i) are elected
c. *Presupposition:* one member of s(X_i) is non-female.
d. *Compatibility:* s is determined by the context of utterance c.

The same treatment can be applied when 'ils' has split antecedents, except that a sum of singular individuals will now appear explicitly in the Logical Form ('eux' is the strong form of 'ils'):

(34) a. Chaque garçon a donné à chaque fille une photo d'eux
[Each boy]$_i$ has given to [each girl]$_k$ a picture of them$_{i+k}$-M. Pl.
a'. *LF:* [$\forall x_i$: boy(x_i)][$\forall x_k$: girl(x_k)] (gave-a-picture-of (x_i, x_k, x_i+x_k {-feminine(x_i+x_k)})
b. *Assertion:* for each boy b and each girl g, b gave g a picture of the plural individual b✚g.
c. *Presupposition:* for each boy b and for each girl g, one member of s[x_i->b, x_k->g](x_i+x_k), i.e.

---

21 The representation is presumably something like the following (using the notation of Heim & Kratzer 1998):
(i) a. You (all) [λX [∀x: x∈ X] x respect your$_x$'s wife]
b. [[[λX [∀x: x∈ X] x respect your$_x$'s wife]]]$^{s, c}$ = λΞ: Ξ is a group and [[[∀x: x∈ X] x respect your$_x$'s wife]]$^{s[X->\Xi], c}$ is defined . [[[∀x: x∈ X] x respect your$_x$'s wife]]]$^{s[X->\Xi], c}$
We then apply to the definedness condition standard rules of presupposition projection and obtain the following result :
(ii) λΞ: Ξ is a group and every ξ in Ξ is a hearer of c. [[[∀x: x∈ X] x respect your$_x$'s wife]]]$^{s[X->\Xi], c}$
By functional application, this can be applied to the plural subject 'you' without yielding a presupposition failure.
The same point could have been made with simpler examples by considering cases of VP-ellipsis, which can be used to determine whether a pronoun is bound or not:
(iii) –You$_i$ λt$_m$ [t$_m$ like your$_m$ parents]
- You$_k$ do too ~~λt$_m$ [t$_m$ like your$_m$ parents]~~
Note that the above example shows (i) that the 2nd person possessive pronoun can be interpreted as a bound variable, and (ii) that *in the antecedent* the bound pronoun does satisfy the requirement that it should range only over hearers. On Heim & Kratzer's implementation, the function in (iii) is partial, with the value:
(iv) λx: x is a hearer. [t$_m$ like your$_m$ parents]$^{s[m->x]}$
Function application can then be applied without difficulty, since the argument 'you$_i$' does in fact refer to a hearer of the speech act. See Appendix II C. for a different implementation of the bound variable reading (without lambdas).

22 Two questions should be distinguished in the analysis of the bound variable reading of 'I$_x$ [t$_x$ did my$_x$ homework], and Peter did too'. (i) Are the 1st person features of 'my' interpreted in the antecedent? The present theory claims that they may be, because the rule of presupposition projection is automatically satisfied anyway (since 'my$_x$' only has to be evaluated under an assignment on which 'x' denotes the speaker) (ii) Are the 1st person features of 'my' interpreted after ellipsis resolution in the consequent? If ellipsis is resolved by copying, the answer must be 'no': 'Peter$_x$ [t$_x$ did my$_x$ homework]' would immediately result in a presupposition failure (note that exactly the same facts hold of bound 'she'). See paragraph 7.1 for further discussion.

one member of b✚g, is non-female.
d. *Compatibility:* s is determined by the context of utterance c.

Once these conditions are in place for gender, essentially nothing needs to be added for indexical features. 'We' is formally analogous to 'ils' except that the condition is now that the plural individual it denotes should have the *speaker* as one of its members:

(35) a. We are elected
a'. *LF:* be-elected (X_i{+author*(X_i)})
b. *Assertion:* the members of s(X_i) are elected
c. *Presupposition:* one member of s(X_i) is the speaker of the context c.
d. *Compatibility:* s is determined by the context of utterance c.

Following the same logic, cases of split antecedents for 'we' can be handled as well, as in the following (simplified) example:

(36) a. Each of my former wives remembers our fights
a'. *LF:* [$\forall x_i$: former-wife(x_i)](remembers-the-fights-of(x_i, x_i+x_k {+author*(x_i+x_k)}))
b. *Assertion:* for each of my former wives f, f remembers the fights of f✚s(x_k).
c. *Presupposition:* for each of my former wives f, one member of the plural individual s[x_i ->p](x_i+x_k), i.e. one member of p✚s(x_k), is the speaker of the context c.
d. *Compatibility:* s is determined by the context of utterance c.

At first sight it might appear that examples such as this require the *context* to be manipulated in such a way that 'us' may denote different plural individuals. But on the present analysis this is not required, since only quantification over *individuals* is involved, just as was the case in the analysis of 'ils'. This is important because it has sometimes been claimed that more complex versions of this example require some kind of 'context shift'. Although I do claim that context shift is needed in some cases, I deny that this is one of them[23].

3.2.4 Extension to tense and mood

This system can be extended to tense and mood. For our purposes it will suffice to assume that the present tense and the indicative each have a 'point' reading: '+present*' is a (context-dependent) predicate, which is true only of the time of utterance; and '+indicative^1*' is true only of the world of utterance (we will later add a second reading for the indicative, '+indicative^2*', which is necessary to handle indicative conditionals).

(37) a. For every time variable t, $[[t \{+present^*(t)\}]]^{s, c}$ is defined only if s(t) is c_T. If so, $[[t\{+present^*(t)\}]]^{s, c}$ =s(t).
b. For every world variable w, $[[w \{+indicative^{1}*(w)\}]]^{s, c}$ is defined only if s(w) is c_W. If so, $[[w \{+indicative^{1}*(w)\}]]^{s, c}$ =s(w).

We do preserve the result that 'I exist' is true whenever it can be uttered felicitously[24]; and that 'It is necessary that I exist' can be false, at least if we stipulate that no indicative mood appears on the verb embedded under 'it is necessary that':

---

[23] (ia) below is due to Partee 1989 and (ib) was suggested by an anonymous reviewer. Both examples are a more complex version of (36) in that they involve donkey anaphora. Any theory of donkey anaphora could presumably be added to the present analysis to handle these cases (in fact, if a DRT-style theory is used, nothing needs to be added since pronouns can be given their standard treatment; things are more complicated on an E-type theory).
(i) a. John often comes over for Sunday brunch. Whenever someone else comes over too, we (all) end up playing trios. (Otherwise we play duets).
b. Every time I get together with a linguist, we argue.

[24] In Appendix II C the indicative [written there as *+indicative^1*(w_i, c*)*] is treated as a shiftable indexical. But the results of this section are preserved.

(38) a. 'I exist' is true whenever it is uttered felicitously.
b. exist(x_i{+author*(x_i)}, t_k{+present*(t_k)}, w_l{+indicative[1]*(w_l)})
b. $[[(b)]]^{c, s}$=# iff $s(x_i)\neq c_A$ or $s(t_k)\neq c_T$ or $s(w_l)\neq c_W$. If ≠#, $[[(b)]]^{c, s}$=1 iff the agent c_A exists at the time c_T in the world c_W, which is true by stipulation.

(39) a. 'It is necessary that I exist' may be uttered felicitously and be false.
b.Nec w_l exist(x_i{+author*(x_i)}, t_k{+present*(t_k)}, w_l)
c. $[[(b)]]^{c, s}$=# iff $s(x_i)\neq c_A$ or $s(t_k)\neq c_T$. If ≠#, $[[(b)]]^{c, s}$=1 iff for *all* w∈ W, the agent c_A exists at the time c_T in the world w, which is unlikely to be the case.

4 Refutation of the Standard Theory

I now come to the refutation of Kaplan's theory. First, I review results that show that Kaplan's theory of indirect discourse is not sufficiently fine-grained to model indirect discourse in natural language. *Semantically*, it is shown that the expressive power of quantification over contexts rather than just quantification over possible worlds is needed. And since each context determines a single world (=the world of that context), while the converse is not true, quantification over contexts will of course allow for a finer semantics than mere quantification over worlds. Second, I adduce new data ('new' in the sense that they haven't been much discussed in the semantic literature, although they have been known for a long time in the typological literature) that suggest that, *morpho-syntactically*, quantification over contexts is also needed since some indexicals appear to be evaluated with respect to the context of the reported rather than of the actual speech act. Together, these observations will build a strong case against Kaplan's theory of indirect discourse, and in favor of the monster-based alternative that I lay out in Section 5.

4.1 Semantic Refutation: some attitude reports can convey indexical distinctions

In a nutshell, the purely semantic refutation of Kaplan's theory proceeds by showing that two thoughts or discourses that express the same Content but different Characters can sometimes be semantically distinguished in a report. This contradicts Kaplan's empirical prediction, which is that only Contents and never Characters should be of any linguistic relevance for indirect discourse.

4.1.1 Attitudes De Se vs. Reports of Attitudes De Se

Two problems should be distinguished at the outset. One is the problem of the 'essential indexical', or of 'attitudes De Se', as it has been discussed by Perry, Lewis and many others. This is a problem about attitudes, not attitude reports. It can be stated either in terms of thought, or in terms of the cognitive significance of sentences. In the first case, one will ask: how fine-grained are thoughts? Are there essential indexicals in thought? In the second case, the question becomes: how fine-grained are discourses? Are there essential indexicals in language? On either construal the argument is the same, and the answer is positive: there are essential indexicals in language and in thought. This, of course, is just a more general version of our earlier conclusion that the diagonal of a Character rather than a Content accounts for the cognitive significance of sentences. There are many ways to make this point. I'll just cite Perry and Lewis on this matter:

"An amnesiac, Rudolf Lingens, is lost in the Stanford library. He reads a number of things in the library, including a biography of himself, and a detailed account of the library in which he is lost... He still won't know who he is, and where he is, no matter how much knowledge he piles up, until that moment when he is ready to say, "*This* place is aisle five, floor six, of Main Library, Stanford. *I* am Rudolf Lingens." [Perry 1977]

Lewis comments:

"It seems that the Stanford library has plenty of books, but no helpful little maps with a dot marked "location of this map." Book learning will help Lingens locate himself in logical space. (...) But none of this, by itself, can guarantee

that he knows where in the world he is. He needs to locate himself not only in logical space but also in ordinary space". [Lewis 1979 p. 138]

From these observations, however, *nothing follows concerning the representation of attitude reports in natural language*. There is no reason to assume that attitude reports should be as fine-grained as attitudes, and superficially it seems that precisely when indexicality is concerned they are *not* as fine-grained (both 'My pants are on fire' and 'His pants are on fire' can be reported with: '$K._i$ believes that his_i pants are on fire'.) It was in fact a great virtue of Kaplan's analysis to show that the two questions are independent: thoughts or direct discourses are as fine-grained as Characters, Kaplan argued, but still a report can only be as fine-grained as a content. This has some apparently counterintuitive consequences, but on second thought they are rather well-motivated from a linguistic standpoint, as was observed above ('... when you wonder, 'Is that me?', it is correct to report you as having wondered whether you are yourself.' (Kaplan 1989, p. 554)).

4.1.2 Castañeda's 'he' and 'he himself'*

Starting from the observation that a report often *fails* to convey indexical distinctions, Castañeda 1968 devised an artificial pronoun, 'he*', to be used solely in a report to indicate that the world 'I' had been used in the original discourse. If the original thought was in the 1st person ('My pants are on fire'), the report will involve 'he*': 'Kaplan believed that his* pants were on fire'. If the thought didn't involve 'I' ('His pants are on fire'), the report won't contain 'he*': '$Kaplan_i$ thought that his_i pants were on fire'.

It was probably assumed even in Castañeda's time that 'he*' had some natural language counterparts, for instance the expression 'he himself'. Furthermore, Anscombe 1975 suggested that the 'indirect reflexives' of Ancient Greek and Latin are in fact the precise counterpart of 'he*' - a point which is not entirely correct since, unlike 'he*', they apparently occurred outside of attitude reports (this point is discussed in some detail in Clements 1975). But a detailed empirical study of these matters came only later.

4.1.3 Clements 1975 and Ewe logophoric pronouns

The remarkable discovery was that 'he*' does in fact exist in natural language. The first great study of this phenomenon was conducted by Clements 1975[25]. Summarizing cross-linguistic data, Clements characterized logophoric pronouns as elements that satisfy conditions (i)-(iii) (Clements 1975 p. 171):

"(i) logophoric pronouns are restricted to *reportive contexts* transmitting the words or thought of an individual or individuals other than the speaker or narrator";
(ii) the antecedent does not occur in the same reportive context as the logophoric pronoun;
(iii) the antecedent designates the individual or individuals whose words or thoughts are transmitted in the reportive context in which the logophoric pronoun occurs."

Clements's characterization is remarkably similar to Castañeda's definition of 'he*' - a point which is all the more striking since Clements (a noted phonologist) apparently did not know of Castañeda's work when he wrote his study[26].

---

[25] A similar topic was investigated in Hagège 1974, which gave 'logophoric' pronouns their name (literally, these are 'pronouns that carry discourse'). Since Clements's study is considerably more detailed and displays a high level of analytical clarity, it is the one I discuss in what follows.

[26] For comparison, here is the definition of 'he*" in Castañeda 1968 (I have adapted expressions between brackets):
"... the attribution of self-knowledge is made by means of the third-person pronoun 'he (himself)' to be abbreviated 'he*', which has here the following characteristics:
(i) it does not express an indexical reference made by the speaker;
(ii) it appears in *oratio obliqua;*
(iii) it has an antecedent, namely [the agent of the report], to which it refers back;
(iv) its antecedent is outside the *oratio obliqua* containing 'he*';

The hypothesis that logophoric pronouns are simply quasi-indexicals was not made by Clements, presumably because he was not aware of the literature on quasi-indexicals and attitudes De Se. However the connection between the two issues did not escape later researchers, among others Heim 1991 and Pan 1997. But although Clements's description of Ewe logophoric pronouns is highly reminiscent of 'he*', he did not provide the crucial semantic judgments that would have proven conclusively that they are actually one and the same thing. As was mentioned above, 'he*' differs from a non-starred 'he' by its distribution, since the former occurs only in attitude reports, while the latter can occur everywhere. But it also differs in its interpretation: whenever it is used, it attributes to the holder/agent of the attitude an indexical reference to him- or herself. If Kaplan thought: 'My pants are on fire', the report will be: 'Kaplan$_i$ thought that his$_i$* pants were on fire', but if his belief was of the form: 'His pants are on fire', the report will not include a star: 'Kaplan$_i$ thought that his$_i$ pants were on fire'.

To my knowledge no detailed study of the crucial semantic facts has been conducted for Ewe. However Kusumoto 1998 cites relevant data from another language with logophoric pronouns, Bafut. Here is the example she provides in order to show that a De Re (non-De Se) reading can *not* be rendered with a logophoric pronoun (her example (15)):

(40) Situation (Kaplan 1977): John is looking at a mirror from a distance and sees a man in the mirror. He notices that the man's pants are on fire. In fact, the man he sees in the mirror is John himself, but he doesn't realize it.

a. John believes that his pants are on fire

b. John wà?àtə mə *yu/à ká khi (Bafut; P.Tamanji,p.c to Kusumoto)
John thinks that self/he FUT burn

'John thinks that he is going to get burnt'

4.1.4 Morgan 1970 and Chierchia 1987: PRO

Decisive examples of unambiguously De Se attitude reports were provided for English by Morgan 1970 and by Chierchia 1987[27]. The crucial observation is that PRO, the unpronounced subject of an infinitive in a 'control structure', can only be used to report a De Se thought. Here is a slightly different version of these examples:

(41) Situation: John is so drunk that he has forgotten that he is a candidate in the election. He watches someone on TV and finds that that person is a terrific candidate, who should definitely be elected. Unbeknownst to John, the candidate he is watching on TV is John himself.

a. True: John hopes that he will be elected

b. False/#: John hopes PRO to be elected [Ok if the thought was: 'I should be elected']

As it appears, then, PRO in an attitude report can only be interpreted 'De Se': roughly, it can be used only in case 'I' was used in the original discourse. In fact the generalization is slightly less restrictive, since as is shown by the following example PRO can also be used in case the original discourse was in the 2nd person (cf. similar examples in Chierchia 1987):

(42) Situation: At a party, John is told that 'Mary' is being particularly obnoxious. He tells the person he is having a conversation with that 'Mary should leave'. But that person is none other than Mary herself.

a. True: John told Mary that she should leave

(v) 'he*' is used to attribute, so to speak, implicit indexical reference to [the agent of the report]; that is, if [the agent of the report] were to assert what, according to [the report], he knows, he would use the indicator 'I' where we, uttering [the report], have used 'he*'" [Castañeda 1968 pp. 440-441]

[27] Only Chierchia made a connection between his examples and 'attitudes De Se'; the term 'attitudes De Se' did not exist in Morgan's time, although Castañeda's work was of course available.

b. False/#: John told Mary to leave [Ok if the discourse was: 'Leave!']

Thus PRO can appear in the report only in case a 1st or a 2nd person pronoun was used (or could have been used[28]) in the original discourse.

4.2 Morphological confirmation: some indexicals can be shifted

4.2.1 The Possibility of an Intermediate Theory: Chierchia 1987 [after Lewis 1979]

While the facts presented above show unambiguously that Kaplan's theory of Indirect Discourse is empirically inadequate, they do not prove that a system with quantification over contexts is actually needed. In fact the most influential analysis of De Se reports, put forth by Chierchia 1987, does without context-shifting. The conceptual point is that it is perfectly possible to quantify simultaneously over each coordinate of a context instead of quantifying directly over the context itself. This is the solution represented in a. below. By contrast, the solution advocated here is represented in b. [for uniformity I use a notation in which the embedded context is bound by a lambda-operator]:

(43) a. John hopes $\lambda x\, \lambda t\, \lambda w$ [$\exists t'$: $t'>t$] x is elected president at t' in w
b. John hopes λc [$\exists t'$: $t'>time(c)$] x is elected president at t' in world(c)

There is a simple correspondence between the two theories: contexts can be identified to triples of the form <x, t, w>, where x is the author, t is the time and w is the world of c; there is a simple correspondence between '$\lambda x\, \lambda t\, \lambda w\, \varphi$' and '$\lambda<x, t, w>\varphi$'[29]; hence there is a simple correspondence between a. and b. Since Chierchia was interested in De Se with respect to individuals, he simply left out the abstraction over times in (43)a. Similarly, Abusch 1997, who was only interested in the interpretation of tense, could leave out the individual coordinate. But when one puts these theories together the result is as in (43)a, which corresponds directly to the solution in b.

Given the facts that were known at the time, the solution in a. was not unreasonable. Its effect was to allow attitude reports to distinguish between thoughts with the same content (contrary to Kaplan's own theory of indirect discourse), without granting that Kaplanian monsters were possible, since no context parameter or variable was manipulated.[30]

---

[28] This proviso is important to capture the following example, which was pointed out to me by J. Almog. Suppose that de Gaulle said (without being in any way drunk): 'De Gaulle should be elected.' One could still say truly that *de Gaulle hoped to be elected.* For even though de Gaulle sometimes chose to speak of himself in the 3rd person, he still didn't doubt in such situations that the speaker was de Gaulle himself. As the formal system below will make clear, what matters in the end is not which words the agent used, but rather which contexts were compatible with his attitude. In this sense the de Gaulle case is crucially different from the example in (41).

[29] The latter is the 'schönfinkelized' version of the former.

[30] It should be noted that Chierchia's analysis added several idiosyncrasies to a Lewisian account of attitude reports.

(i) First, Chierchia claimed that embedded clauses could *either* be treated as propositions or as properties. Thus a proposition-based system was preserved for De Re (non-De Se) readings, while a property-based account was used for De Se reports. This is thus a mixed theory, which Lewis 1979 wouldn't have endorsed. Lewis's point was precisely that one could and should do everything in terms of properties, since the latter are strictly more fine-grained than propositions. Why, then, have a mixed account?

(ii) Second, Chierchia claimed, in effect, that attitude reports are *not* special. He argued that De Se readings are simply a by-product of the syntactic appearance of property-denoting clauses, such as infinitives. He had no way to single out a natural class of 'attitude reports', something which was a virtue given the data he had access to. But as soon as one considers logophoric pronouns (and shifted indexicals), this becomes a liability, for these are items that can *only* occur in attitude reports.

These idiosyncracies are not shared by Heim 1991b, who in addition develops a system in which indexicals, like De Se pronouns, are bound by an operator in Comp. The crucial move is to posit that not just embedded clauses, but also matrix clauses can contain a λ-operator in their complementizer position. (Heim notes that there is a 'premonition' of this proposal in Koopman & Sportiche 1989). See also Kratzer 1997 and Heim 2002.

4.2.2 Why an intermediate theory is undesirable: shifted indexicals

This intermediate theory is undesirable, however. As it turns out, there are a number of indexicals cross-linguistically which *can* be shifted in attitude reports[31]. Adopting a solution based on quantification over contexts derives this fact straightforwardly. By contrast, the intermediate theory would have to make special stipulations for these cases. In what follows I go through a list of examples that show that some indexicals can be shifted in attitude reports.

(i) Consider the expressions 'two days ago' / 'in two days' in English and their French counterparts 'il y a deux jours' / 'dans deux jours' (for 'in two days', I have in mind the reading: 'at the end of a period of two days', not 'within a period of two days'. The ambiguity leads to additional complexities; the French expression is unambiguous in the desired way, which is why I use it). First, it is clear from their meanings that these expressions are context-dependent, since 'two days ago' and 'in two days' uttered on Monday do not refer to the same day as the same expressions uttered on Tuesday. In this respect they are analogous to the expressions 'the day before yesterday' and 'the day after tomorrow', which appear to have exactly the same meanings. But there is one difference, which I illustrate using 'two days ago'. While 'the day before yesterday' is a well-behaved Kaplanian indexical, which can only be evaluated with respect to the context of the actual speech act, such is not the case of 'two days ago':

(44) John has told me repeatedly over the years: 'I was still sick two days ago/the day before yesterday'
a. #John has told me repeatedly over the years that he was still sick the day before yesterday.
b. John has told me repeatedly over the years that he was still sick two days ago

In b. it is apparent that 'two days ago' can be evaluated with respect to the context of the *reported* speech act. But this is impossible in a. Note that 'two days ago' can *also* be interpreted with respect to the actual speech act, which is why it is crucial that attitude verbs be treated as quantifiers over contexts rather than as context-shifting modal operators. (If we had made the latter choice the original context of utterance would systematically become unavailable in the scope of an attitude verb, which predicts that an indexical should be forced to be shifted in such environments.)

One could suspect that 'two days ago' is not a *strict* indexical, and that its ability to be shifted simply shows that it is context-dependent by accident, so to speak. It could be pointed out that there are expressions such as 'before' (and 'earlier' or 'later'), which can be used indexically and can also be shifted, although it is unlikely that they are strict indexicals (that is, elements that are lexically specified as indexicals):

(45) John has told me repeatedly over the years: 'I have been sick before'
=> John has told me repeatedly over the years that had been sick before

Clearly 'before' uttered on Monday does not have the same semantic contribution as 'before' uttered on Tuesday. Furthermore, the preceding example shows that it can have a shifted reading when it appears under an attitude verb. Still, there is every reason to think that 'before' is indexical 'by accident'. Its other uses ('before Monday', etc.) suggest that it has an unpronounced temporal

[31] These cases should not be confused with 'Free Indirect Discourse', a literary style of considerable interest discussed, among others, by Banfield 1982, Reinhart 1983 and Doron 1991. In Free Indirect Discourse as well, indexicals can be evaluated with respect to a non-actual speech act, as in the following example:

(i) a. Tomorrow was Monday, Monday, the beginning of another school week! [Lawrence, *Women in Love,* p. 185, London, Heinemann 1971; cited in Banfield 1982 p. 98 (her (54)]
b. #He thought: 'Tomorrow was Monday, Monday, the beginning of another school week!'
c. #He thought that tomorrow was Monday, Monday, the beginning of another school week!

One major difference between the cases discussed in this paper and Free Indirect Discourse is that the latter cannot normally appear as part of a 'that'-clause, as shown in c. Furthermore, the indexicals that can be shifted in Free Indirect Discourse and in regular Indirect Discourse are not the same. 'Tomorrow' and 'yesterday' can be shifted in Free Indirect Discourse, but not in regular Indirect Discourse. The relation between shiftability in attitude reports and Free Indirect Discourse is discussed in Sharvit 2001.

argument: 'before t'. Presumably 't', a null temporal pronominal, may refer to any element made salient in the discourse[32]. Among these is the time of utterance, hence the appearance of indexicality. If this analysis is on the right track, we should expect that the temporal argument of 'before t' can in fact refer to any salient moment, whether the expression is embedded under an attitude operator or not. This is indeed what we find, as is demonstrated by the following cases of discourse anaphora:

(46) a. A week ago I met a man who was still sick two days before.
b. A week ago I met John. He was still sick two days before.

Thus 'two days before' cannot be taken to falsify Kaplan's theory of demonstratives, since it does not appear that the expression qualifies as a strict indexical to start with. But 'two days ago' is different in that it cannot so easily depend on any salient moment. Rather, 'two days ago' is easily evaluated only with respect to the context of some reported speech act. This explains why c. below is degraded (we will return to the fact that it is not entirely unacceptable):

(47) a. I met John last week. Two days before he was still sick.
b. *I met John last week. The day before yesterday he was still sick.
c. ?I met John last week. Two days ago he was still sick.

The facts might be even clearer with French 'dans deux jours' ('in two days'), as contrasted with 'après-demain' ('the day after tomorrow')[33]:

(48) a. J'ai rencontré Jean la semaine dernière. Deux jours plus tard il était malade.
I have met Jean the week last. Two days later he is dead.
'I met Jean a week ago. Two days later he was sick.'
b. *J'ai rencontré Jean il y a une semaine. Après-demain il était malade.
I have met Jean there is a week. The-day-after-tomorrow he was sick.
c. *J'ai rencontré Jean il y a une semaine. Dans deux jours il était malade.
I have met Jean there is a week. In two days he was sick.

As one would expect from its indexical nature, 'two days ago' (and 'dans deux jours') can only be read 'De Se' with respect to time. In other words, the time with respect to which 'two days ago' is evaluated when it is shifted can only be the 'now' of the agent (to use Abusch's terminology). Here again it contrasts minimally with 'two days earlier/before', which has the ability to be evaluated with respect to any salient antecedent. The crucial test is provided by the following example, in which the De Se and the De Re readings are truth-conditionally distinguished:

(49) Situation: 3 days ago (on Wednesday), John told me: 'According to the newspaper, it rained in L.A. on Monday'. John erroneously thought that he was speaking on Thursday (i.e. 3 days after Monday). He was in fact talking on Wednesday.
a. John said that it had rained two days earlier/before
b. #John said that it had rained two days ago

Why should (49)b not be true/acceptable? Simply because John's thought was of the form 'It rained on Monday', and certainly not of the form 'It rained two days ago/the day before yesterday', with a temporal indexical. For if John had been asked to state his belief in indexical terms he would have

[32] Exactly the same facts –and the same analysis- hold of the expressions 'earlier' and 'later' (here the comparative suggests that there is a concealed temporal argument, e.g. 'earlier ~~than t~~').

[33] These data can be replicated with English 'in two days', but one must be careful to block the unwanted reading of 'within two days'. This can be done by using 'in at least two days', which is uninformative and hence infelicitous on the reading: 'within a period of at least two days' (the reading 'at the end of a period of at least two days' is, by contrast, informative). Here are the relevant examples:

(i) a. I met John a week ago. At least two days later he was sick.
b. *I met John a week ago. The day after tomorrow he was sick.
c. <*>I met John a week ago. In at least two days he was sick.

said: 'It rained three days ago' (since he believed that the conversation was taking place on Thursday).

All these facts are entirely unexpected for a theory that prohibits indexicals from being shifted. By contrast, they follow straightforwardly on the monster-based theory advocated here. 'Two days ago' is specified as taking a context variable as argument; in other words, the expression is well-formed only in case it appears in a logical form as 'two-days-ago(c)'. Since attitude verbs are analyzed as quantifiers over contexts, the only two representations of the preceding sentences are as follows:

(50) John said that it had rained two days ago

a. Past t_k $SAY_{<John,t_k,actually>}c_i$ rain(two-days-ago(c_i), world(c_i))

b. Past t_k $SAY_{<John,t_k,actually>}c_i$ rain(two-days-ago(c*), world(c_i))

c. $[[(a)]]^{c,s}$ =1 iff for some t∈ T before c_T, for all c' compatible with John's claim at t_i in c_W, $[[rain(two\text{-}days\text{-}ago(c_i), world(c_i))]]^{c,\, s[ci\text{–}>c']}$=1

Here *c** is a free variable whose value is contextually supplied, and which by convention must denote the context of the actual speech act c (this convention must be integrated to the 'Compatibility condition' between the initial assignment of evaluation s and the context of utterance c which was discussed in Section 3.2.3: it must always be the case that s(*c**)=c). The crucial point is that in a., in which 'two days ago' is shifted, the expression must take the embedded context as argument, which automatically yields a 'De Se', shifted reading, as in c. As for the unshifted reading, it is straightforwardly derived by b.

Do other cases of shifting occur? Not if the only way to introduce a context variable (other than *c**) in a representation is by quantifying over it, i.e. by introducing an attitude operator. However the marginal acceptability of (47)c suggests that it might be possible to leave a context variable free if the discourse situation provides it with a salient value. There appear to be special constraints on this mechanism, as is shown by the fact that (48)c can be made to have roughly the status of (47)c, but only if the tense of the second sentence is modified:

(51) ?J'ai rencontré Jean il y a une semaine. Dans deux jours (*après-demain) il serait malade.
*I have met Jean there is a week. In two days (*the-day-after-tomorrow) he would-be sick.*
'I met Jean a week ago. In two days he would be sick'

I conclude that under ill-understood conditions a context variable other than *c** may be left free in a Logical Form, which leads to some marginal cases of shifting of 'two days ago'/'in two days' outside of attitude reports[34].

Let us observe, finally, that the present theory predicts that in cases of iterated attitude reports a shiftable indexical could depend on *any attitude verb in whose scope it appears.* This is because attitude verbs are analyzed here as quantifiers over contexts rather than as modal operators. The following sentence suggests that the desired ambiguity is real:

(52) My brother has informed me repeatedly over the years that my mother had asked the night before where I had been two days ago.

Assuming that 'the night before' is evaluated with respect to my brother's speech act, there appear to be two readings for 'two days ago': it may be evaluated with respect to the time either of my mother's or of my brother's speech act, as is predicted[35].

(ii) The same analysis can be given for indexicals in languages other than English. Each time the argument should in principle have three steps: (a) show that, semantically, the element is context-dependent; (b) prove that the element is lexically specified as being context-dependent, i.e. that it

[34] Thanks to François Recanati and an anonymous reviewer for helpful comments and suggestions on this topic.
[35] Thanks to an anonymous reviewer for suggesting that this prediction be tested.

behaves like 'two days ago' rather than like 'earlier'; and finally, (c) show that the element in question can only be interpreted 'De Se' when it is shifted. I only provide arguments (a) and (b) because the judgments in (c) were not clear-cut for my informant.

In the following example the 1st person embedded pronoun is clearly evaluated with respect to the context of the reported rather than of the actual act[36]:

(53) Situation: John says: 'I am a hero' [D. Petros, p.c.]

ǰon	*ǰəgna*	*nə-ññ*	*yil -all*
John	*hero*	*be.PF-1SO*	*3M.say-AUX.3M*

"John says that he is a hero"

Typically one reacts to such examples by claiming that the embedded clause must be quoted. But in at least some examples this can be shown not to be the case. Consider the following sentence, which involves a shifted 2nd person pronoun:

(54) *mɨn amt' -a* *ɨnd-al-ə -ññ* *al-səmma-hu-mm* (Leslau 1995 p. 779)
what bring.IMPER-2M *COMP-say.PF-3M-1SO* *NEG-hear.PF-1S-NEG*
'I didn't hear what he told me to bring' (lit. I didn't hear that he said to me bring what)

If the embedded clause had been quoted, the original discourse should have been of the form: 'bring what!'. But as the translation shows, this is not the correct reading (in fact, such a direct discourse would presumably be meaningless). Rather, the report means that he told me 'Bring X!', and I didn't hear what X was. The fact that there is an indirect question shows that the embedded clause is not quoted. Despite this, the embedded 2nd person pronoun can be evaluated with respect to the context of the reported speech act[37]. One could explain away these examples by claiming that only part of the embedded clause is quoted - in our case, only the embedded 2nd person morpheme. But crucially this is *not* a possibility that is normally open in cases of quotation. Thus in English 'I' or 'you' cannot normally be quoted in indirect discourse - otherwise 'Mary didn't hear what John told her 'you' should bring' should mean 'Mary didn't hear what John told her *she* (=Mary) should bring', contrary to fact.

With this background in mind, we can ensure that it is *only* in attitude reports that Amharic 1st and 2nd person pronouns can be shifted. That this is the case is suggested by the following contrasts, where Amharic behaves exactly like English, for the simple reason that no attitude verb is involved, and thus no context-shifting is possible (Degif Petros, p.c.):

(55) a. [My$_k$ brother]$_i$ found a girl that he$_{i,\ *k}$ likes
b. [My$_k$ brother]$_i$ found a girl that I$_{*i,\ k}$ like

a'. *wəndɨm - e*	*yəmm -i -wəd -at -ɨn*	*lɨǰǰ*	*agəññə*
brother-POSS.1S	REL-3M-love.IMP-3FO-ACC	girl	find.PF.3M

'My brother found a girl he likes/loves'[38]

b'. *wəndɨm - e*	*yəmm -ɨ -wəd -at -ɨn*	*lɨǰǰ*	*agəññə*
brother-POSS.1S	REL-1-love.IMP-3FO-ACC	girl	find.PF.3M

'My brother found a girl I like/love'

---

[36] Thanks to Degif Petros, Mengistu Amberber, Delombera Negga and Makonnen Argaw for help with the data and the transcriptions.

[37] As the gloss indicates, the embedded verb is in the imperative. This is surprising since in other languages (e.g. English) the imperative indicates that an order is given by the speaker of the *actual* speech act rather than by the speaker of a reported speech act. It might be that the typological differences we observed with indexicals also exist with imperatives (as suggested by U. Larsen, p.c.).

[38] As pointed out by M. Amberber and M. Argaw, the word *lɨǰǰ* (pronounced *lɨǰ* by M. Argaw) can in other contexts mean "child" or "boy". "*set lɨǰǰ*" (lit. female child) could be used to disambiguate.

It should be noted that similar facts have been noted for Engenni in Thomas 1978; for Aghem in Hyman 1979; and for Navajo in Hale & Platero 1998 and Speas 1999[39].

Exactly the same form of argument can be applied to the case of the Russian present tense. In (56)a the tense of the embedded clause can be evaluated with respect to the context of the reported speech act; but the use of a 3rd person pronoun to refer to the author of the reported speech act (=Petja) shows that the embedded clause is not quoted (Petja would have used 'I' to refer to himself):

(56) a. petja$_i$ skazal, c#to on$_i$ plac#et [Russian]
Petja$_i$ said that he$_i$ is-crying
'Petja said that he was crying [at the time of his utterance]'
b. petja$_i$ vstretil c#eloveka, kotoryj plac#et [Russian]
Petja met person, who is-crying
'Petja met a person who is crying / cries'
NOT: 'Petja met a person who was crying [at the time of the meeting]'
[Similar examples in Kondrashov & Kondrashova 1999. See also Kusumoto 1998]

Crucially, shifting is possible *only* in attitude reports. As shown in b., shifting is impossible in a relative clause. Furthermore, it can be shown that the generalization is semantic rather than syntactic in nature. Other complement clauses behave differently from the complement of 'say' in case they appear under verbs that are not attitude operators:

(57) a. c#asto sluc#alos', c#to mis#a plakal / *plac#et (Janssen 1996)
often happened, that Misha cried / is-crying
b. It often happened that Misha cried / *is crying

In other words, Russian behaves just like English when a clause is embedded under the verb 'happen'. From the present perspective, this is entirely predicted on the assumption that the Russian present tense is a shiftable indexical. Being an indexical, it must take as argument a context variable. But 'happen', which is not an attitude verb, fails to introduce a context variable, and therefore the Russian present tense cannot be shifted when it appears under such a verb.[40]

---

[39] Thomas 1978 uses the term 'semi-indirect discourse' to refer to such examples, which are difficult to classify as involving either 'direct' or 'indirect' discourse, at least on a traditional analysis of indexicals. As she observes, in (ia) the embedded 2nd person pronoun suggests that direct discourse is used, while the embedded 3rd person pronoun argues for indirect discourse. The same point is made by Hyman 1979 concerning (ib) (the 2nd person pronoun is evaluated with respect to the context of the reported speech act, but the presence of the 3rd person logophoric pronoun indicates that the embedded clause is not quoted).

(i) a. ò wei ga ... bhú tou eì ka òkì nàà ìwó wu zà
2-sub 3(-ref)-obj 3-ref-sub 2-obj
he say [sp you should-take him seq he and you should-die stay]
'He said, "Look after me, and I will die with you"' or
'He said that she should look after him, and he would die with her' [Engenni, Kwa; Thomas 1978]
b. wìzÍn 'vÚ ndzÈ à wÍn NÍ'á **é** Ngé 'lÍghá **wò**
[woman that] said to him [that LOG-3 much like you]
'the woman said to him that she liked him a lot', or
'the woman said to him "I like you"' [Aghem, Bantu; Hyman 1979]

Note also that in Amharic shifting seems to occur only after an all-purpose attitude verb which originally means 'say'. Why this is so is unclear to me, and should definitely be investigated.

[40] Abusch 1997 gave interesting examples that suggest (in our terms) that the Russian facts can be replicated within a fragment of English (Abusch did not mention Russian, however). As Abusch observed, the point of evaluation of modals such as 'might' and 'ought' can be shifted, but only in 'intensional' environments. This is a slightly different generalization from the one we are suggesting here, since for Abusch intensional operators also include 'will', which is taken to shift the time coordinate (the 'now parameter') with respect to which tense is evaluated. Here are Abusch's original examples:

5 A Monster-Based Account

In 'Where Monsters Dwell', Israel & Perry observed that Kaplan's prohibition against monsters was nothing more than a stipulation, and that one would expect to find monsters in the domain of attitude reports. While they did not know examples of monstrous (shifted) indexicals, their conjecture was exactly correct, and essentially for the right reasons[41]. I follow their lead, and give two versions of a theory that makes crucial use of quantification over contexts. The first version, which is developed in greater detail in Schlenker 1999, relies on heavily indexed logical forms. The second version seeks to eliminate some of the syntactic stipulations of the first theory, and to provide a semantic account which is more elegant and makes different, and hopefully better, predictions.

5.1 Three Components

In both versions, the analysis has three components.

5.1.1 Quantification Over Contexts

Since we justified earlier the use of context variables, we can directly define the syntax and semantics of (simplified) attitude verbs, construed as quantifiers over contexts. For simplicity, we concentrate on the verb 'say'. As can be seen, the analysis is entirely similar to the quantificational version of a modal analysis, except that quantification over possible worlds is replaced with quantification over contexts[42]:

(58) a. Syntax
If ϕ is a well-formed formula, if c_i is a context variable, and if α', β', γ' are respectively an individual, a time and a world term, then $\ulcorner SAY_{<\alpha', \beta', \gamma'>}\ c_i \phi \urcorner$ is a well-formed formula.

(i) a. John married a woman who might become rich (Abusch's (40))
ok John's bride could become rich at the time of utterance
* John's bride could become rich at the time of John's wedding
b. John believed that his bride might become rich (Abusch's (41))
ok John believed that, at the time of this thought, his bride could become rich

(Note that the contrast between (ia) and (ib) involves not just a shift in temporal perspective, but also in epistemic perspective: in (ia) 'might' is evaluated with respect to the speaker's Common Ground, while in (ib) it is evaluated with respect to John's epistemic alternatives).

Future research should determine which version of the generalization is correct. Different predictions are made for the following examples:

(ii) a. John married a woman who wanted to have children who might become rich.
b. (#)John married a woman who would later have children who might become rich.

On both theories (iia) is predicted to be grammatical. (iib) is predicted by Abusch to be fine on the reading: '... who would later, at t', have children who, at t', might become rich'. I predict that this reading should be unavailable. Preliminary data suggest that this is indeed the case. (Heim 1994b (fn. 22) claims that Abusch's prediction is borne out. However her example involves embedding, not under 'would', but under 'will', which in any event allows tenses –e.g. the present- to be 'shifted'. An account of this fact that does not hinge on Abusch's assumption concerning the 'now' parameter is offered in Schlenker 1999).

[41] "But now we should remind ourselves of the following facts about actual utterances and the contexts in which they are produced: one might not know who the agent of c is; one might not know when the time of c is; one might not know what the place of c is; one might not know what the world of c is. (...) So given a type for an utterance, that is, given a sentence Φ, other contexts for Φ are epistemic alternatives. (...) In conformity with Kaplan's restriction, and supposing for simplicity that indices proper -circumstances- are just worlds, all he would allow us is this:

c, w |= KΦ iff $\forall$w': w' R_K w $\Rightarrow$ c, w' |= Φ

But to capture the facts about ignorance, what we need is rather more like this:

c, w |= KΦ iff $\forall$c', w': <c, w> R_K <c', w'> $\Rightarrow$ c', w' |= Φ; where c'=<c'_a, c'_t, c'_p, c'_w>.

This, of course, is monstrous." (Perry & Israel 1996, pp. 314-315)

[42] Special provisions must be made for quantifying in, but these are no different in this theory from what they are elsewhere.

b. Semantics

$[[SAY_{<\alpha', \beta', \gamma'>}\ c_i\ \phi]]^{c,s}=1$ iff for all c' compatible with the claim made by $[[\alpha']]^{c,s}$ at time $[[\beta']]^{c,s}$ in world $[[\gamma']]^{c,s}$: $[[\phi]]^{c,s[ci \rightarrow c']}=1$

(Although for 'say' the contexts compatible with the agent's claim are context of *speech*, for verbs such as 'hope' the relevant entities are contexts of *thought*[43].)

There is an analytical choice to be made at this point. We may either take contexts as primitive, or construct them out of entities such as 'speaker', 'time of utterance', 'world of utterance'. Minimally one needs an individual, a time and a world to define a context. Two individuals are sometimes more convenient, to represent both the speaker and the hearer of the speech act. This choice can be important when morpho-syntactic rules must be defined which narrowly depend on the way in which context variables are represented. If contexts are primitive, context variables such as 'c' can be used directly. Otherwise contexts will have to be represented as tuples of the form <x, (y), t, w>, and attitude verbs will in effect quantify simultaneously over all three or four coordinates.

5.1.2 Filtering Mechanism: Shiftable vs. Non-Shiftable Indexicals. Logophoricity.

Once this semantics is in place, we will have to address Kaplan's original motivation for prohibiting quantification over contexts. The problem is that although some indexicals can be shifted, as our theory predicts, others cannot be. These expressions are well-behaved from a Kaplanian standpoint, but they are unexpected on the present theory since in principle every indexical should have the ability to be shifted under an attitude verb. Some lexical stipulations are needed to insure that English 'I' or 'yesterday', unlike Amharic 'I', or 'two days ago', should never be shifted. Under the 1st theory to be presented below, the filtering is done in the morpho-syntax; a distinction is enforced between 'matrix' and 'embedded' context variables, and certain sequences of symbols (e.g. 'tomorrow(c)', where 'c' is an embedded context variable) are stipulated to be ill-formed. In the 2nd theory to be discussed, the filtering is effected in the semantics: some lexical specification prevents 'I' from having the actual speaker as its semantic value[44]. As we will see, the predictions are slightly different, in a rather interesting way.

On either theory, two natural classes are defined: roughly, that of elements that can *only* be evaluated with respect to the actual speech act (English 'I'), and the class of elements that can be evaluated with respect to any context whatsoever (Amharic 'I'). A natural morphological implementation of this distinction is to posit a binary feature ±actual speech act. Both English 'I' and Amharic 'I' are indexical expressions (they have, if you will, a feature '+contextual'). But in addition, English 'I' is specified as '+actual speech act', while Amharic 'I' is simply underspecified for ±actual speech act. This makes a typological prediction: there should be elements that are +contextual and -actual speech act. How would such expressions behave? Since they are +contextual expressions, they must either depend on the actual speech act, or else be embedded under an attitude verb. But since they are -actual speech act expressions, the first option is precluded. Thus they should be elements that may only appear in the scope of an attitude operator, and must thus be semantically dependent on the context variable introduced by such operators. These expressions

---

[43] It may in the end be necessary to countenance improper contexts as well, for instance (a) contexts whose author does not exist at the time and world of the context, or (b) contexts without an author coordinate. This would seem to be important to handle (i) below, which was suggested by Irene Heim and Jay Rifkin (p.c.):

(i) I would prefer PRO not to exist

Because the embedded verb involves PRO, the sentence should be interpreted De Se, and thus involve (in our terms) quantification over contexts. The desired reading should presumably be something like: 'in every (possibly improper) context c compatible with what I would prefer, the author of c does not exist in the world of c at the time of c (or: c has no author coordinate)'. See Predelli 1998 for a defense of improper contexts.

[44] The necessary feature might be similar to Recanati 1993's 'REF' ["A term is (type)-referential if and only if its linguistic meaning includes a feature, call it 'REF', by virtue of which it indicates that the truth-condition (or, more generally, satisfaction-condition) of the utterance where it occurs is singular" (p. 17)].

exist: they are the logophoric pronouns whose behavior was described earlier; they are also the natural-language counterparts of Castañeda's 'he*'. But instead of being stipulated, their very existence is now predicted by the filtering mechanism which is needed independently to capture the distinction between English and Amharic 'I'.

It should be noted that logophoric forms are not limited to the person domain. There is in German a tense/mood called 'Konjunktiv I' (1st subjunctive[45]), which (i) can (almost) solely occur under attitude verbs, and which furthermore (ii) must be *interpreted* as dependent on an attitude verb, as shown by the following facts:

(59)	Der Peter	meint,	a.	es	sei später, als es tatsächlich	ist
	the Peter	*thinks*		*it*	*be later than it really*	*is*
			b.	es	ist später, als es tatsächlich	ist
				it	*is later than it really*	*is*
			c.	* es	sei später, als es tatsächlich	sei
				it	*be later than it really*	*be*
			d.	* es	is später, als es tatsächlich	sei
				it	*be later than it really*	*be*

Although the German indicative (glossed as 'is') may be interpreted either inside or outside the scope of an attitude verb, the 2nd possibility is precluded for the 'Konjunktiv I' (glossed as 'be'). This directly accounts for the ungrammaticality of c. and d. The natural suggestion is that the German Konjunktiv I is nothing but a temporal or modal version of the logophoric pronouns that are found in Ewe: it is an indexical expression which can only depend on the context of a reported speech act.

Another way to make the same point is to observe that, in those cases where the Konjunktiv I is not syntactically embedded, it forces a reading of modal subordination, as in the following examples:

(60) a. Er sagte, sie sei schön. Sie habe grüne Augen. [Jäger 1971] *He said she be pretty. She have green eyes.*
b. Er sagte, sie sei schön. Sie hat grüne Augen. [Jäger 1971]
He said, she be pretty. She has green eyes

As Jäger 1971 observes, in a., which involves a Konjunktiv I form of 'have', the 2nd sentence must be read from the standpoint of the attitude holder, so that it is interpreted as: 'He says/thinks that she has green eyes'. No such reading is forced in b. This effect is rather strikingly similar to one found in Ewe under similar circumstances:

(61) "The antecedent of the logophoric pronoun in Ewe need not occur in the same sentence, but may occur several sentences earlier. In such cases (...) the subsequent sentences of the discourse will continue to present the events described by the narrator from the point of view of the same individual or individuals. " [Clements 1975 p. 170]

The final typology, then, will be the following ('+C' is short for '+contextual'):

(62)

Amharic 'I'	Russian Present	+C	±actual
English 'I'	English Present	+C	+actual
Logophoric pronouns	Logophoric tense/mood	+C	-actual

It is a non-trivial result to *derive* the existence of logophoric pronouns from the feature mechanism needed to get the difference between English 'I' and Amharic 'I'. Still, there is a major stipulation

[45] The name 'subjunctive' is partly misleading because the Konjunktiv I displays very different properties from the Romance subjunctive. For instance, (i) the Romance subjunctive is not restricted to attitude reports; (ii) it cannot appear embedded under 'say', unlike the Konjunktiv I.

left in this analysis, since the feature ±actual does not appear to be related to anything else in the grammar. The 2nd theory to be presented below will seek to eliminate this stipulation, in the sense that it will reduce the feature ±actual to another feature for which there is independent motivation.

5.1.3 Agreement Mechanism

A final element of the analysis must be introduced. In a nutshell, every analysis that allows an embedded pronoun (or tense) to be dependent on a context variable must explain why the same pronoun can still inherit the syntactic features that appear on the matrix subject. The problem appears clearly in the following examples:

(63) a. John hopes PRO to buy himself a car.
b. Mary hopes PRO to buy herself a car.
c. You hope PRO to buy yourself a car.
d. I hope PRO to buy myself a car.

In each case the reflexive pronoun, and therefore its local binder PRO, bears the features of the matrix subject. But in this kind of logical forms, *there is no relation between PRO and the matrix subject* (the following is a simpler example, from Appendix II B.)*:*

(64) $\text{HOPE}_{<\text{John,now,actually}>} c_i$ be-elected(agent(c_i), time(c_i), world(c_i))

This problem is not new, and it has appeared under different forms in other researchers' work, e.g. Chierchia and Heim (the latter stated an explicit agreement rule to account for the percolation of features in similar examples in Heim 1994a). In this case as well, two theories will be presented below. The first theory will claim that the agreement must essentially be stipulated, in that in such cases the relevant features simply fail to be interpreted. The stipulation is not entirely idle, since it is supposed to account for cases of semantically empty agreement in the person but also in the tense domain ('Sequence of Person', 'Sequence of Tense'). The second theory, by contrast, will seek to do without this stipulation altogether, which will involve a slight revision of the semantic analysis.

5.2 Rich LFs (Schlenker 1999)

I review briefly the theory presented in Schlenker 1999, simplifying somewhat the notation (this is possible in part because I ignore the issue of Free Indirect Discourse, which Schlenker 1999 tried to analyze within the same analytical framework as shifted indexicals; since the two phenomena display very different properties, this made for a rather baroque system).

In this system it is crucial that a distinguished context variable, *c**, be set aside to refer to the actual context. The formation rules are as before, as is the semantics for attitude verbs. Two points are special about this theory: the analysis of the filtering mechanism, and the account of agreement.

5.2.1 Filtering Mechanism

The necessary stipulation is that some indexicals can only take *c** as argument, while others may take any context variable, and still others can only take non-*c** variables as arguments. In other words, the feature ±actual speech act is reanalyzed as ±c*. The lexical entries in English, Amharic and Ewe are as follows:

(65) a. English 'I': +indexical, +c*
b. Amharic 'I': +indexical [underspecified]
c. Ewe logophoric pronoun: +indexical, -c*

Thus it is in the logical syntax that certain representations are blocked that would make English 'I' a shifted indexicals:

(66) $\text{Say}_{<x1,\ x2,\ t1,\ w1>}$ c hero(I(c), time(c), world(c))

This sentence would be perfectly interpretable, but the lexical entry of 'I' does not allow it to take a 'c' argument. All it can take is a 'c*' argument, and since by hypothesis the value of the latter is

always the actual speech act, the effect is that English 'I' is a Kaplanian indexical, one that can never be shifted. By contrast, an equivalent sentence in Amharic is well-formed, because Amharic 'I' does *not* have such a restriction:

(67) a. $SAY_{<x1,\ x2,\ t1,\ w1>}$ c hero(I_{Amh}(c), time(c), world(c))
b. [[$SAY_{<x1,\ x2,\ t1,\ w1>}$ c I_{Amh}(c) be a hero at time(c) in world(c)]]$^{s,\ C^*}$=1
if and only if for every context C compatible with what s(x_1) tells s(x_2) at s(t_1) in s(w_1), the author of C is a hearo at the time of C in the world of C.

5.2.2 Agreement Mechanism

With English and Amharic 1st person pronouns taken care of, it must still be explained how PRO may inherit features from the matrix subject even though it is bound by the embedded context variable.

Consider the following example in a., with the simplified LF in b. (note that standard locality conditions on the reflexive 'himself' force it to be bound by PRO):

(68) a. John hopes PRO_k to buy $himself_k$/*$herself_k$ a car
b. $HOPE_{<x_1,\ t_1,\ w_1>}$ c_i buy-for(agent(c_i), agent(c_i), time(c_i) in world(c_i))

The first thing to observe is that the features on PRO do not appear to be interpreted. Let us again analyze gender features in presuppositional terms. The contribution of a masculine feature is thus that the value of the variable under any assignment may only be a male individual:

(69) $[[he_i]]^s$ is defined only if s(i) is a male individual. If defined, $[[he_i]]^s$=s(i).

In a simple case, this appears to make correct predictions. Using the rule of presupposition projection for universally quantified statements, we determine that every context c compatible with the agent's hope must satisfy the presupposition of the nuclear scope 'buy-for(agent(c_i), agent(c_i), time(c_i), world(c_i))', where c is the value of the variable $c_{i..}$ If 'agent(c_i)' is decorated with a masculine feature, this will yield the result that all contexts compatible with the person's hope have a male agent. In simple cases this does not appear to be a problem; John is male, and all contexts compatible with his hope are likely to be contexts whose author is male as well. But now consider the following variation of the previous example:

(70) John (a transsexual) hopes to become a woman, and he hopes PRO to buy himself/*herself a car

The discourse has been set up in such a way that all contexts compatible with John's hope are now contexts with a female author. The presuppositional analysis of gender features predicts that 'himself' should yield a presupposition failure in this case; by contrast, 'herself' should be used unproblematically. But exactly the opposite is the case.

Entirely parallel examples can be constructed in the tense domain (see also Heim 1994b for discussion):

(71) In 1999 John believed that he was already in year 2005, and that Clinton was still president

By stipulation all contexts compatible with John's thought are contexts whose time coordinate must follow the time of the actual speech act. But the presuppositional analysis predicts that the preceding sentence should then be incoherent:

(72) $Bel_{<John,\ 1999,\ w^*>}$c (time(c)=2005 $\wedge$ president (Clinton, time(c), world(c))

Due to the past tense, the presupposition on time(c) is that its value should precede the time of utterance, i.e. 2002. This should make the sentence a presupposition failure, contrary to fact. Thus tense features do not appear to be interpreted when a variable is semantically dependent on an embedded context.

In the tense literature the solution would be to posit a Sequence of Tense rule, which allows tense features to remain uninterpreted in certain environments. In fact quite independently of our somewhat exotic example a Sequence of Tense rule is known to be necessary (even when none of the agents is confused about the time):

(73) Abusch's version of Kamp & Rohrer's French examples
John decided a week ago that in ten days at breakfast he would say to his mother that they were having their last meal together.

The tense of the most deeply embedded clause ('were having') refers to a moment which is *after* any other time mentioned in the discourse. Thus any theory of tense that requires the past tense features to be interpreted in this case is bound to go wrong. The quasi-consensus in the literature is that there are Sequence of Tense rules, which leave certain tense features uninterpreted when they appear under attitude verbs that bear themselves the same features.

It is no trivial matter to implement the agreement rule which is needed both in the case of tense and in the case of person. The solution adopted in Schlenker 1999 was to decompose contexts into tuples of variables of the form $<x^1, y^2, t^0, w^0>$, where by convention the superscripts 0, 1, 2 indicate that a variable is a coordinate of a context (note that a variable could have both a superscript and a subscript. As was mentioned above, any number of context variables could be needed in a given sentence; hence within this new notational framework any number of author, hearer, time or world coordinates might appear in a representation). It was further stipulated that embedded contexts inherit their features from the matrix clause: the author coordinate inherits the features of the speaker/thinker argument, the hearer coordinate (if any) inherits the features of the hearer argument, and similarly for the time and world arguments. The agreement was enforced as a part of the construction rule for attitude operators:

(74) Syntax
Let x', y', t' and w' be any variables or constants, and let α, β, γ, δ be the features borne by these elements [we abbreviate this as: $x'(\alpha)$, $y'(\beta)$, $t'(\gamma)$, $w'(\delta)$]
If ϕ is a well-formed formula, then so is '$\text{Say}_{<x'(\alpha),\ y(\beta),\ t(\gamma),\ w(\delta)>}$ $<x^1(\alpha), y^2(\beta), t^0(\gamma), w^0(\delta)>$ ϕ'

The details of this system are rather complex, and will not be considered further here. Suffice it to say that that this system has to be further supplemented with rules that stipulate that features inherited through agreement are not interpreted. The system is admittedly stipulative, but part of the motivation was to show that exactly the same stipulations could account for Sequence of Person and for Sequence of Tense. (A further argument was given in Schlenker 1999 to show that in other cases - involving 'only' – some features had to be disregarded anyway. However this argument was based on a dubious generalization, which is discussed later in this paper).

While the system developed so far is roughly adequate, it is also stipulative. This is because both the filtering device and the agreement mechanism are postulated in an entirely *ad hoc* fashion. There are in fact two distinct problems:

-In the case of the filtering mechanism, the problem is not that *some* device be postulated. This is presumably necessary on anybody's theory, for the simple reason that there appear to be minimal differences between English 'I' and Amharic 'I', or between 'yesterday' and 'two days ago'. The problem, however, is that the feature we postulated to encode these lexical stipulations is not motivated *independently*. Ideally, one would like to couch these stipulations in a theoretical vocabulary which is independently needed for other parts of the grammar; this does not seem to be the case here.

-In the case of the agreement mechanism, it is not clear that *any* device is needed at all. As we will see shortly, it is because of our particular implementation of the semantics that we had to stipulate that features should not be interpreted in certain environments. We discuss the prospects for eliminating this stipulation in what follows.

6 An Alternative to the Filtering Mechanism: a semantic derivation of ±actual

The alternative to be presented here is semantically based. This new theory suggests that the apparent contrast between matrix and embedded context variables is in fact the by-product of a semantic difference which is independently motivated. In a nutshell, the semantic distinction between 'he' and 'I', which is needed on everyone's theory, will be seen to be sufficient to derive the distinction between English 'I', Amharic 'I', and the Ewe and Gokana logophoric markers. The new theory makes one surprising new prediction, which appear to be borne out.

6.1 Person

6.1.1 A simple version of the system

Consider first how the distinction between 'he' and 'I' can be captured in the present system. Let us assume that 1st and 2nd person features contribute a presupposition that the value of the relevant variable is speaker/hearer of the actual speech act:

(75) a. $[[x\ \{+\text{author}^*(x)\}]]^{s,c}$ is defined only if s(x) is the author of c. If so, $[[x\ \{+\text{author}^*(x)\}]]^{s,c} = s(x)$
b. $[[x\ \{+\text{hearer}^*(x)\}]]^{s,c}$ is defined only if s(x) is a hearer of c. If so, $[[x\ \{+\text{hearer}^*(x)\}]]^{s,c} = s(x)$

It is tempting to extend this system to 3rd person pronouns by positing a negative presupposition: a 3rd person feature presupposes that the variable it appears on denotes neither the author nor a hearer of the context c. We may decompose this into two independent statements of the following form, with 3rd person=[-1st person, -2nd person]:

(76) a. $[[x\ \{-\text{author}^*(x)\}]]^{s,c}$ is defined only if s(x) is not the author of c. If so, $[[x\ \{-\text{author}^*(x)\}]]^{s,c} = s(x)$
b. $[[x\ \{-\text{hearer}^*(x)\}]]^{s,c}$ is defined only if s(x) is not a hearer of c. If so, $[[x\ \{-\text{hearer}^*(x)\}]]^{s,c} = s(x)$

Although this system will have to be amended, it suffices to provide a derivation of the existence of logophoric pronouns. Following our previous practice, we posit that indexicals of the Amharic variety are simply specified for a feature +author, defined as a relation between an individual and a context, and written as: +author(x, c). As with other grammatical features, we assume that this is interpreted as a presupposition on the value of a variable:

(77) a. $[[x\ \{+\text{author}(x, c_i)\}]]^{s,c}$ is defined only if s(x) is the author of $s(c_i)$. If so, $[[x\ \{+\text{author}(x, c_i)\}]]^{s,c} = s(x)$
b. $[[x\ \{+\text{hearer}(x, c_i)\}]]^{s,c}$ is defined only if s(x) is a hearer of $s(c_i)$. If so, $[[x\ \{+\text{hearer}(x, c_i)\}]]^{s,c} = s(x)$

As before, we translate each pronoun with a variable, followed by a presupposition that represents the contribution of the grammatical features of the pronoun. But we also want to say that in Amharic a shifted 1st person pronoun can be dependent on the context variable introduced by an attitude operator. If we stopped here this would not be possible, for we would have representations such as the following, in which the variable 'x' appears free:

(78) a. John_i says that I_i am a hero (Amharic)
b. $\text{SAY}_{<\text{John}, \ldots, \ldots>} c_i$ hero(x {+author(x, c_i)}, ..., ...)

Since 'x' is free, its value cannot be affected by the attitude operator, contrary to what we want. To solve this technical problem I introduce a mechanism of 'definite closure' by which an ι-operator can

close off the free variable (the presupposition that follows the variable is then interpreted as the restriction of the definite description; this is defined in Appendix II, C)[46]:

(79) $SAY_{<John,\ ...,\ ...>}c_i$ hero(ιx{+author(x, c_i)}, ..., ...)

If we don't add any further stipulations, the only way we may define an element which is an author of a speech act, but not the author of the actual speech act, is to assume that the relevant pronoun is an indexical (i.e. a shiftable indexical, since this is now the only category there is) which is specified as -author*. The relevant combination is thus the conjunction: (+author(x, c_i) ∧ -author*(x)), which (when preceded by the ι-operator) is interpreted as follows:

(80) $[[\iota x \{+author(x, c_i) \wedge -author^*(x)\}]]^{c,s}$ is undefined unless there is exactly one individual that is both (i) author of $s(c_i)$ and (ii) not the author of c.

This would appear to be a good definition of logophoric pronouns, since these can only be used to refer to the author of a reported speech- or thought- act. On the present account, this means that logophoric pronouns may refer to the author of any context, except if that person is also the author of the actual speech act (without the latter requirement the pronouns in question would simply be shiftable 1st person pronouns of the Amharic variety). This makes an interesting prediction, which did not follow from the preceding system, based on richly annotated Logical Forms. Previously there was nothing to prevent a 1st person embedded verb from taking logophoric marking, since a pronoun could bear logophoric marking if it was syntactically bound by a coordinate of an embedded context. But on the present theory a 1st person logophoric pronoun would be a contradiction in terms, for by definition a logophoric pronoun must have as its denotation something different from the author of the actual speech act. This is best seen if we add the semantic condition on a 1st person pronoun (of the English variety) to a logophoric pronoun. The result is:

(81) ιx {+author(x, c_i) ∧ -author*(x) ∧ +author*(x)}

But of course this yields a referential failure in all cases, because of the contradiction between -author*(x) and +author*(x)[47].

The prediction, then, is that (author-denoting) logophoric pronouns should never appear in the 1st person. Typologically this prediction appears to be borne out, as this pattern is exceedingly rare. Roncador 1988, who provides a survey of the literature, notes only one apparent exception to this general absence: the logophoric marker which appears on the verb in Gokana can in principle be applied to all persons[48]. Roncador relies on the description of Hyman & Comrie 1981. The latter point out, however, that although logophoric marking is morphologically possible in the first person, it is 'dispreferred', so that (82)b is degraded by comparison with (82)a:

(82) Gokana (Hyman & Comrie's (11))

a. Ok:	mm~	kO	mm~	dO~
	I	*said*	*I*	*fell*
b. 'Dispreferred to [a]':	mm~	kO	mm~	dO~-E~
	I	*said*	*I*	*fell-LOG*

The fact that (82)b appears to be relatively degraded is all the more striking since in the other persons logophoric marking is *preferred* whenever it is possible; the opposite pattern is thus found in the first

[46] Thanks to A. von Stechow for pointing out an important mistake in a previous version of this work.

[47] See Heim 2002 for an attempt to derive this prediction within a system in which indexical features are not interpreted, but only provide an indication on the nature of the binders.

[48] He also discusses another potential case (Ngbaka), but concludes that we do not have clear evidence that logophoric marking really occurs in the first person (Roncador 1988 p. 279).

person[49]. Although far more fieldwork is needed to confirm these data, I take them to be preliminary evidence in favor of the proposed theory[50].

6.1.2 Refinements

The presuppositions we have stated for -1st and -2nd person turn out to be too strong. Consider a situation in which the identity of a person seen in a mirror is in doubt. I could very well utter the following:

(83) He looks like me... Mmh, he must be me, in fact he *is* me![51]

Clearly we do not want to make such a statement a presupposition failure, as is predicted on the current theory. Thus it does not have to be presupposed that a 3rd person pronoun does not denote the speaker or hearer of the actual speech act. But it also wouldn't do to say that 'he' carries no presupposition whatsoever, for in a normal context 'he' cannot refer to the speaker (nor to the hearer). The generalization appears to be that 'he' can be used *whenever it is not presupposed that the variable denotes the speaker or hearer of the actual speech act.* In effect, then, 'he' is used as a semantic default. It would be an important enterprise to determine in more detail what it means to presuppose that a variable denotes the speaker or hearer[52]. But for our purposes the following statement should suffice:

(84) a. Use -author* just in case +author* would have yielded a presupposition failure.
b. Use -hearer* just in case +hearer* would have yielded a presupposition failure.

These conditions have the further advantage of allowing a variable marked with a 3rd person feature to refer to the speaker under some (though not all) assignments. Consider for instance the following sentence:

(85) Every man (including me) likes his mother.

Certainly the bound variable in a. should be allowed to range, among others, over the speaker. Under the previous theory this would have been impossible, for a presupposition failure would have ensued. But this is no longer the case under the revised theory. It is clear that using a +author* feature on the variable x would have resulted in presupposition failure. Therefore the use of -author* was licensed, and similarly for -hearer*.

Of course in simpler cases we preserve our initial results: when there is no uncertainty about who is who, and when the variable is not bound by a quantifier, a 3rd person feature forces the variable to denote someone other than the actual speaker or hearer. And this suffices to define logophoric pronouns. In addition to being empirically more adequate, the revised theory also has the advantage of extending to an analysis of mood in German.

6.2 Extension to Mood

We take Stalnaker's semantics of conditionals as our starting point, in that we define for every context c a Common Ground CG(c), which is the set of worlds compatible with everything the

[49] Hyman & Comrie 1981 and Roncador 1988 give a functional explanation of this asymmetry.

[50] A similar prediction is made about hearer-denoting logophoric pronouns, which have been claimed to exist in Mupun, a Chadic language described by Frajzingier 1985 (a hearer-denoting logophoric pronoun refers to the hearer of a reported speech act; see also Frayzingier 1993). If the present theory is on the right track, hearer-denoting logophoric pronouns *should never appear in the 2nd person.* I do not know whether this prediction is borne out.

[51] If one prefers 'this is me' to 'he is me', one may consider negative statements, as suggested by an anonymous reviewer: 'He looks like me... but he isn't me'. The second sentence should not come out as trivially true, since the speaker asserts something non-trivial when he utters this in the mirror situation.

[52] This could be done along the following lines. Evaluate sentences with respect to a set G of assignments rather than with respect to a single assignment g. A pronoun bearing index i would be presupposed to refer to a female individual just in case for every assignment function g in G, g(i) is female. See Dekker 2000 for an analysis.

speech act participants presuppose. In line with our treatment of the pronominal case, we posit the following presuppositions for the indicative and subjunctive markings:

(86) a. $[[w \{+indicative^2*(w)\}]]^{s, c}$ is defined only if s(w) is in the Common Ground of c. If defined, $[[w \{+indicative^2*(w)\}]]^{s, c}$=s(w)
b. 'w {-indicative^2*(w)}' can be used only if marking 'w' with +indicative^2* *in the same Logical Form* would result in a presupposition failure.

This semantics is very close in spirit to Stalnaker 1975 and von Fintel 1997. In particular, non-indicative marking is taken to have a very weak semantic contribution, since it only means that indicative marking couldn't have been used. The difference, however, is that the present analysis is simply the application to mood of the analysis of person laid out in the preceding section.

In line with our treatment of logophoricity in the pronominal case, let us now see how we can analyze the Konjunktiv I in German. First, we define a feature $+indicative^1(w, c_i)$, which indicates that w is the world of c_i (this is just the 'shiftable' version of the feature $+indicative^1(w)$ defined in Section 3.2.4; in the Appendix we use $+indicative^1(w)$ throughout). By analogy with our analysis of logophoric pronouns, we posit the following feature combination for the German Konjunktiv I:

(87) $\iota w\{+indicative^1(w, c_i) \wedge -indicative^2*(w)\}$

As in the pronominal case, this revised theory derives predictions which were not made by the system based on richly annotated LFs. In a nutshell, whenever it is presupposed that s(w) is in the Common Ground, the use of -indicative^2*, and hence of the Konjunktiv I, should be precluded. This appears to be the case, as demonstrated by the following (fairly robust) generalization:

(88) The Konjunktiv I cannot be used in the 1st person present singular of 'believe'.

Consider the following asymmetries:

(89) a. *Ich glaube, daß Maria krank sei
I believe that Maria sick is-KONJ1
b. Ich glaubte, daß Maria krank sei
I believed that Maria sick is-KONJ1
'I believed that Maria was sick'
c. Peter glaubt, daß Maria krank sei
Peter believes that Maria sick is-KONJ1
'Peter believes that Maria is sick'
d. Peter glaubte, daß Maria krank sei
Peter believed that Maria sick is-KONJ1
'Peter believes that Maria is sick'

The Konjunktiv I can be used in every case except in the first person present. Why should that be? The answer becomes apparent when it is observed that the set of worlds compatible with the speaker's belief at the time of utterance is *always a subset of the Common Ground itself.* As a result, the presupposition projection rule for universally quantified statements entails that the indicative *can* be used. But due to the definition of -indicative^2*, if +indicative^2* can be used, it *must* be used. This derives the generalization (See Appendix II, C for an explicit derivation)[53].

[53] I leave two further points for future research. First, the same facts appear to hold for 'say' as for 'believe', as shown in the following:

(i) a. *Ich sage, daß Maria krank sei
I say that Maria sick is-KONJ1
b. Ich sagte, daß Maria krank sei
I said that Maria sick is-KONJ1
'I said that Maria was sick'
c. Peter sagt, daß Maria krank sei

7 Is there an alternative to the Agreement Mechanism?

Can we also get rid of our last stipulation, which was the agreement mechanism? The answer will only be partly positive. In the person case, we may apparently do without an agreement mechanism by introducing a slightly different semantics. But in the case of tense only part of the data will be covered in this way. The data that motivated the introduction of 'Sequence of Tense' rules will remain unaccounted for under the new system.

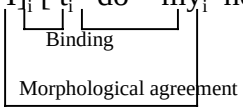
7.1 Reasons to get rid of the agreement mechanism

The agreement mechanism that was postulated in Section 5 had a major weakness: it was an outright stipulation. The only motivation given was that such a mechanism was needed in order to make the correct predictions; but this is just another way of saying that the theory was by itself unable to account for part of the data, and had therefore to be supplemented with a stipulation.

7.1.1 Attempt at motivating the stipulation [Kratzer 1998, Schlenker 1999]

There have been several attempts at motivating this stipulation. One mechanism was offered in Kratzer 1998, and criticized in Schlenker 1999. The latter suggested that the same kind of agreement rule was needed to account for cases involving 'only' (first brought up by Heim 1991) and for the case of attitude reports. As Heim observed, a 1st person pronoun can sometimes be used as a bound variable. Furthermore, in such cases the person features do not appear to be interpreted at all, for if they were they would presumably constrain the value of the pronoun (under an assignment) to be the speaker of the actual speech act. But such is not always the case:

(90) a. Only I do my homework [... therefore Peter doesn't do his]
b. [only $I]_i$ [t_i do my_i homework]



On the relevant reading 'my_i' is treated as a variable bound by the subject trace, and thus indirectly by the generalized quantifier 'only I'. If the 1st person feature were interpreted, it would yield a presupposition failure, since the bound variable 'my_i ' must range over non-speakers:

(91) For a referential NP, $[[[\text{only NP}]_i \phi]]^{s,\,c}$ =1 iff for d=$[[\text{NP}]^{s,\,c}$, $[[\phi]]^{s[i->d]\,c}$=1 and for all individual d'≠d, $[[\phi]]^{s[i->d'],\,c}$=0

As can be seen, if 'my_i' carries a presupposition that i denotes the speaker, the above semantics will yield a presupposition failure.

Peter says that Maria sick is-KONJ1
'Peter says that Maria is sick'

Second, it appears that the Konjunktiv I can never be used after factives, as shown in the following:

(ii) a. Peter weiß, daß Maria krank ist
Peter knows that Maria sick is
b. *Peter weiß, daß Maria krank sei
Peter knows that Maria sick be

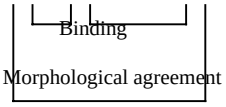
An appealing line of explanation would run as follows. Consider the semantics of 'know', analyzed as a quantifier over contexts:

(iii) $[[\text{KNOW}_{<x,\,t,\,w>}\ c_i\text{: } \phi]]^{s,\,c}$=1 iff every context c' compatible with what s(x) knows at s(t) in s(w) is such that: $[[\phi]]^{s[ci->c'],\,c}$=1.

This naive semantics seems to derive the result. Since 'know' is factive, a context can be compatible with what the agent knows only if that context is part of the Common Ground. As a consequence, indicative marking can be used in the embedded clause, and therefore it must be used. The problem, however, is that this naive semantics appears to be incorrect, among others because it fails to yield an entailment from *John knows that φ* to *John believes that φ*. The reason is that $\text{KNOW}_{<x,\,t,\,w>}$ only quantifies over contexts that lie in the Common Ground; whereas $\text{BELIEVE}_{<x,\,t,\,w>}$ has no such restriction, unless x's beliefs entail those of the speaker, which isn't in general true.

The problem is not unique to indexical features. Gender features on bound pronouns fail to be interpreted in exactly the same environment [Heim & Kratzer 1998]:

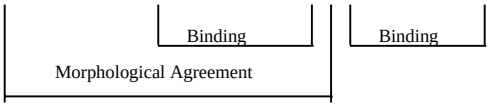
(92) a. Only she did her homework [... therefore Peter doesn't do his]
b.[only she]$_i$[t$_i$ did her$_i$ homework]



As in the preceding case, the semantics of 'only' implies that 'her$_i$' must be evaluated under assignments that do assign male individuals to the variable, which should yield a presupposition failure, contrary to fact.

These examples are somewhat similar to those that were discussed above with respect to De Se readings:

(93) John hopes λ<x$_i$, t, w> PRO$_i$ to buy himself$_i$ a car



In both cases the morphological rule has the same flavor: an element inherits features that remain uninterpreted from an element which is not its binder, but is syntactically close to its binder.

7.1.2 Why this motivation was dubious

Unfortunately it does not appear that the similarity goes any further than this. When the data are considered in greater detail, it appears that (apart from De Se reports) there are two kinds of environments in which features of bound pronouns may remain uninterpreted: (a) focus-sensitive constructions, and (b) ellipsis resolution. Attitude reports do not appear to form a natural class with either (a) or (b), and thus it is not obvious that a unification can be achieved[54].

As has often been noted, a treatment of 'only NP' as a generalized quantifier is rather implausible[55], and can advantageously be replaced with a focus-based analysis. For our purposes this has the added benefit of predicting other cases in which some feature may remain uninterpreted, as in the following examples:

(94) a. Only I did my homework [Therefore John didn't do his]
b. Even I did my homework [Therefore John also did his]
c. Even Mary did her homework [Therefore John also did his]
d. *I* did my homework. [Implicature: John didn't do his]
e. *Mary* did her homework. [Implicature: John didn't do his]

In a focus-based analysis (*à la* Rooth 1996) (94)a can be given the following treatment, where 'o' stands for 'ordinary value' and 'f' stands for 'focus value'. The important point is that the mechanism that computes the focus value must be allowed to disregard the 1st person feature of 'my$_x$', or else a presupposition failure will ensue[56]:

(95) a. 'Only' combining with a clause ϕ yields the assertion $\forall p\ [p \in [[\phi]]^{f, s, c} \ \& \ p(c_w)] \Rightarrow p=[[\phi]]^{o, s, c}$
b. The focus value $[[I_F\ (\lambda x\ t_x$ did my$_x$ homework $)]]^{f, s, c}$ is {[[I did my homework]], [[John did

[54] However see Heim 2002 for a recent attempt in that direction. Heim argues (*contra* the present view) that certain features must remain uninterpreted in a broader range of cases, for instance in 'Few men brought their children', where the plural features on the bound variable 'their' makes no semantic contribution. Heim attempts to analyze attitude verbs as 'verbal quantifiers' which, like 'few students', transmit certain features in the phonological component only.

[55] There are several reasons for this. (i) 'only' would be a determiner taking a referential expression rather than a predicate as its first argument; (ii) in other cases, in which it *would* take a predicate as an argument (e.g. 'Only students came'), it would have to be non-conservative.

[56] Rooth's intensional analysis would have to be slightly modified to be compatible with our extensional theory.

his homework]], [[Mary did her homework]], ...}, i.e.[57]
{[[t_x did my_x homework)]]$^{f, s[y->c_A], c}$, [[t_x did my_x homework)]]$^{f, s[y->John], c}$, [[t_x did my_x homework)]]$^{f, s[y->Mary], c}$ ...}

(96) a. ϕ=’Only I did my homework [Therefore John didn’t do his]’
b. LF: Only(C) [I_F [λx t_x did my_x homework]]
c. Presupposition: [[I [λx t_x did my_x homework]]]]$^{o, c, s}$=1
d. Assertion: ∀p [[p∈[[ϕ]]$^{f, c, s}$ & $p(c_w)$=1]⇒p=λw c_A did c_A’s homework in w]

Features that appear on bound pronouns may similarly be ignored in ellipsis constructions, at least on the assumption that ellipsis is resolved by copying part of the antecedent at LF:

(97) a. I did my homework. John did too (i.e. John did his homework)
b. Mary did her homework. John did too (i.e. John did his homework)

‘John did ~~my homework~~ too’ and ‘John did ~~her homework~~ too’ would obviously yield a presupposition failure unless ‘my’ and ‘her’ could somehow be ignored (syntacticians call this phenomenon ‘vehicle change’).

How should these facts be handled? Two theories seem plausible:
-Theory A posits that the mechanisms that compute (a) focus values and (b) ellipsis resolution may disregard features of bound pronouns. Elsewhere (in particular in sentences of the form NP λx ϕ) features must be interpreted.
-Theory B posits that features present at LF are always interpreted, but that in constructions of the form NP λx ϕ the variable x may inherit the features of NP in the morphological component only, so that these features are invisible at LF.

Theory A, which is a natural extension of the analysis presented here, posits that the features of ‘my’ are interpreted on the bound variable in ‘I λx [t_x did my_x homework]’ (as was pointed out in paragraph 3.2.3, this is perfectly compatible with standard rules of presupposition projection). Only in the computation of ‘John did too’ is the 1st person feature ignored, somewhat along the following lines[58]:

(98) (I like my mother and) John does too (bound variable reading)
a. Step 1 (Copying): John λx [t_x did my_x homework]
b. Step 2 (Vehicle change): John λx [t_x did x’s homework]

What about focus constructions in this theory? If the focus value is construed as a set of propositions, as in (99)b_1, we must stipulate that the mechanism that computes it may disregard certain features. If it is construed as a set of logical forms, as in (99)b_2, we may simply apply once again the mechanism of vehicle change:

(99) I_F did my homework (implicature: John didn’t do his)
a. Assertion: I λx [t_x did my_x homework]
b_1. Focus value, construed as a set of propositions:
A={ [[y λx [t_x did x’s homework]]]$^{c, s[y->d]}$(d): d an individual which is a contextually given alternative to the speaker}
b_2. Focus value, construed as a set of logical forms:
A={⌜d λx [t_x did x’s homework]⌝: d a denoting expression which is a contextually given

[57] This is a slight simplification. Rooth’s procedure would first yield: {[[y (λx t_x did my_x homework)]]$^{o, s[y->c_A], c}$, [[y (λx t_x did my_x homework)]]$^{o, s[y->John], c}$, [[y (λx t_x did my_x homework)]]$^{o, s[y->Mary], c}$, ...}.
[58] In Appendix II C (98)a would be written without λ as: John x_m{+author*(x_m)} do-the-homework-of(x_m{+author*(x_m)}, x_m{+author*(x_m)}) (I have further simplified the representation by omitting time and world variables).

alternative to 'I'}

d. Implicature: No element of A is both (i) strictly more informative than a. and (ii) true.

On Theory B, things are somewhat simpler. 'I did my homework' may appear at LF without indexicals features in the variables, yielding: I λx [t_x did x's homework]. No problem arises either in the computation of focus or in the resolution of ellipsis (note that although Theory B is compatible with the rest of our theory, it makes our earlier observation on presupposition projection in bound variable constructions unnecessary).

Neither Theory A nor Theory B extends naturally to De Se reports, and as a result the agreement mechanism posited there stands as an outright stipulation, which it would be desirable to eliminate (see Heim 2002 for a different view).

7.2 Alternative: De Re and Modes of Presentation

Let us now explore the possibility that, despite appearances, person features *are* always interpreted in attitude reports. First, we note that in De Re readings of definite descriptions the descriptive content is evaluated with respect to the actual world rather than with respect to the agent's belief worlds. The same observation applies to the presuppositions of gender features: 'Ralph believes of Ortcutt that he is a woman' does not attribute to Ralph the contradictory thought that someone is both male and female. Thus with De Re readings the gender features *can* in fact be interpreted, though not with respect to the agent's belief worlds/contexts. We can extend this solution to De Se readings by positing that a De Se reading is just a De Re reading where the mode of presentation of the *res* to the agent is linguistically specified in the report (the mode of presentation must be indexical)[59].

7.2.1 Quantifying In with Modes of Presentation

Let us first consider the issue of Quantifying In independently of the De Se problem. The initial observation, due to Quine 1956, was that both of the following sentences may be simultaneously true:

(100) a. Ralph believes, of Ortcutt, that he is a spy (*qua* the man in the brown hat)

b. Ralph believes, of Ortcutt, that he is not a spy (*qua* the man seen at the beach).

On the assumption that beliefs are closed under conjunction, a straightforward analysis would risk attributing irrationality to Ralph. For instance, in a possible worlds framework Ralph would have to believe that he lives in a world where (i) Ortcutt is a spy, and (ii) Ortcutt is not a spy. But there are no such worlds, and therefore Ralph would believe the empty proposition. This fails to distinguish irrational beliefs from cases of mistaken identity. Kaplan's solution in 'Quantifying In' was to reintroduce in his truth-conditions the mode of presentation under which Ralph held the relevant belief. Simplifying somewhat, Kaplan's analysis was the following: Ralph believes, of Ortcutt, that he is a spy just in case there is some acquaintance relation α between Ralph and Ortcutt, and Ralph believes (De Dicto) that whoever falls under α is a spy. This is represented more formally in the following:

(101) a. Ralph believes that Ortcutt is a spy

b. ∃α R(α, Ortcutt, Ralph) & Ralph believes ⌜α is a spy⌝

c. R(α, x, Ralph) ['α represents x to Ralph'] if and only if:

(i) α denotes x (ii) α is a name of x for Ralph and (iii) α is sufficiently vivid

Of course if we were to consider the De Re reading with a definite description (e.g. 'Ralph believes that the spy Ortcutt isn't a spy'), it is clear that the content of the description would be

[59] Abusch 1997 assumes that modes of presentation in attitude reports are left for the context to determine. She does not assume that they can be linguistically specified. See also Reinhart 1991.

evaluated with respect to the actual world rather than with respect to Ralph's belief worlds. The same result carries over to gender features. Given the notorious difficulty of giving a compositional treatment of Kaplan's suggestions, I simply stipulate that the 'final' logical form is not that in b., but rather that in c., where quantification over modes of presentation is explicitly represented – and where the variable x appears in an extensional position:

(102) a. Ralph believes (of Bernard J. Ortcutt) that ^{Ok}he /$^{\#}$she is woman.
b. BELIEVE$_{<\text{Ralph, ..., ...}>}$ c woman (x{^{Ok}-feminine(x) /$^{\#}$+feminine ...}, ..., ...)
c. [$\exists\alpha_{<c,\, e>}$: R(α, x {^{Ok}-feminine(x) /$^{\#}$+feminine ...}, Ralph,...)]
BELIEVE$_{<\text{Ralph, ..., ...}>}$ c woman(α(c), ..., ...)

Crucially, the presupposition that follows the variable 'x' does *not* have to apply to the mode of presentation α which is Ralph's description of Bernard J. Ortcutt (Ralph may well think, incorrectly, that Ortcutt is a woman; this won't affect the fact that 'x' refers to a male individual).

The conclusion is that for De Re attitudes no problem arises with gender features, since these are always interpreted outside the scope of the attitude operator. So why did we have a problem with De Se readings? This was because we assumed that these were *not* De Re beliefs, and had a representation in which the De Se pronoun was solely dependent on the embedded context (for simplicity the embedded predicate is analyzed as taking a simple individual argument; the reflexive is used only to show that PRO has masculine features):

(103) a. Ralph hopes (to be a woman, and) PRO to be worthy of ^{Ok}himself/$^{\#}$herself
b. *LF:* HOPE$_{<\text{Ralph, ..., ...}>}$c worthy-of-self($\iota$x[{+author(c, x) $\wedge$ ^{Ok}-feminine(x)/$^{\#}$+feminine(x)}, ..., ...)

If, as is plausible in a., every context c compatible with Ralph's hope is one whose agent c_A is a woman, then the feature +feminine should be grammatical, while –feminine should yield a presupposition failure. But exactly the opposite appears to hold.

There might be a way out, however. If we could treat De Se readings as a variety of De Re readings, as has been suggested repeatedly in the literature, we could get out of this problem. Just as in the De Re case considered above, the gender features that appear on the pronoun would then be evaluated outside the scope of the attitude operator, and the unwanted predictions would disappear. Of course since there exist De Se *readings* we have to ensure that something in the logical form differentiates a simple De Re reading from a De Re reading which is also De Se. Given that there is a quasi-consensus on the fact that a De Se reading entails the corresponding De Re reading[60], this can be effected by simply *adding in the logical form a predicate that specifies the nature of the acquaintance relation* (this partly follows a suggestion by B. Schein (p.c.)). In effect, then, this theory says that 'John hopes PRO to be elected' is true if and only if Smith hopes, of Smith, that he should be elected, under the description (the acquaintance relation): 'I'. This solution has a cost, however: we have to stipulate that part of the features of PRO (the gender features) can be 'exported' with the variable they appear on, while others (the shifted indexical features) must remain in the scope of the attitude operator to constrain the choice of the mode of presentation (in the type-theoretic notation, 'c' is the type of contexts):

(104) [$\exists\alpha_{<c,\, e>}$: R(α, x {^{Ok}-feminine(x) /$^{\#}$+feminine(x)...}, Ralph,...)]
HOPE$_{<\text{Ralph, ..., ...}>}$c worthy-of-self($\alpha$(c){+author(c, x)}, ..., ...)

If we are ready to split the grammatical features of PRO in this fashion, no problem will arise. There is now a direct dependency between PRO and its antecedent 'Ralph' (they are coindexed), so that the variable x must have the individual Ralph as a value. Since Ralph is male (even if he wants to become a woman), the feature +feminine is correctly excluded; only –feminine is allowed. Further, since +author(c, x) appears as a presupposition on the value of the *description* α(c), it makes the

[60] For a dissenting view, see Schlenker 1999 (Appendix II to Chapter 2).

existential quantification over α felicitous and true only in case Ralph has a hope of the form: 'I am worthy of myself', in the 1st person. This appears to be the correct result, with the benefit that all the grammatical features that appear on PRO are in fact interpreted semantically, contrary to our previous assumptions. The unsettling part, however, is that these features do not get interpreted *in the same place*.

7.2.2 Remaining Problems with Tense

There is another problem with this solution. Although we may have now solved the problem of interpreting gender features on pronouns, we have lost our account of Sequence of Tense. This is because, in the famous Kamp-Rohrer-Abusch example (repeated below), a De Re interpretation of the past tense will still not yield the correct result:

(105) Abusch's version of Kamp & Rohrer's French examples
John decided a week ago that in ten days at breakfast he would say to his mother that they were having their last meal together.

Even in case all the agents involved have perfect knowledge of the time, so that there is no distinction between De Se and De Re with respect to time, the past tense features on the most embedded tense cannot be interpreted, since this would imply that the time of the last meal is *before* the utterance time, contrary to fact. Thus the theory presented here does not extend to all temporal cases. By contrast, the goal in Schlenker 1999 was to show that the *same* rules can account for tense agreement and for person agreement. The latter is now accounted for without agreement rule at all, but the theory of tense is back where it was before this entire endeavor – with an unpleasant need for special stipulations. I conclude that an alternative to the agreement mechanism has yet to be fully worked out. By contrast, the alternative to the filtering mechanism discussed in Section 6 did appear to be a net improvement over the previous version of the theory[61].

8 Extensions

I now discuss very briefly two possible extensions of the present theory.

8.1 Vagueness[62]

In unpublished work, Yael Sharvit has suggested that standards of measurement for vague predicates should be treated as shiftable indexicals. An argument for doing so could be derived from the following contrast:

(106) a. #John's height is 1m50 and he is thus short, and rather unhappy. But if he were the same height but were tall, he would be much happier.
b. John's height is 1m50 and he is thus short. But although his mother knows his real height she thinks that he is tall.

The second sentence of (106)a is infelicitous because it would seem that there are no worlds (compatible with what is presupposed at that point) in which John is 1m50 and is tall: the first sentence has already established that 1m50 makes one short, not tall. This reasoning, however, can go through only on the assumption that the standard of tallness is the same across possible worlds. This, in turn, suggests that the standard is fixed by the context of speech rather than by a world of evaluation. But on the present theory this still leaves two possibilities open: the standard may be fixed by *any* context or only by the context of the actual speech act. (106)b suggests that the first possibility is the correct one: John's mother isn't mistaken about John's height, but rather about what

[61] U. Sauerland (p.c.) notes that a theory without an Agreement Mechanism may systematically fail for examples with plurals such as the following:
(i) We all sometimes believe that we're the only person in the world. See also Heim 2002.
I do not know how such an example could be treated in the theory of Section 7.2.

[62] Thanks to Yael Sharvit and to Chris Barker for discussions on this topic.

the standard of tallness is. This may be analyzed as follows, where the standard of tallness d_{tall} is treated as a shiftable indexical, which may thus take the context variable c_i as argument (*tall(John, d, time(c_i), world(c_i))* should be read as: *d is the maximal degree such that John is tall to that degree at time(c_i) in world(c_i)*]

(107) $\text{BELIEVE}_{<x^*, t^*, w^*>}\ c_i\ ([\imath d\text{: tall(John, d, time}(c_i)\text{, world}(c_i))] > d_{tall}(c_i))$

8.2 Expressives

Kaplan has recently tried to extend the logical insights of *Demonstratives* to the study of expressives (Kaplan 2001). Within the present system, his primary example, 'oops', can be analyzed as introducing a presupposition that the author of the actual speech (at the time and world of that speech act) has 'just observed a minor mishap':

(108) $[[\text{oops } \phi]]^{s,\,c}$ is undefined unless the agent of c has just observed a minor mishap at the time of c in the world of c. If defined, $[[\text{oops } \phi]]^{s,\,c}=[[\phi]]^{s,\,c}$

I wrote this rule in a Kaplanian fashion, with 'oops' depending only on the context of the actual speech act. The connection with the semantics of indexicals should be obvious. Unsurprisingly, Kaplan makes anew the observation that expressives, like indexicals, do not appear to be shiftable. But is this always true? Maybe not. While (109)a is definitely odd, and suggests that (as Kaplan claims) 'honky' introduces a presupposition that the speaker has a negative attitude towards Caucasians at the time and world of the speech act, (109)b sounds far more coherent:

(109) a. #I am not prejudiced against Caucasians. But if I were, you would be the worst honky I know.
b. I am not prejudiced against Caucasians. But John, who is, thinks/claims that you are the worst honky he knows.

The same point could be made with examples suggested by A. Kratzer, which involve (German equivalents of) the word 'really', used to indicate that the speaker asserts a sentence with particular force. Although it isn't clear how 'really' should be analyzed, it *is* clear that whatever attitude is ascribed to the speaker in (110)a is ascribed to the agent John in (110)b, which would suggest that 'really' is a shiftable expressive:

(110) a. Direct discourse: 'Mary is really pregnant'.
b. Indirect discourse: John said that Mary was really pregnant.

While this is all very preliminary –not least because the theory of expressives is in its infancy- it appears that we might be able to replicate with expressives the typology we observed with indexicals: some are unshiftable, but others can be shifted in attitude reports.

9 Conclusion

Monsters *do* exist in natural language. Each time we report somebody's speech or thought we use one. While the facts are particularly striking in Amharic, they are just as real in English, as is witnessed by the shiftability of 'two days ago'. The general picture is now the following:
(i) Why do monsters exist? Because Frege's insight was in fact correct: even when indexicality is brought into the picture, *the same semantic object* is responsible for the cognitive significance of a matrix sentence and for the truth-conditional contribution of a clause embedded under an attitude verb. In a possible-worlds framework this object can be identified with be a property of contexts, as was suggested originally in Lewis 1979. The implementation of this insight requires a system which has at least the expressive power of full quantification over contexts, but this technical fact is not special to indexicality, since the grammar of tense and mood also requires the apparatus of full quantification over times and possible worlds.

(ii) Why are some indexicals (e.g. English ‘I’) unshiftable? This turned out to be the complement of the question: Why do some indexicals (namely the logophoric pronouns) *have* to be shifted? While the problem can be treated either in syntactic terms (with richly annotated logical forms) or with a semantic stipulation, the latter solution appears preferable, among others because it derives certain puzzling asymmetries in the use of logophoric pronouns and of logophoric mood.

(iii) How are the agreement facts with De Se readings to be accounted for? For person no special mechanism is needed *if* one is willing to change slightly the semantics of De Se readings, and to posit a somewhat stipulative LF syntax for Quantifying In. In the case of tense, however, this only accounts for part of the data, and some Sequence of Tense rule is still needed in the end.

Appendix I.
Expressive Power Arguments: Translation Procedures

In the following we provide translation procedures from First-Order Logic into Temporal English, Modal English and Attitudinal English[63]. In the first two cases the existence of the translation derives the conclusion drawn by Cresswell 1990: the full power of quantification over times and worlds is needed to analyze temporal and modal talk in English. The present argument is somewhat simpler that Cresswell's, however, since our fragments do not involve any individual quantifiers. By contrast, Cresswell 1990 had to consider the interaction between individual and temporal or modal quantifiers to reach his conclusion.

♦ From First-Order Logic into Temporal English

(i) $T(Rx_{i_1}...x_{i_n})$ = *... then*$_{i_1}$ *... then*$_{i_n}$*...*
where '...' stands for a translation of R, for instance (for n=2): *it rains more then*$_{i_1}$ *than it rains then*$_{i_n}$
(ii) $T(\neg\phi)$ = *it is not the case that* T(ϕ)
(iii) $T(\phi\wedge\psi)$ = *T(ϕ) and T(ψ)*
(iv) $T(\forall x_i\phi)$ = *always*$_i$*, when 1=1, it is the case that T(ϕ)*
-Examples:
$T(\forall x_1Px_1)$=*always*$_1$*, when 1=1, ...then*$_1$*...*
e.g. *always*$_1$*, when 1=1, it is the case that then*$_1$ *it rains.*
$T(\forall x_1\forall x_2Rx_1x_2)$= *always*$_1$*, when 1=1, it is the case that always*$_2$*, when 1=1, it is the case that ... then*$_1$ *... then*$_2$*...*
e.g. *always*$_1$*, when 1=1, it is the case that always*$_2$*, when 1=1, it is the case that it rains then*$_1$ *more than it rains then*$_2$*.*

♦ From First-Order Logic into Modal English

(i) $T(Rx_{i_1}...x_{i_n})$ = *... then*$_{i_1}$ *... then*$_{i_n}$*...*
where '...' stands for a translation of R, for instance (for n=2): *it rains more then*$_{i_1}$ *than it rains then*$_{i_n}$
(ii) $T(\neg\phi)$ = *it is not the case that* T(ϕ)
(iii) $T(\phi\wedge\psi)$ = *T(ϕ) and T(ψ)*
(iv) $T(\forall x_i\phi)$ = *necessarily, if 1=1, it is the case that T(ϕ)*
-Examples:
$T(\forall x_1Px_1)$ = *necessarily*$_1$*, if 1=1, ...then*$_1$*...*
e.g. *necessarily*$_1$*, if 1=1, it is the case that then*$_1$ *it rains.*
$T(\forall x_1\forall x_2Rx_1x_2)$= *necessarily*$_1$*, if 1=1, it is the case that necessarily*$_2$*, if 1=1, it is the case that ... then*$_1$ *... then*$_2$*...*
e.g. *necessarily*$_1$*, if 1=1, it is the case that necessarily*$_2$*, if 1=1, it is the case that it rains then*$_1$ *more than it rains then*$_2$*.*

♦ From First-Order Logic into Attitudinal English

The difficulty here is to find an attitude verb in English that quantifies *un*restrictedly over all possible worlds (rather than over the possible worlds *compatible with the agent's attitude).* If we could stipulate that, say, the verb 'conceives' had this property (*John conceives that ϕ* is true iff some world satisfies ϕ), we could simply use the translation T* below, replacing line (iv) with: $T^*(\exists x_i\phi)$=*he conceives*$_i$ *that* T*(ϕ). Alternatively, we can follow the procedure outlined below, which

[63] Note that these fragments of English include indices to disambiguate binding dependencies. If one wanted to translate First-Order Logic into Individual English (=a fragment that includes individual quantifiers such as 'everything'), these indices would of course be needed as well.

yields a translation T that is adequate only on those models in which it is possible for an individual (call him X) to have no beliefs whatsoever. This assumption makes the accessibility relation 'compatible with X's beliefs' vacuous, since *all* worlds should be compatible with the beliefs of someone who suspends judgment on everything...

-Auxiliary Translation:

(i) $T^*(Rx_{i_1}...x_{i_n})$ = *the way things are*$_1$ *and ... and the way things are*$_n$ *stand in relation R* if $n \geq 2$
the way things are$_1$ *has property R* if $n=1$

(ii) $T^*(\neg\phi)$ = *it is not the case that* $T^*(\phi)$

(iii) $T^*(\phi \wedge \psi)$ = $T^*(\phi)$ *and* $T^*(\psi)$

(iv) $T^*(\forall x_i \phi)$ = *he believes*$_i$ *that* $T^*(\phi)$

-Final Translation: For any formula ϕ, $T(\phi)$=*Suppose someone doesn't have any beliefs whatsoever. Then* $T^*(\phi)$.

-Examples:

$T(\forall x_1 Rx_1)$=*Suppose someone doesn't have any beliefs whatsoever. Then he believes*$_1$ *that the way things are*$_1$ *has property R.*

$T(\forall x_1 \forall x_2 Rx_1x_2)$=*Suppose someone doesn't have any beliefs whatsoever. Then he believes*$_1$ *that he believes*$_2$ *that the way things are*$_1$ *and the way things are*$_2$ *stand in relation R.*

Appendix II.
Formal Systems

A. ELD: An Extensional Version of Kaplan's Logic of Demonstratives (with 'Say')

♦ Primitive Symbols

-Punctuation: (,)

-Variables: an infinite set of individual variables: $x_i \in V_{individuals,}$; an infinite set of time variables: $t_i \in V_{times}$; an infinite set of world variables: $w_i \in V_{worlds}$

-Predicates: for all natural numbers p, q, r, an infinite number of p-q-r-place predicates. Among them: 'exist', 'be-elected', which are 1-1-1 predicates

-Names: John, Napoleon, Hume

-Indexicals: I, now, actually

-Sentential connectives: $\wedge$, $\vee$, $\neg$, $\Rightarrow$, $\Leftrightarrow$

-Identity: =

-Quantifiers: $\forall$, Alw, Nec

-Attitude operator: SAY

♦ Syntax

- *Terms*

-If $\alpha \in V_{individuals,}$ or α='John', 'Napoleon, 'Hume' or α='I', α is an individual term.

-If $\beta \in V_{times,}$ or β='now', β is a time term.

-If $\gamma \in V_{worlds,}$ or γ='actually', γ is a world term.

- *Formulas*

-If π is an p-q-r place predicate, α_1, ..., α_p are individual terms, β_1, ..., β_q are time terms, and γ_1, ..., γ_r are world terms, then $\pi(\alpha_1, ..., \alpha_p, \beta_1, ..., \beta_q, \gamma_1, ..., \gamma_r)$ is a formula.

-If ϕ and ψ are formulas, then $(\phi \wedge \psi)$, $(\phi \vee \psi)$, $\neg\phi$, $(\phi \Rightarrow \psi)$, $(\phi \Leftrightarrow \psi)$ are formulas.

-If ϕ is a formula and $\alpha \in V_{individuals}$, $\beta \in V_{times}$, $\gamma \in V_{worlds}$, and $\alpha' \in T_{individuals}$, $\beta' \in T_{times}$, $\gamma' \in T_{worlds}$,then $\forall\alpha\phi$, Alw$\beta\phi$, Nec$\gamma\phi$ and $SAY_{<\alpha', \beta', \gamma'>}\ \gamma\phi$ are formulas.

♦ Semantics

- *Definition*

A is a model for the extensional language of demonstratives iff there are X, T, W, C such that:
(i) A=<X, T, W, C, I>
(ii) C, the set of contexts, is non-empty, and furthermore if $c \in C$:
a. $c_A \in X$ (the agent of c)
b. $c_T \in T$ (the time of c)
c. $c_W \in W$ (the world of c)
(iii) X, T, W, C are non-empty sets and have pair-wise an empty intersection with each other.
(iv) $\boldsymbol{I}$, the interpretation function, assigns
-to each predicate an extension, as follows:
If π is an p-q-r place predicate, for all contexts c, $\boldsymbol{I}_c(\pi) \subseteq X^p \times T^q \times W^r$
-to each name α an extension, for all contexts c: $\boldsymbol{I}_c(\alpha) \in X$
-to each indexical, an extension (which depends on the context):
$\boldsymbol{I}_c(I) = c_A$
$\boldsymbol{I}_c(now) = c_T$
$\boldsymbol{I}_c(actually) = c_W$
-It is further stipulated that for all $c \in C$,
$< c_A, c_T, c_W > \in \boldsymbol{I}_c(exist)$ (the agent of a context exists at the time of the context in the world of the context)
(v) K assigns to each triple $<x, t, w> \in X \times T \times W$ a set of Characters (the set of thoughts compatible with x's claim at t in w), where a Character is any element of $(\{0,1\}^W)^C$
Note: Instead of taking as primitive the set of Characters compatible with someone's claim, we could derive it from the notion: 'context c is compatible with x's claim at t in w'. The definition would run as follows (von Stechow):
k is compatible with x's claim at t in w=: for all c compatible with x's claim at t in w, $k(c)(c_W)=1$.

1. If α is a variable, $[[\alpha]]^{c, s}=s(\alpha)$. If α is a name or an indexical, $[[\alpha]]^{c, s}=\boldsymbol{I}_c(\alpha)$
2. $[[\pi(\alpha_1, ..., \alpha_p, \beta_1, ..., \beta_q, \gamma_1, ..., \gamma_r)]]^{c, s}=1$ iff $<[[\alpha_1]]^{c, s}, ..., [[\alpha_p]]^{c, s}, [[\beta_1]]^{c, s}, ..., [[\beta_q]]^{c, s}, [[\gamma_1]]^{c, s}, ..., [[\gamma_r]]^{c, s}> \in \boldsymbol{I}_c(\pi)$
3. $[[(\phi \wedge \psi)]]^{c, s}=1$ iff $[[\phi]]^{c, s}=1$ and $[[\psi]]^{c, s}=1$
$[[(\phi \vee \psi)]]^{c, s}=1$ iff $[[\phi]]^{c, s}=1$ or $[[\psi]]^{c, s}=1$
$[[(\phi \Rightarrow \psi)]]^{c, s}=0$ iff $[[\phi]]^{c, s}=1$ and $[[\psi]]^{c, s}=0$
$[[(\phi \Leftrightarrow \psi)]]^{c, s}=1$ iff $[[\phi]]^{c, s}=[[\psi]]^{c, s}$
$[[\neg\phi\ ]]^{c, s}=1$ iff $[[\phi\ ]]^{c, s}=0$
4. Quantification
a. $[[\forall\alpha\phi]]^{c, s}=1$ iff for all $x \in X$, $[[\phi]]^{c, s[\alpha \rightarrow x]}=1$
b. $[[Alw\beta\phi]]^{c, s}=1$ iff for all $t \in T$, $[[\phi]]^{c, s[\beta \rightarrow t]}=1$
c. $[[Nec\gamma\phi]]^{c, s}=1$ iff for all $w \in W$, $[[\phi]]^{c, s[\gamma \rightarrow w]}=1$
5. Attitude operators
$[[SAY_{<\alpha', \beta', \gamma'>}\ \gamma\phi]]^{c, s}=1$ iff there exists k in $K(<[[\alpha']]^{c, s}, [[\beta']]^{c, s}, [[\gamma']]^{c, s}>)$ such that
(i) $k(<[[\alpha']]^{c, s}, [[\beta']]^{c, s}, [[\gamma']]^{c, s}>) = \lambda w[[\phi]]^{c, s[\gamma \rightarrow w]}$
(ii) $[[\alpha']]^{c, s}$ says k at $[[\beta']]^{c, s}$ in $[[\gamma']]^{c, s}$

Remarks

(i) 'I actually exist now' is true in every context, since:
$[[\text{exist(I, now, actually)}]]^{c, s}=1$ iff $< c_A, c_T, c_W> \in \mathcal{I}$(exist), which is true by stipulation.

(ii) 'I necessarily exist now' may be false, since:
$[[\text{Nec } w_i \text{ exist(I, now, } w_i)]]^{c, s}=1$ iff for all $w \in W$, $[[\text{exist(I, now, } w_i)]]^{c, s[wi->w]}=1$, which does not have to be true.

(iii) 'John says that John (he) is elected' can be true (among others) if John's utterance was of the form 'I am elected' or 'John/he is elected', since:
$[[\text{SAY}_{<\text{John, now, actually}>} w_i \text{ be-elected(John, now, } w_i)]]^{c, s}=1$ iff there exists k in $K(<[[\text{John}]]^{c, s}, [[\text{now}]]^{c, s}, [[\text{actually}]]^{c, s}>)$, i.e. in $K(<\text{John}, c_T, c_W>)$ such that:
(i) $k(<\text{John}, c_T, c_W>)= \lambda w[[\text{be-elected(John, now, } w_i)]]^{c, s[wi->w]}$
(ii) John says k at c_T in c_W
• If John says: 'I am elected' [i.e. be-elected(I, now, actually)], he says the character: $k^*=\lambda c\, \lambda w\, [[\text{be-elected(I, now, } w_i)]]^{c, s[wi->w]}$. Since $k^*(<\text{John}, c_T, c_W>)=\lambda w\, [[\text{be-elected(I, now, } w_i)]]^{<\text{John}, c_T, c_W>, s[wi->w]}= \lambda w\, [[\text{be-elected(John, now, } w_i)]]^{<\text{John}, c_T, c_W>, s[wi->w]}$, the report is true.
• If John says: 'John is elected' [i.e. be-elected(John, now, actually)], he says the character: $k^*=\lambda c\, \lambda w\, [[\text{be-elected(John, now, } w_i)]]^{c, s[wi->w]}$. Since $k^*(<\text{John}, c_T, c_W>)=\lambda w\, [[\text{be-elected(John, now, } w_i)]]^{<\text{John}, c_T, c_W>, s[wi->w]}$, the report is true.

(iv) Zeevat's Observation: 'John says that he (John) is Hume' is true iff 'John says that he (John) is Napoleon' is true, since:
$[[\text{SAY}_{<\text{John, now, actually}>} w_i \text{ John=Hume}]]^{c, s}=1$ iff $[[\text{SAY}_{<\text{John, now, actually}>} w_i \text{ John=Napoleon}]]^{c,s}=1$ (this is because $\lambda w[[\text{John=Hume}]]^{c, s[wi->w]}=\emptyset=\lambda w[[\text{John=Napoleon}]]^{c, s[wi->w]}$).

B. MEL: A Monster-Friendly Extensional Logic of Demonstratives with Attitude Operators
(This system illustrates in a simplified form the main semantic ideas of the paper. However it deals neither with the Agreement Problem nor with the Filtering Problem).

Same as I., except:

♦ Primitive Symbols

-Add the 1-place context functors :
agent
time
world
two-days-ago
-Add the 0-1-1 predicate: rain
-Add the time quantifier: PAST
-Add the attitude operator HOPE
- Add an infinite set of context variables: $c_i \in V_{\text{contexts}}$

♦ Syntax

Replace the definitions with:

- *Terms*

-If $\alpha \in V_{\text{individuals,}}$ or α='John', 'Napoleon, 'Hume' or α='I', α is an individual term.
-If $\beta \in V_{\text{times,}}$ or β='now', β is a time term.

-If $\gamma \in V_{worlds,}$ or γ='actually', γ is a world term.
-If $\delta \in V_{contexts}$, δ is a context term, agent(δ) is an individual term, time(δ) and two-days-ago(δ) are time terms, world(δ) is a world term.

- *Formulas*

-If π is an p-q-r-s place predicate, α_1, ..., α_p are individual terms, β_1, ..., β_q are time terms, γ_1, ..., γ_r are world terms, and δ_1, ..., δ_s are context terms, then $\pi(\alpha_1, ..., \alpha_p, \beta_1, ..., \beta_q, \gamma_1, ..., \gamma_r, \delta_1, ..., \delta_s)$ is a formula.
-If ϕ and ψ are formulas, then $(\phi \wedge \psi)$, $(\phi \vee \psi)$, $\neg\phi$, $(\phi \Rightarrow \psi)$, $(\phi \Leftrightarrow \psi)$ are formulas.
-If ϕ is a formula and $\beta \in V_{times}$, Past $\beta\phi$ is a formula.
-If ϕ is a formula and $\alpha \in V_{individuals}$, $\beta \in V_{times}$, $\gamma \in V_{worlds}$, $\delta \in V_{contexts}$, and $\alpha' \in T_{individuals}$, $\beta' \in T_{times}$, $\gamma' \in T_{worlds}$,then $\forall\alpha\phi$, Alw$\beta\phi$, Nec$\gamma\phi$, $SAY_{<\alpha', \beta', \gamma'>}\delta\phi$ and $HOPE_{<\alpha', \beta', \gamma'>}\delta\phi$ are formulas.

♦ Semantics

- *Definition*

(iv) Replace the first line with:
-to each predicate an extension, as follows:
If π is an p-q-r-s place predicate, for all contexts c, $\boldsymbol{I}_c(\pi) \subseteq X^p \times T^q \times W^r \times C^s$

(v) Replace with:
K^{SAY} and K^{HOPE} each assigns to each triple <x, t, w> $\in$ X× T× W a set of contexts (=those contexts compatible with x's claim/hope at t in w).

Functors
Add:
$\boldsymbol{I}$, the interpretation function, assigns for all contexts c:
$\boldsymbol{I}_c$(agent)=that function f defined over contexts such that for each c'$\in$C, f(c')= c'$_{A.}$
$\boldsymbol{I}_c$(time)=that function f defined over contexts such that for each c'$\in$C, f(c')= c'$_{T.}$
$\boldsymbol{I}_c$(world)=that function f defined over contexts such that for each c'$\in$C, f(c')= c'$_{W.}$
$\boldsymbol{I}_c$(two-days-ago)=that function f defined over contexts such that for each c'$\in$C, f(c')= c'$_{T.}$ –2 days (this assumes that the set of moments is ordered, etc.)

Add 1'. and replace the other lines by their new version:
1'. If α is a term and f is a functor, $[[f\alpha]]^{c, s} = \boldsymbol{I}_c(f)([[\alpha]]^{c, s})$
2. $[[\pi(\alpha_1, ..., \alpha_p, \beta_1, ..., \beta_q, \gamma_1, ..., \gamma_r, \delta_1, ..., \delta_s)]]^{c, s}=1$ iff $<[[\alpha_1]]^{c, s}, ..., [[\alpha_p]]^{c, s}, [[\beta_1]]^{c, s}, ..., [[\beta_q]]^{c, s}, [[\gamma_1]]^{c, s}, ..., [[\gamma_r]]^{c, s}, [[\delta_1]]^{c, s}, ..., [[\delta_s]]^{c, s}> \in \boldsymbol{I}_c(\pi)$
4. Quantification
Add:
d. $[[Past\beta\phi]]^{c, s}=1$ iff for some t$\in$T such that t< c_T, $[[\phi]]^{c, s[\beta \rightarrow t]}=1$
5. Attitude operators
Replace with:
$[[SAY_{<\alpha', \beta', \gamma'>}\ \delta\phi]]^{c, s}=1$ iff for all c'$\in K^{SAY}$(<$[[\alpha']]^{c, s}$, $[[\beta']]^{c, s}$, $[[\gamma']]^{c, s}$>): $[[\phi]]^{c, s[\delta \rightarrow c']}=1$
$[[HOPE_{<\alpha', \beta', \gamma'>}\ \delta\phi]]^{c, s}=1$ iff for all c'$\in K^{HOPE}$(<$[[\alpha']]^{c, s}$, $[[\beta']]^{c, s}$, $[[\gamma']]^{c, s}$>): $[[\phi]]^{c, s[\delta \rightarrow c']}=1$

Derivations

(i) a. John hopes to be elected *(is unambiguously 'De Se')*
b. $HOPE_{<John, now, actually>}c_i$ be-elected(agent(c_i), time(c_i), world(c_i))

c. $[[(b)]]^{c,s}$ =1 iff for all c'$\in K^{HOPE}$(<John, c_T, c_W>), [[be-elected(agent(c_i), time(c_i), world(c_i))]]$^{c, s[ci->c']}$=1, iff for all c'$\in K^{HOPE}$(<John, c_T, c_W>), <c'$_A$, c'$_T$, c'$_w$>$\in \boldsymbol{I}_c$(be-elected)

(ii) a. John hopes to be Hume *(is not equivalent to: John hopes to be Napoleon)*
b. HOPE$_{<John,now,actually>}c_i$ agent(c_i)=Hume
c. $[[(b)]]^{c,s}$ =1 iff for all c'$\in K^{HOPE}$(<John, c_T, c_W>), [[agent(c_i)=Hume]]$^{c, s[ci->c]}$=1, iff for all c'$\in K^{HOPE}$(<John, c_T, c_W>), c'$_A$= $\boldsymbol{I}_c$(Hume)

(iii) a. John$_i$ hopes that he$_i$, John, is elected (*has a De Re, non-De Se reading)*
b. HOPE$_{<John,now,actually>}c_i$ be-elected(John, time(c_i), world(c_i))
c. $[[(b)]]^{c,s}$ =1 iff for all c'$\in K^{HOPE}$(<John, c_T, c_W>), [[be-elected(John, time(c_i), world(c_i))]]$^{c, s[ci->c']}$=1, iff for all c'$\in K^{HOPE}$(<John, c_T, c_W>), <John, c'$_T$, c'$_w$>$\in \boldsymbol{I}_c$(be-elected)

(iv) a. John hopes that he himself is elected / John hopes that he* is elected (Ewe)
b. Same as (ib)
c. Same as (ic)

(v) a. John$_i$ says that I$_i$ am elected (Amharic) *(has a De Se reading with a shifted indexical)*
b. SAY$_{<John,now,actually>}c_i$ be-elected(agent(c_i), time(c_i), world(c_i))
c. $[[(b)]]^{c,s}$ =1 iff for all c'$\in K^{SAY}$(<John, c_T, c_W>), [[be-elected(agent(c_i), time(c_i), world(c_i))]]$^{c, s[ci->c]}$=1, iff for all c'$\in K^{SAY}$(<John, c_T, c_W>), <c'$_A$, c'$_T$, c'$_w$>$\in \boldsymbol{I}_c$(be-elected)

(vi) a. John says that I am elected (English) *(does not have a shifted reading)*
b. SAY$_{<John,now,actually>}c_i$ be-elected(I, time(c_i), world(c_i))
c. $[[(b)]]^{c,s}$ =1 iff for all c'$\in K^{SAY}$(<John, c_T, c_W>), [[be-elected(I, time(c_i), world(c_i))]]$^{c, s[ci->c]}$=1, iff for all c'$\in K^{SAY}$(<John, c_T, c_W>), <c_A, c'$_T$, c'$_w$>$\in \boldsymbol{I}_c$(be-elected)

(vii) a. John said that it had rained two days ago (English) *(has a De Se reading with a shifted indexical)*
b. Past t_k SAY$_{<John,tk,actually>}c_i$ rain(two-days-ago(c_i), world(c_i))
c. $[[(b)]]^{c,s}$ =1 iff for some t$\in$T before c_T, for all c'$\in K^{SAY}$(<John, t, c_W>), [[rain(two-days-ago(c_i), world(c_i))]]$^{c, s[ci->c']}$=1, iff for some t$\in$T before c_T, for all c'$\in K^{SAY}$(<John, t, c_W>), <c'$_T$-2days, c_w>$\in \boldsymbol{I}_c$(rain)

C. MELP: A Monster-Friendly Logic of Demonstratives with Partiality and Attitude Operators
(This system presents a more sophisticated version of the theory than MEL. The Filtering Problem is addressed in a semantic fashion, along the lines of Section 6. The Agreement Problem, discussed in Section 7, is not addressed).
Primitive Symbols
-Punctuation: (,)
-Variables: an infinite set of individual variables, $V_{individuals,}=\{x_i: i\in|N\}$; an infinite set of time variables, $V_{times}=\{t_i: i\in|N\}$; an infinite set of world variables: $V_{worlds}=\{w_i: i\in|N\}$; an infinite set of context variables: $V_{contexts}=\{c_i: i\in|N\}$
-Individual functor: mother is 1-place individual functor.
-Predicates: for all natural numbers p, q, r, s, an infinite number of p-q-r-s place predicates. Among them: 'exist', 'speak', 'be-elected', 'be-addressed', 'female', 'sick', which are 1-1-1-0 predicates; 'like' is a 2-1-1-0 predicate.
-Features:
a. +feminine, -feminine, +author*, -author*, +hearer*, -hearer*, which are 1-0-0-0 predicates.

b. +author, -author, +hearer, -hearer, which are 1-0-0-1 predicates
c. +present, -present, which are 0-1-0-1 predicates
d. +present*, -present*, which are 0-1-0-0 predicates
e. +indicative^1, -indicative^1, which are 0-0-1-1 predicates
f. +indicative^1*, -indicative^1*, +indicative^2*, -indicative^2*, which are 0-0-1-0 predicates
-Names: John, Napoleon, Hume, Mary
-Sentential connectives: $\wedge$, $\neg$
-Identity: =
-Quantifiers: $\forall$, Alw, Nec
-Definite closure operator: ι
-Attitude operators: SAY, HOPE, BELIEVE, KNOW

♦ Syntax

- *Complex features*

-If π is an p-q-r-s place feature, α_1, ..., α_p are individual variables, β_1, ..., β_q are time variables, γ_1, ..., γ_r are world variables, and δ_1, ..., δ_s are context variables, then $\pi(\alpha_1, ..., \alpha_p, \beta_1, ...\beta_q, \gamma_1 ...\gamma_r, \delta_1, ... \delta_s)$ is a complex feature.
-If ϕ and ψ are complex features, then $(\phi \wedge \psi)$ is a complex feature.
Note: Only conjunctions of features are allowed. Features cannot be negated.

- *Terms*

Let ϕ be any complex feature.
-If $\alpha \in V_{individuals,}$ α, $\alpha\{\phi\}$ and $\iota\alpha\{\phi\}$ are individual terms.
-If α is a name, α is an individual term
-If f is a one place individual functor, and if α is an individual term, so is $f\alpha$
-If $\beta \in V_{times,}$ β, $\beta\{\phi\}$ and $\iota\beta\{\phi\}$ are time terms.
-If $\gamma \in V_{worlds,}$ γ, $\gamma\{\phi\}$ and $\iota\gamma\{\phi\}$ are world terms.
-If $\delta \in V_{contexts}$, δ is a context term.
Note: $\{\phi\}$ is a presupposition on the value of the preceding element.

- *Formulas*

-If π is an p-q-r-s place predicate, α_1, ..., α_p are individual terms, β_1, ..., β_q are time terms, γ_1, ..., γ_r are world terms, and δ_1, ..., δ_s are context terms, then $\pi(\alpha_1, ..., \alpha_p, \beta_1, ...\beta_q, \gamma_1 ...\gamma_r, \delta_1, ... \delta_s)$ is a formula.
-If ϕ and ψ are formulas, then $(\phi \wedge \psi)$, $\neg\phi$ are formulas.
-If ϕ is a formula, ATT an attitude operator, and $\alpha \in V_{individuals}$, $\beta \in V_{times}$, $\gamma \in V_{worlds}$, $\delta \in V_{contexts}$, and $\alpha' \in T_{individuals}$, $\beta' \in T_{times}$, $\gamma' \in T_{worlds}$,then $\forall\alpha\phi$, Alw$\beta\phi$, Nec$\gamma\phi$, $ATT_{<\alpha', \beta', \gamma'>}\delta\phi$ are formulas.
- If ϕ is a formula and $\alpha \in V_{individuals}$, $\beta \in V_{times}$, $\gamma \in V_{worlds}$, and if α', β' and γ' are an individual, a time and a world term respectively, then:
$\alpha'\alpha\phi$, $\beta'\beta\phi$, $\gamma'\gamma\phi$ are formulas.

Note: Terms can optionally be used as variable binding devices. This is useful to analyze bound variable readings.

♦ Semantics

- *Definition*

A is a model for the extensional language of demonstratives iff there are X, T, W, C, I, K such that:
(i) A=<X, T, W, C, ***I***, K>
(ii) C, the set of contexts, is non-empty, and furthermore if $c \in C$:
a. $c_A \in X$ (the agent of c)

b. $c_T \in T$ (the time of c)
c. $c_W \in W$ (the world of c)
(iii) X, T, W, C are non-empty sets and have pair-wise an empty intersection with each other.
(iv) For each attitude operator ATT, K^{ATT} assigns to each triple $<x, t, w> \in X \times T \times W$ a set of contexts (=those contexts compatible with x's attitude at t in w). It is further assumed that:
(v) $\boldsymbol{I}$, the interpretation function, assigns
-to each predicate an extension, as follows:
If π is an p-q-r-s place predicate, for all contexts c, $\boldsymbol{I}_c(\pi) \subseteq X^p \times T^q \times W^r \times C^s$
-to each name an extension, for all contexts c: $\boldsymbol{I}_c$(John), $\boldsymbol{I}_c$(Napoleon), $\boldsymbol{I}_c$(Hume) $\in X$
-It is further stipulated that for all $c \in C$, $x \in X$, $t \in T$, $w \in W$:
$< c_A, c_T, c_W > \in \boldsymbol{I}_c$(exist) (the agent of a context exists at the time of the context in the world of the context)
$< c_A, c_T, c_W > \in \boldsymbol{I}_c$(speak) (the agent of a context speaks at the time of the context in the world of the context)
-Extension of the features:
$\boldsymbol{I}_c$(+feminine)={x: $<x, c_T, c_W> \in \boldsymbol{I}_c$(female)}
$\boldsymbol{I}_c$(+author*)={c_A}
$\boldsymbol{I}_c$(+hearer*)$\subseteq${x: $<x, c_T, c_W> \in \boldsymbol{I}_c$(be-addressed)}
$\boldsymbol{I}_c$(+present*)={c_T}
$\boldsymbol{I}_c$(+indicative^1*)={c_W}
$\boldsymbol{I}_c$(+author)={$<c'_A, c'>$: $c' \in C$}
$\boldsymbol{I}_c$(+hearer)$\subseteq${$<x, c'>$: $<x, c'_T, c'_W> \in \boldsymbol{I}_c$(be-addressed)}
$\boldsymbol{I}_c$(+present)={$<c'_T, c'>$: $c' \in C$}
$\boldsymbol{I}_c$(+indicative^1)={$<c'_W, c'>$: $c \in C$}
$\boldsymbol{I}_c$(+indicative^2*)={c'_W: $c' \in K^{BELIEVE} < c_A, c_T, c_W >$}
The negative features have as extension the entire relevant domain (i.e. they have no semantic contribution. There is, however, a restriction on their use, stated below).
Note: indicative^1* is a predicate that asserts that a world w is the world of the context of the actual utterance. indicative^1 is a relation that asserts that a world w is the world of a context c. Finally, indicative^2* is a predicate that asserts of a world w that w is in the Common Ground (in Stalnaker's sense), i.e. that it is one of those worlds that are compatible with the speaker's belief at the time of utterance.

(vi) For all $c \in C$, $\boldsymbol{I}_c$(mother) $\in$ {x: $x \in \boldsymbol{I}_c$(+feminine)}^{X}

- *Reference and Satisfaction*

1. If α is a name, $[[\alpha]]^{c, s} = \boldsymbol{I}_c(\alpha)$
If α is a variable of type τ (τ=individual, time, world) and if ϕ is a complex feature:
$[[\alpha]]^{c, s} = s(\alpha)$
$[[\alpha\{\phi\}]]^{c, s} = \#$ iff $[[\phi]]^{c, s} = 0$.
$[[\alpha\{\phi\}]]^{c, s} = s(\alpha)$ iff $[[\phi]]^{c, s} \neq 0$
$[[\iota\alpha\{\phi\}]]^{c, s} = \#$ iff there is zero or strictly more than one d in X (if τ=individual) or T (if τ=time) or W (if τ=world) satisfying: $[[\phi]]^{c, s[\alpha \to d]} = 0$. Otherwise,
$[[\iota\alpha\{\phi\}]^{c, s}$= the unique d in X (if τ=individual) or T (if τ=time) or W (if τ=world) satisfying: $[[\phi]]^{c, s[\alpha \to d]} = 0$.
Note: The definite closure operator is a variable binding device, which has on {ϕ} the same effect as a standard definite description operator.

If f is a 1-place individual functor, $[[f\alpha]]^{c, s}= \boldsymbol{I}_c(f)([[\alpha]]^{c, s})$

2. $[[\pi(\alpha_1, ..., \alpha_p, \beta_1, ...\beta_q, \gamma_1 ...\gamma_r, \delta_1, ... \delta_s)]]^{c, s}=1$ iff none of $[[\alpha_1]]^{c, s}$, ..., $[[\alpha_p]]^{c, s}$, $[[\beta_1]]^{c, s}$, ..., $[[\beta_q]]^{c, s}$, $[[\gamma_1]]^{c, s}$, ..., $[[\gamma_r]]^{c, s}$, $[[\delta_1]]^{c, s}$, ..., $[[\delta_s]]^{c, s}$ is #, and $<[[\alpha_1]]^{c, s}$, ..., $[[\alpha_p]]^{c, s}$, $[[\beta_1]]^{c, s}$, ..., $[[\beta_q]]^{c, s}$, $[[\gamma_1]]^{c, s}$, ..., $[[\gamma_r]]^{c, s}$, $[[\delta_1]]^{c, s}$, ..., $[[\delta_s]]^{c, s}>\in \boldsymbol{I}_c(\pi)$

$[[\pi(\alpha_1) ...(\alpha_p) (\beta_1) ...(\beta_q) (\gamma) ...(\gamma_r) (\delta_1) ... (\delta_s)]]^{c, s}=0$ iff none of $[[\alpha_1]]^{c, s}$, ..., $[[\alpha_p]]^{c, s}$, $[[\beta_1]]^{c, s}$, ..., $[[\beta_q]]^{c, s}$, $[[\gamma_1]]^{c, s}$, ..., $[[\gamma_r]]^{c, s}$, $[[\delta_1]]^{c, s}$, ..., $[[\delta_s]]^{c, s}$ is #, and $<[[\alpha_1]]^{c, s}$, ..., $[[\alpha_p]]^{c, s}$, $[[\beta_1]]^{c, s}$, ..., $[[\beta_q]]^{c, s}$, $[[\gamma_1]]^{c, s}$, ..., $[[\gamma_r]]^{c, s}$, $[[\delta_1]]^{c, s}$, ..., $[[\delta_s]]^{c, s}>\notin \boldsymbol{I}_c(\pi)$

3. $[[(\phi \wedge \psi)]]^{c, s}=1$ iff $[[\phi]]^{c, s}=1$ and $[[\psi]]^{c, s}=1$
$[[(\phi \wedge \psi)]]^{c, s}=0$ iff $[[\phi]]^{c, s}=0$ or $[[\psi]]^{c, s}=0$
$[[\neg\phi]]^{c, s}=1$ iff $[[\phi]]^{c, s}=0$
$[[\neg\phi]]^{c, s}=0$ iff $[[\phi]]^{c, s}=1$

4. Quantification
a. $[[\forall\alpha\phi]]^{c, s}=1$ iff for all $x\in X$, $[[\phi]]^{c, s[\alpha->x]}=1$
$[[\forall\alpha\phi]]^{c, s}=\#$ iff for some $x\in X$, $[[\phi]]^{c, s[\alpha->x]}=\#$
b. $[[Alw\beta\phi]]^{c, s}=1$ iff for all $t\in T$, $[[\phi]]^{c, s[\beta->t]}=1$
$[[Alw\beta\phi]]^{c, s}=\#$ iff for some $t\in T$, $[[\phi]]^{c, s[\beta->t]}=\#$
c. $[[Nec\gamma\phi]]^{c, s}=1$ iff for all $w\in W$, $[[\phi]]^{c, s[\gamma->w]}=1$
$[[Nec\gamma\phi]]^{c, s}=\#$ iff for some $w\in W$, $[[\phi]]^{c, s[\gamma->w]}=\#$
d. Let α' be a term and α a variable.
$[[\alpha'\alpha\phi]]^{c, s}=1$ iff for the d such that $d=[[\alpha']]^{c, s}$, $[[\phi]]^{c, s[\alpha->d]}=1$
$[[\alpha'\alpha\phi]]^{c, s}=\#$ iff for the d such that $d=[[\alpha']]^{c, s}$, $[[\phi]]^{c, s[\alpha->d]}=\#$

5. Attitude operators
Let ATT be an attitude operator.
$[[ATT_{<\alpha', \beta', \gamma'>}\ \delta\phi]]^{c, s}=1$ iff for all $c'\in K^{ATT}(<[[\alpha']]^{c, s}, [[\beta']]^{c, s}, [[\gamma']]^{c, s}>)$: $[[\phi]]^{c, s[\delta->c']}=1$
$[[ATT_{<\alpha', \beta', \gamma'>}\ \delta\phi]]^{c, s}=\#$ iff for some $c'\in K^{ATT}(<[[\alpha']]^{c, s}, [[\beta']]^{c, s}, [[\gamma']]^{c, s}>)$: $[[\phi]]^{c, s[\delta->c']}=\#$
Use of negative features
Negative features do not have a semantic contribution. However, they are subject to the following condition:
A negative feature can appear in a logical form only if the corresponding positive feature would have yielded a presupposition failure (i.e. would have yielded the value #).

Compatibility condition relating the initial assignment and the context of utterance.
s is an assignment determined by c [abbreviation: 's(c)'] only if s properly represents the referential intentions of c_A at c_T in c_W

Derivations

(i) a. I exist *(is true in every context in which it can be felicitously uttered)*
b. $exist(x_i\{+author^*(x_i)\}, t_k\{+present^*(t_k)\}, w_l\{+indicative^1(w_l, c^*)\})$
c. $[[(b)]]^{c, s}=\#$ iff $s(x_i)\neq c_A$ or $s(t_k)\neq c_T$ or $s(w_l)\neq c_W$. If $\neq\#$, $[[(b)]]^{c, s}=1$ iff $<s(x_i), s(t_k), s(w_l)>$ $\in \boldsymbol{I}_c(exist)$, iff $< c_A, c_T, c_W>\in \boldsymbol{I}_c(exist)$, which is true by stipulation.

Note: The compatibility condition on s and c forces s(c*) to be c, hence the results above.

(ii) a. I am speaking *(is true in every context in which it can be felicitously uttered)*
b. $speak(x_i\{+author^*(x_i)\}, t_k\{+present^*(t_k)\}, w_l\{+indicative^1(w_l, c^*)\})$

c. same derivation as in (i), with 'speak' replacing 'exist'.

(iii) a. It is necessary that I exist *(may be false)*
b.Nec w_l exist(x_i{+author*(x_i)}, t_k{+present*(t_k)}, w_l)
c. $[[(b)]]^{c,\,s}$=# iff $s(x_i)\neq c_A$ or $s(t_k)\neq c_T$ If ≠#, $[[(b)]]^{c,\,s}$=1 iff for all w∈ W, $[[\text{exist}(x_i\{+\text{author}^*(x_i)\}, t_k\{+\text{present}^*(t_k)\}, w_l\{+\text{indicative}^2{}^*(w_l)\})]]^{c,\,s[wi->w]}$=1, which does not have to be the case.

Note: It is crucial here that no 'indicative' feature appear on the embedded verb.

(iv) a. It is necessary that I be speaking *(may be false)*
b. Nec w_l speak(x_i{+author*(x_i)}, t_k{+present*(t_k)}, w_l)
c. same derivation as in (iii), with 'speak' replacing 'exist'.

Note: It is crucial here that no 'indicative' feature appear on the embedded verb.

(v) a. You are being addressed *(true when uttered felicitously)*
b. addressed(x_i{+hearer*(x_i)}, t_k{+present*(t_k)}, w_l{+indicative^1(w_l, c*)})
c. $[[(b)]]^{c,\,s}$=# unless $s(x_i)\in \boldsymbol{I}_c$(+hearer*), $s(t_k)=c_T$, $s(w_l)=c_W$. If ≠#, $[[(b)]]^{c,\,s}$=1 iff $\langle s(x_i), s(t_k), s(w_l)\rangle \in \boldsymbol{I}_c$(be-addressed). But since $\boldsymbol{I}_c$(+hearer*)⊆{x: <x, c_T, c_W>∈ $\boldsymbol{I}_c$(be-addressed)}, this is always the case when the value of the sentence is not #.

(vi) a. She is female *(true when uttered felicitously)*
b. female(x_i{+feminine(x_i)}, t_k{+present*(t_k)}, w_l{+indicative^1(w_l, c*)})
c. $[[(b)]]^{c,\,s}$=# unless $s(x_i)\in \boldsymbol{I}_c$(+feminine), $s(t_k)=c_T$, $s(w_l)=c_W$. If ≠#, $[[(b)]]^{c,\,s}$=1 iff $\langle s(x_i), s(t_k), s(w_l)\rangle \in \boldsymbol{I}_c$(female). But since $\boldsymbol{I}_c$(+feminine)⊆{x: <x, c_T, c_W>∈ $\boldsymbol{I}_c$(female)}, this is always the case when the value of the sentence is not #.

(vii) a. [Pointing] You are elected and (but) you are not elected *(may be true)*
b. elected(x_i{+hearer*(x_i)}, t_k{+present*(t_k)}, w_l{+indicative^1(w_l, c*)})
& ¬elected(x_m{+hearer*(x_m)}, t_k{+present*(t_k)}, w_l{+indicative^1(w_l)})
c. $[[(b)]]^{c,\,s}$=# unless $s(x_i)$, $s(x_m)\in \boldsymbol{I}_c$(+hearer*), $s(t_k)=c_T$, $s(w_l)=c_W$. If ≠#, $[[(b)]]^{c,\,s}$=1 iff $[[\text{elected}(x_i\{+\text{hearer}^*(x_i)\}, t_k\{+\text{present}^*(t_k)\}, w_l\{+\text{indicative}^1(w_l)\})]]^{c,\,s}$=1 and $[[\text{elected}(x_m\{+\text{hearer}^*(x_m)\}, t_k\{+\text{present}^*(t_k)\}, w_l\{+\text{indicative}^1(w_l)\})]]^{c,\,s}$=0, iff $\langle s(x_i), s(t_k), s(w_l)\rangle \in \boldsymbol{I}_c$(be-elected) and $\langle s(x_m), s(t_k), s(w_l)\rangle \notin \boldsymbol{I}_c$(be-elected), which may both be true if $s(x_i)\neq s(x_m)$

(viii) a. I like my mother *(can have a bound variable reading)*
b. x_i{+author*(x_i)} x_m{+author*(x_m)} like(x_m{+author*(x_m)}, mother x_m{+author*(x_m)}, t_k{+present* (t_k)}, w_l{+indicative^1(w_l, c*)})
Note: x_i{+author*(x_i)} is the variable binder, and x_m{+author*(x_m)} is the bound variable[64].
c. $[[(b)]]^{c,\,s}$=# iff $s(x_i)\neq c_A$ or if for d=$s(x_i)$, $[[\text{like}(x_m\{+\text{author}^*(x_m)\}, \text{mother } x_m\{+\text{author}^*(x_m)\}, t_k\{+\text{present}^*(t_k)\}, w_l\{+\text{indicative}^1(w_l, c^*)\})]]^{c,\,s[xm->d]}$=#.
In sum, $[[(b)]]^{c,\,s}$=# iff $s(x_i)\neq c_A$ or $s(t_k)\neq c_T$ or $s(w_l)\neq c_W$

[64] In a system with λ-abstraction, this would be written as:
x_i{+author*x_i} λx_m{+author*x_m}like(x_m{+author*x_m}, mother x_m{+author*x_m}, ...)

If ≠#, $[[(b)]]^{c, s}$=1 iff [[like(x_m{+author*(x_m)}, mother x_m{+author*(x_m)}, t_k{+present*(t_k)}, w_l{+indicative^1(w_l)})]]$^{c, s[xm->s(xi)]}$=1. Since $[[(b)]]^{c, s}$≠# only if $s(x_i)=c_A$, $s(t_k)=c_T$ and $s(w_l)=c_W$, $[[(b)]]^{c, s}$=1 iff

$<c_A, \boldsymbol{I}_c(\text{mother})(c_A), c_T, c_W> \in \boldsymbol{I}_c(\text{like})$.

(ix) a. John hopes to be elected

b. $\text{HOPE}_{<\text{John, } tk\{+\text{present*(tk)}\}, wl\{+\text{indicative1*(wl)}\}>} c_i$ be-elected(ιx$_m${+author(x_m, c_i)}, ιt_n{+present(t_n, c_i)}, ιw_p{+indicative^1(w_p, c_i)})

c. $[[(b)]]^{c,s}$ =# iff $s(t_k)\neq c_T$ or $s(w_l)\neq c_w$ [65]. If $[[(b)]]^{c,s}$ ≠#,
$[[(b)]]^{c,s}$ =1 iff for all c'∈ K^{HOPE}(<John, c_T, c_W>), [[be-elected(ιx$_m${+author(x_m, c_i)}, ιt_k{+present(t_k, c_i)}, ιw_l{+indicative^1(w_l, c_i)})]]$^{c, s[ci->c']}$=1, iff for all c'∈ K^{HOPE}(<John, c_T, c_W>), $<c'_A, c'_T, c'_w> \in \boldsymbol{I}_c$ (be-elected)

Note: (i) All embedded elements are treated as 'De Se'. (ii) The LF in b. does *not* capture the fact that the embedded elements *must* take as argument the closest context. This should presumably be derived from a syntactic condition on Control.

(x) a. I am elected (Amharic)

b. be-elected(x_m{+author(x_m, c_i)}, t_n{+present(t_n, c_i)}, w_p{+indicative^1(w_p, c_i)})

c. $[[(b)]]^{c,S}$ =# iff $s(x_m)\neq (s(c_i))_A$ or $s(t_n)\neq (s(c_i))_T$ or $s(w_p)\neq (s(c_i))_W$. If $[[(b)]]^{c,s}$ ≠#,
$[[(b)]]^{c,s}$ =1 iff $<s(x_m), s(t_n), s(w_p)> \in \boldsymbol{I}_c$(be-elected)

d. The compatibility condition requires that s properly represent the referential intentions of c_A at c_T in c_W. This must be appealed to in order to ensure that $s(c_i)$=c.

Note: (i) The latter observation predicts that there could be differences between Amharic 'I' and English 'I' even in unembedded sentences (under certain conditions it could be that $s(c_i)\neq c$, if such are the referential intentions of the speaker). I do not know whether this is correct.

(ii) For simplicity Amharic tense and mood have been treated as shiftable indexicals (by analogy with Amharic 'I').

(xi) a. John_i says that I_i am elected (Amharic)

b. $\text{SAY}_{<\text{John, } tk\{+\text{present*(tk)}\}, wl[+\text{indicative1*(wl)}\}>} c_i$ be-elected(ιx$_m${+author(x_m, c_i)}, ιt_n{+present(t_n, c_i)}, ιw_p{+indicative^1(w_p, c_i)})

c. $[[(b)]]^{c,s}$ =# iff $s(t_k)\neq c_T$ or $s(w_l)\neq c_w$ [66]. If $[[(b)]]^{c,s}$ ≠#,
$[[(b)]]^{c,s}$ =1 iff for all c'∈ K^{SAY}(<John, c_T, c_W>), [[be-elected(ιx$_m${+author(x_m, c_i)}, ιt_k{+present(t_k, c_i)}, ιw_l{+indicative^1(w_l, c_i)})]]$^{c, s[ci->c']}$=1, iff for all c'∈ K^{SAY}(<John, c_T, c_W>), $<c'_A, c'_T, c'_w> \in \boldsymbol{I}_c$(be-elected)

(xii) a. John says that he* is elected (Ewe/Gokana)

b. $\text{SAY}_{<\text{John, } tk\{+\text{present*(tk)}\}, wl\{+\text{indicative1*(wl)}\}>} c_i$ be-elected(ιx$_m${(+author(x_m, c_i) ∧ -author*(x_m))},

---

[65] The full condition is: $[[(b)]]^{c,s}$ =# iff $s(t_k)\neq c_T$ or $s(w_l)\neq c_w$ or for some c'∈ K^{HOPE}(<John, c_T, c_W>), it is not the case that: there is exactly one x in X and there is exactly one t in T and there ix exactly one w in W satisfying: $[[+\text{author}(x_m, c_i)]]^{c,s[ci->c'][xm->x]}$=1 and $[[+\text{present}(t_n, c_i)]]^{c,s[ci->c'][tn->t]}$=1 and and $[[+\text{indicative}^1(w_p, c_i)]]^{c,s[ci->c'][wp->w]}$=1. But this condition is always met given the existence and uniqueness of one agent, time and world per context.

[66] The full condition is: $[[(b)]]^{c,s}$ =# iff $s(t_k)\neq c_T$ or $s(w_l)\neq c_w$ or for some c'∈ K^{HOPE}(<John, c_T, c_W>), it is not the case that: there is exactly one x in X and there is exactly one t in T and there ix exactly one w in W satisfying: $[[+\text{author}(x_m)(c_i)]]^{c,s[ci->c'][xm->x]}$=1 and $[[+\text{present}(t_n)(c_i)]]^{c,s[ci->c'][tn->t]}$=1 and and $[[+\text{indicative}^1(w_p)(c_i)]]^{c,s[ci->c'][wp->w]}$=1. But this condition is always met given the existence and uniqueness of one agent, time and world per context.

$\iota t_n\{+present(t_n, c_i)], \iota w_p[+indicative^1(w_p, c_i)\})$
c. $[[(b)]]^{c,s} = \#$ iff $s(t_k) \neq c_T$ or $s(w_l) \neq c_w$. If $[[(b)]]^{c,s} \neq \#$,
$[[(b)]]^{c,s} = 1$ iff for all $c' \in K^{SAY}(<John, c_T, c_W>)$, $[[be\text{-}elected(\iota x_m\{(+author(x_m, c_i) \wedge -author^*(x_m)\}, \iota t_k\{+present(t_k, c_i)\}, \iota w_l\{+indicative^1(w_l, c_i)\})]]^{c, s[ci \rightarrow c']} = 1$. Given the semantic vacuousness of negative features, this is equivalent to: for all $c' \in K^{SAY}(<John, c_T, c_W>)$, $<c'_A, c'_T, c'_w> \in \boldsymbol{I}_c(be\text{-}elected)$
d. Condition of use of $-author^*(x_m)$: replacing $-author^*(x_m)$ with $+author^*(x_m)$ should have produced a presupposition failure. This is indeed the case if for some $c' \in K^{SAY}(<John, c_T, c_W>)$, $c'_A \neq c_A$ (this always holds unless John is claiming that he is me, the agent of the actual speech act).

Note: For simplicity, the embedded tense and mood have been treated as shifted indexicals.

(xiii) a. #I say that I* am elected (Gokana)
b. $SAY_{<xh\{author^*(xh)\},\ tk\{+present^*(tk)\},\ wl\{+indicative1^*(wl)\}>} c_i$ $be\text{-}elected(\iota x_{m\{}((+author(x_m, c_i) \wedge -author^*(x_m)) \wedge +author^*(x_m))\}, \iota t_n\{+present(t_n, c_i)\}, \iota w_p\{+indicative^1(w_p, c_i)\})$
c. (b) cannot be used felicitously because of the clause:
$\{((+author(x_m, c_i) \wedge -author^*(x_m)) \wedge +author^*(x_m))\}$
'$-author^*(x_m)$' can be used only if '$+author^*(x_m)$' yields a presupposition failure.
But since '$+author^*(x_m)$' also appears in the same clause (so that it is evaluated under the same assignments), the sentence cannot be uttered felicitously.

Note: The ungrammaticality/oddity of a. could be derived in some cases even if 'I*' were not given a specification for +author*. The embedded clause would be represented as:
$be\text{-}elected(\iota x_m\{(+author(x_m)(c_i) \wedge -author^*(x_m))\}, \iota t_n\{+present(t_n, c_i)\}, \iota w_p\{+indicative^1(w_p, c_i)\})$
In case the speaker has no uncertainty concerning who he is (a common situation), it will be the case that:
for all $c' \in K^{SAY}(<c_A, c_T, c_W>)$, $c'_A = c_A$. As a result, $+author^*(x_m)$ could be used instead of $-author^*(x_m)$ without causing a presupposing failure. But if it *can* be used, it *must* be used, which explains the deviance of (a). [This predicts that (a) should become grammatical if the speaker is uncertain of his identity].

(xiv) a. John believes that Mary be* sick (German)
b. $BELIEVE_{<John,\ tk\{+present^*(tk)\},\ wl\{+indicative1^*(wl)\}>} c_i$ $sick(Mary, \iota t_n\{+present(t_n, c_i)\}, \iota w_p\{(+indicative^1(w_p, c_i) \wedge -indicative^2*(w_p))\})$
c. $[[(b)]]^{c,s} = \#$ iff $s(t_k) \neq c_T$ or $s(w_l) \neq c_w$. If $[[(b)]]^{c,s} \neq \#$,
$[[(b)]]^{c,s} = 1$ iff for all $c' \in K^{BELIEVE}(<John, c_T, c_W>)$, $[[sick(Mary, \iota t_n\{+present(t_n, c_i)\}, \iota w_p\{(+indicative^1(w_p, c_i) \wedge -indicative^2*(w_p))\})]]^{c, s[ci \rightarrow c']} = 1$. Given the semantic vacuousness of negative features, this is equivalent to: for all $c' \in K^{BELIEVE}(<John, c_T, c_W>)$, $<Mary, c'_T, c'_w> \in \boldsymbol{I}_c(sick)$
d. Condition of use of $-indicative^2*(w_p)$: replacing $-indicative^2*(w_p)$ with $+indicative^2*(w_p)$ should have produced a presupposition failure. This is indeed the case if for some $c' \in K^{BELIEVE}(<John, c_T, c_W>)$, $c'_A \notin K^{BELIEVE}(<c_A, c_T, c_W>)$ (this holds unless John's epistemic alternatives are a subset of my, the speaker's, epistemic alternatives).

(xv) a. #I believe that Mary be* sick (German)
b. $BELIEVE_{<xh\{author^*(xh)\},\ tk\{+present^*(tk)\},\ wl\{+indicative1^*(wl)\}>} c_i$ $sick(Mary, \iota t_n\{+present(t_n, c_i)\}, \iota w_p\{(+indicative^1(w_p, c_i) \wedge -indicative^2*(w_p))\})$

c. $[[(b)]]^{c,\,s}$=# if $s(x_h)\neq c_A$ or $s(t_k)\neq c_T$ or $s(w_l)\neq c_W$. But if $s(x_h)=c_A$ and $s(t_k)=c_T$ and $s(w_l)=c_W$, for every $c'\in K^{BELIEVE}(<s(x_h), s(t_k), s(w_l)>)$ (i.e. for every $c'\in K^{BELIEVE}(<c_A, c_T, c_W>)$), $[[\iota w_p\{(+\text{indicative}^1(w_p, c_i) \wedge \text{-indicative}^2{*}(w_p))\}]]^{c,\,s[ci\to c']}= c'_W$
But since $<s(x_h), s(t_k), s(w_l)>= <c_A, c_T, c_W>$, for every $c'\in K^{BELIEVE}(<s(x_h), s(t_k), s(w_l)>)$: $c'\in K^{BELIEVE}(<c_A, c_T, c_W>)$, and thus $[[\text{indicative}^2{*}(w_p)]]^{c,\,s[wp\to c'w]}=1$, which entails that $[[\iota w_p\{(+\text{indicative}^1(w_p, c_i) \wedge +\text{indicative}^2{*}(w_p))\}]]^{c,\,s[ci\to c']}= c'_W$. In other words, $+\text{indicative}^2{*}(w_p)$ could be used instead of $\text{-indicative}^2{*}(w_p)$ without causing a presupposition failure. But if it *can* be used, it *must* be used, which explains the deviance of (a).

References

Abusch, Dorit (1997) 'Sequence of Tense and Temporal *De Re'*, *Linguistics & Philosophy* 20: 1-50

Anderson, S. R. & E. L. Keenan (1985) 'Deixis', in Shopen, M. (ed.) *Language Typology and Syntactic Description,* vol. 3: 'Grammatical Categories in the Lexicon', Cambridge University Press.

Anscombe, G. E. M (1975) 'The First Person', reprinted in Cassam, Quassim, ed. *Self-Knowledge.*

Banfield, Ann (1982) *Unspeakable Sentences (Narration and Representation in the Language of Fiction),* Routledge & Kegan Paul.

Barentsen, Adrian (1996) 'Shifting points of orientation in Modern Russian: Tense selection in 'reported perception'', in Janssen & van der Wurff, ed. *Reported Speech*

Bäuerle, R. (1983) 'Pragmatisch-semantische Aspekte der NP-Interpretation', in M. Faust et al. (eds.) *Allgemeine Sprachwissenschaft, Sprachtypologie und Textlinguistic,* Tübingen 1983, 121-131.

Benthem, J. van (1978) 'Tense logic and standard logic', *Logique et Analyse* 80, 47-83.

Blackburn, P., M. de Rijke & Y. Venema (2001) *Modal Logic,* Cambridge University Press

Boolos, George (1984) 'To Be is to Be a Value of a Variable (or to Be Some Values of Some Variables', *Journal of Philosophy,* 81: 430-449.

Burge, Tyler (1974) 'Demonstrative Constructions, Reference, and Truth', *The Journal of Philosophy* 71, 7: 205-223

Cassam, Quassim [ed.] (1994) *Self-Knowledge,* Oxford University Press.

Castañeda, Hector-Neri (1966) 'He: A Study in the Logic of Self-Consciousness', *Ratio* 7.

Castañeda, Hector-Neri (1967) 'Indicators and Quasi-Indicators', *American Philosophical Quaterly,* Volume 4, Number 2.

Castañeda, Hector-Neri (1968) 'On the Logic of Attributions of Self-Knowledge to Others', *The Journal of Philosophy,* Volume 65, Number 15.

Chierchia, Gennaro (1987) 'Anaphora and Attitudes *De Se',* in Bartsch, van Benthem & van Emde Boas (eds), *Language in Context,* Doris.

Clements, George N. (1975) 'The logophoric pronoun in Ewe: Its role in discourse', *Journal of West African Languages* 10: 141-77

Cooper, Robin (1983) *Quantification and Syntactic Theory,* Reidel.

Cresswell, Maxwell (1990) *Entities and Indices,* Kluwer

Dekker, Paul (2000) 'Coreference and Representationalism', in K. von Heusinger and U. Egli (eds.), *Reference and Anaphorical Relations,* Kluwer, Dordrecht.

Doron, Edit (1991) 'Point of View as a Factor of Content', SALT I.

Farkas, Donka (1997) 'Evaluation Indices and Scope', in Szabolcsi, ed. *Ways of Scope-Taking,* Kluwer.

Fintel, Kai von (1997) 'The Presupposition of Subjunctive Conditionals', MIT Working Papers in Linguistics 25, eds. Orin Percus and Uli Sauerland.

Frajzyngier, Zygmunt (1985) 'Logophoric Systems in Chadic', *Journal of African Languages and Linguistics* 7: 23-37

Frajzyngier, Zygmunt (1993) *A Grammar of Mupun,* Dietrich Reimer Verlag

Frege, Gottlob (1892) 'Über Sinn und Bedeutung', Zeitschrift für Philosophie und philosophische Kritik, pp. 25-50. Reprinted in Frege, 'Funktion, Begriff, Bedeutung', edited by G. Patzig, Göttingen, 1986.

Haas-Spohn, Ulrike (1995) *Versteckte Indexicalität und subjektive Bedeutung,* Studia Grammatica 38, Akademie Verlag.

Hagège, Claude (1974) 'Les Pronoms Logophoriques', *Bulletin de la Société de Linguistique de Paris* 69, 287-310

Hale, Kenneth & Paul Platero (1996) 'Negative Polarity Expressions in Navajo', ms., MIT

Heim, Irene (1991a). 'Artikel und Definitheit.' In A. v. Stechow and D. Wunderlich (eds) *Handbuch der Semantik.* Berlin: de Gruyter, 1991.
Heim, Irene (1991b) Class hand-outs on 'Control' (Seminar taught by Heim & Higginbotham, Spring 1991): 'Interpretation of PRO' (Feb. 22), 'The first person' (March 8), and a hand-out on 'connectedness' (April 12).
Heim, Irene (1994a) 'Puzzling reflexive pronouns in *de se* reports', hand-out, Bielefeld Conference, March 1994.
Heim, Irene (1994b) 'Comments on Abusch's Theory of Tense', in H. Kamp (ed.), *Ellipsis, Tense and Questions,* pp. 143-170, U. of Amsterdam.
Heim, Irene (2002) 'Features of Pronouns in Semantics and Morphology', hand-out of a talk given at USC on 01.31.02.
Heim, Irene & Angelika Kratzer (1998) *Semantics in Generative Grammar*. Oxford: Blackwell
Higginbotham, James (2000) 'Competence with Demonstratives', to appear in B. Ramberg & M. Hahn (eds.), *Festschrift for Tyler Burge*
Hyman, Larry M., ed. (1979). *Aghem Grammatical Structure*. Southern California Occasional Papers in Linguistics No. 7
Hyman, Larry M. & Bernard Comrie (1981) 'Logophoric Reference in Gokana', *Journal of African Languages and Linguistics* 3: 19-37
Iatridou, Sabine (1994) 'On the Contribution of Conditional *Then'*, *Natural Language Semantics* 2: 171-199
Iatridou, Sabine (1998) 'The Grammatical Ingredients of Counterfactuality', ms., MIT
Israel, David & John Perry (1996) 'Where Monsters Dwell', in Seligman & Westerståhl, ed. *Logic, Language and Computation,* Volume 1.
Jäger, Siegfried (1971) *Der Konjunktiv in der deutschen Sprache der Gegenwart,* Max Hueber Verlag.
Jacobson, Pauline (1999) 'Towards a Variable-Free Semantics', *Linguistics and Philosophy* 22: 117-184.
Janssen, Theor & Wim van der Wurff, ed. (1996) *Reported Speech,* John Benjamins
Kadmon, Nirit (2001) *Formal Pragmatics*, Blackwell
Kaplan, David (1969) 'Quantifying In', in *Words and Objections,* ed. D. Davidson & J. Hintikka, Dordrecht: Reidel.
Kaplan, David (1990) 'Thoughts on Demonstratives', in ed. P. Yourgrau, *Demonstratives*, Oxford University Press, 1990.
Kaplan, David (1977/1989) 'Demonstratives', in *Themes from Kaplan,* ed. Almog/Perry/Wettstein, Oxford University Press, 1989
Kaplan, David (1989) 'Afterthoughts', in *Themes from Kaplan,* ed. Almog/Perry/Wettstein, Oxford University Press, 1989
Kaplan, David (2001) 'The Meaning of Ouch and Oops (Explorations in the theory of *Meaning as Use*', Draft #3, ms., UCLA
Kondrashov, D. & N. Kondrashova (1999). 'Double-access double-decker: Where semantics can get a free ride on syntax', hand-out, Thermi Conference on Tense, Aspect and Mood.
Koopman, Hilda & Dominique Sportiche (1989) 'Pronouns, Logical Variables, and Logophoricity in Abe', *Linguistic Inquiry,* 20: 4
Kratzer, Angelika (1991) 'Conditionals', in A. von Stechow, D. Wunderlich (ed.) *Semantik: ein internationales Handbuch der zeitgenössischen Forschung.*
Kratzer, Angelika (1997) 'Impersonal Pronouns and Logophoricity'. Hand-out of a talk given at *Sinn und Bedeutung* II, Berlin, December 1997.
Kratzer, Angelika (1998) 'More Structural Analogies Between Pronouns and Tenses', *Proceedings of SALT VIII*
Kuhn, Steven (1980) 'Quantifiers as Modal Operators', *Studia Logica* 39:145-158.

Kuno, Susumu (1972) 'Pronominalization, Reflexivization, and Direct Discourse' *Linguistic Inquiry* 3(2): 161-195

Kusumoto, Kiyomi (1998). 'Tenses as logophoric pronouns'. Hand-out of talk given at the MIT/UConn/UMass Semantics Workshop, October 31, 1998 (UConn)

Leslau, Wolf (1995). *Reference Grammar of Amharic.* Harrassowitz - Wiesbaden.

Lewis, David (1970) 'General Semantics'. Reprinted in *Philosophical Papers,* Vol. I, Oxford University Press, 1983

Lewis, David (1979) 'Attitudes De Dicto and De Se'. *The Philosophical Review* 88, 513-43. Reprinted in Lewis 1983

Morgan, Jerry (1970) 'On the Criterion of Identity for Noun Phrase Deletion', CLS 6.

Pan, Hai-Hua (1997) *Constraints on Reflexivization in Mandarin Chinese,* Garland.

Partee, Barbara (1973) 'Some Structural Analogies Between Tenses and Pronouns in English', *The Journal of Philosophy* 70, 601-609.

Partee, Barbara (1984) 'Nominal and Temporal Anaphora', *Linguisics & Philosophy* 7, 243-286.

Partee, Barbara (1989) 'Binding Implicit Variables in Quantified Contexts', CLS 25.

Perry, John (1977) 'Frege on Demonstratives', *Philosophical Review,* 86. Reprinted in Yourgrau 1990.

Perry, John (1979) 'The Problem of the Essential Indexical', *Noûs* 13: 3-21. Reprinted in N. Salmon & S. Soames (eds) *Propositions and Attitudes,* oxford University Press, 1988.

Predelli, Stefano (1998) 'Utterance, interpretation and the logic of indexicals' *Mind and Language* 13, 400-414

Quine, W. V. (1956) 'Quantifiers and Propositional Attitudes', *Journal of Philosophy,* 53.

Quine, W. V. (1960) 'Variables Explained Away', reprinted in Quine, *Selected Logic Papers,* Harvard University Press, 1995.

Recanati, François (1993) *Direct Reference: From Language to Thought,* Blackwell.

Reinhart, Tanya (1983) 'Point of View in Language - The Use of Parentheticals', in Gisa Rauh (ed.) *Essays on Deixis,* Gunter Narr Verlag Tübingen.

Reinhart, Tanya (1991) 'Self-Representation'. Manuscript.

Roncador, Manfred von (1988) *Zwischen direkter und indirekter Rede,* Niemeyer.

Rooth, Mats (1996) 'Focus', in Lappin, ed. *Handbook of Contemporary Semantic Theory,* Blackwell.

Scott, Dana (1970) 'Advice on Modal Logic', in K. Lambert (ed.) *Philosophical Problems in Logic. Recent Developments.* Reidel, Dordrecht, 143-173.

Sharvit, Yael (2001) 'Hebrew 'Past-under-Past' sentences and the SOT Parameter', hand-out of a talk given at the Conference on the Syntax and Semantics of Semitic Languages, USC, May 4-5 2001.

Speas, Peggy (1999) 'Person and Point of View in Navajo Direct Discourse Complements', ms., University of Massachusetts, Amherst

Stalnaker, Robert (1975) 'Indicative Conditionals', *Philosophia* 5: 269-86

Stalnaker, Robert (1978) 'Assertion', in *Syntax and Sematnics* 9, Academic Press.

Stalnaker, Robert (1981) 'Indexical Belief', *Synthese* 49

Stalnaker, Robert (1999) *Context and Content,* Oxford University Press.

Stechow, Arnim von (2001) 'Schlenker's Monsters', hand-out of a talk given at UCLA, April 4th, 2001 (available at http://www2.sfs.nphil.uni-tuebingen.de/arnim/Schlenkers.Monster.pdf)

Stone, Matthew (1997) 'The Anaphoric Parallel between Modality and Tense', IRCS Report 97-06, University of Pennsylvania

Thomas, Elaine (1978). *A Grammatical Description of the Engenni Language.* Summer Institute of Linguistics Publication Number 60.

Weinstein, Scott (1974) 'Truth and Demonstratives', *Noûs* 8: 179-184. Reprinted in Ludlow 1997

Zimmermann, Thomas Ede (1991) 'Kontextabhängigkeit', in A. von Stechow, D. Wunderlich (ed.) *Semantik: ein internationales Handbuch der zeitgenössischen Forschung.*