

Remembering, Imagining, and the First Person*

James Higginbotham

University of Southern California

1. Introduction.

It is widely supposed that certain uses of anaphoric forms (pronouns, reflexives, and others to be discussed below) give rise to peculiarly "first-personal" interpretations, and it has become customary following David Lewis (1979) to call these interpretations *de se*. Assuming that *de se* interpretations do indeed contrast with interpretations along the familiar, if not necessarily pellucid, *de dicto-de re* axis, there are then four questions about the *de se*, the first two more philosophical, and the latter two more linguistic: (i) what is the nature of *de se* interpretations? (ii) what relation do they bear to ordinary uses of the first-person pronoun? (iii) why are they triggered by the particular linguistic items that trigger them? and (iv) are they universal in human language, and what relation, if any, do they bear to logophoric phenomena in languages having special logophoric forms?¹ Here I consider almost exclusively the first question, hazarding only a few remarks about the second and third; and I omit the fourth, most properly linguistic, question entirely. I believe that questions (iii) and (iv) in particular are more deserving of philosophical attention than it might at first seem; but that discussion will have to await another occasion.

Constructions such as those in (1)-(3) below, stretching back to the original work of Castañeda (1966) and (1967), have been discussed at length over the years:

(1) John/Each man expects (that) he will win.

(2) John/Each man expects (that) he himself will win.

(3) John/Each man expects himself to win.

To these constructions must be added the case where the subject of the infinitival complement is understood, as in *John/Each man expects to win*. Throughout this article, I represent the understood subject by the element PRO, as in Chomsky (1981). PRO is an expression having an interpretation (in fact, necessarily anaphoric to the main clause subject), but no phonetic realization. Alternative views of the understood subject are possible. My choice may not ultimately affect the semantic question (i) above, but it will have consequences for (ii)-(iv).² Given PRO as the understood subject, to the examples (1)-(3) we may add (4):

(4) John/Each man expects PRO to win.

The datum to be considered, then, is that there is something first-personal about the contents of the complements in (2)-(4) that need not obtain in (1); that is, that there are contexts in which the assertion (1) is true (with *he* anaphoric to the subject, *John* or *each man*) that are not first-personal, but no such contexts for the assertion of any of (2)-(4). I need not give a preliminary rehearsal of the familiar scenarios given in support of this view; see below for discussion of examples.³

The contrast to be considered is between the ways in which (2)-(4) must be understood, and (1) need not be understood, so that (1) possesses an ambiguity that (2)-(4) lack.⁴

There would appear to be three, and only three, modes of explanation of the observed contrast.

First (as in Higginbotham (1989) and (1991), for example) it may be suggested that the contexts in which (1) is not first-personal are concealed *de dicto* contexts, where some

Fregean “mode of presentation” attends the bound pronoun *he*. Consider Castañeda's example of the amnesiac war hero, who is reading about himself and his heroic exploits, all unaware, indeed disbelieving, that it is he himself that he is reading about. We are licensed to say truly (5):

(5) The war hero thinks that he is a hero.

but not (6):

(6) The war hero thinks that he himself is a hero.

How can these be distinguished? According to the first method, they are distinguished in that, while the pronoun in (5) need not be taken as a mere variable, having for its antecedent the expression *the war hero*, the emphatic reflexive must be so taken.

Suppose that we understand the ordinary pronoun as if, while referring to the war hero, it presented him as, say, *the person that he (the war hero) is reading about*. Then we have a sense in which (5) can be true although (6), being for whatever reason strictly *de re*, is false. So the interpretations of (2)-(4) must be, and that of (1) can be, first-personal simply because *de re*.

The above view is rejected in Lewis (1979) for substantially theoretical reasons. It may also be questioned on the ground that it doesn't reveal what is peculiarly first-personal about, say, John's or each man's expectations in (3). The latter point will chiefly occupy us here.

Second, following the path taken by Lewis, and in a simplified form by Chierchia (1990), it may be proposed that the complements of the *de se* constructions are of a higher logical type, properties instead of propositions. Thus, according to Chierchia, (2) has it that, for John or each man= x , x expects the property of being a thing y such that y

wins; or, to put it perhaps more naturally, x expects of x that property. If x 's expectations come to pass, then x has the property, and so x wins. Likewise, when the war hero thinks he is a hero, but not that he himself is a hero, then for $\text{him}=x$, x thinks that x is a hero, but it is not the case that for $\text{him}=x$, x thinks, or ascribes to x , the property of being a thing y such that y is a hero.

I shall elaborate somewhat upon this second proposal, and what I will argue is its chief shortcoming, in what follows. For the moment, I note that Stalnaker (1981) responded to Lewis in such a way as to defend a variant of the first view above, albeit from a different angle. For Stalnaker, the modes of presentation relevant to (5) do not figure in the logical form of the examples, but rather in setting up the possible worlds in which John has expectations about himself, but doesn't realize that it is himself he has expectations about. I discuss this view in section 7 below.

Third, there is a possibility, sympathetically elucidated in Perry (1983) (but leading in the end to a skeptical conclusion), and notoriously having roots in Frege's (1918) discussion of the first person, as well as in Castañeda's work: it may be that there is a *special*, first-personal interpretation of the emphatic reflexive as in (2), of the reflexive as in (3), and of PRO as in (4), contributing its meaning to the propositional complement; and perhaps this special interpretation is also involved in the first-person forms themselves. The special interpretation would be available for (1), but not obligatory. The problem then is to elucidate what this interpretation is.

In this article I will argue that the third view is correct, and I will offer a particular way of understanding it. If it is correct, then *de se* interpretations of embedded clauses, obligatory in (4) and (6), optional in (1) and (5) above, are neither the result of

suppressing a conceptualized constituent in favor of the bare object, as in the first view, nor do they call for a reconstrual of clauses as expressions of properties, as in the second, but rather they have their own conceptualized constituents (analogous to Fregean senses, but only analogous, since they may, and in the view I advance, will, contain objects that are not senses), whose properties explain why they are different from *de re* contexts.

One approach to what is special about first-person contexts is already illustrated in Peacocke (1981) as well as Perry (1983), cited above. I use Peacocke's exposition: suppose that there is a special mode of presentation *self* that a thinker *x* can employ in thinking about himself, and no one else, and others cannot therefore employ in thinking about *x*. A particular first-person thought will employ a token [*self*_{*x*}], indexed by *x*. The problem will now be to elucidate the empirical as well as the formal content of *self* and [*self*_{*x*}].

The view that I advance in section 6 below will not be Peacocke's, for reasons having to do with some empirical peculiarities of first-personal embedded constructions that I outline in earlier sections. Like his, however, it may be thought of as sympathetic to a latter-day Fregean perspective; and there are some direct points of contact, as well as room for differences. My remarks will in any case cash out some promissory notes in Higginbotham (1995) and (forthcoming).

2. Gerundive Complements

I have outlined a general view, that there is something distinctive in the beliefs or expectations about oneself that are alleged in assertions of examples such as (2)-(4), or (6). The phenomena may be illustrated with other types of examples, where the anaphora having distinctive interpretations appear within complement clauses, and so on a classical

view occur as constituents of linguistic elements referring to propositions. The linguistic contexts hosting these examples would include predicates expressing epistemic states, as in knowing oneself to be so and so; indirect discourse; and other predicates involving states, conditions, or activities, such as dreaming, fantasizing, or pretending. Here, however, I turn to a different domain, that of the contexts of remembering, imagining, and the like, where the anaphor is the subject of a verbal gerundive complement, as in (7)-(10):

(7) John/Each man remembered/imagined [his going to the movies]

(8) John/Each man remembered/imagined [him, himself going to the movies].

(9) John/Each man remembered/imagined [himself going to the movies]

(10) John/Each man remembered/imagined [PRO going to the movies]

The above examples are parallel, or as parallel as can be, to (1)-(4) above.⁵ I call particular attention to (10).

The first-personal character of such reports as (10) is foreshadowed in the notorious example (11), which arose some years ago in discussion between Jerry Fodor and Judith Thomson:

(11) Only Churchill remembers giving the speech (about blood, toil, tears, and sweat).

As Fodor (1975, p.133 ff.) remarks, (11) is true provided that (a) only Churchill gave the speech, and (b) he remembers doing so. Thus (12) is a valid argument (with PRO explicit in the premisses and conclusion):

(12) Only Churchill gave the speech;

Churchill remembers [PRO giving the speech]; therefore,

Only Churchill remembers [PRO giving the speech].⁶

Compare the conclusion of (12) to (13):

(13) Only Churchill remembers his giving the speech.

The latter is, at first sight, obviously false, and is therefore not implied by the premisses of (12). Indeed, it appears that, for the simple reason that those people who listened to Churchill's speech (in person, or on the radio) at the time he gave it, were aware in the ordinary way of what was going on, and subsequently remembered it, are all of them counterexamples to the truth of (13). If so, (13) is not true on any interpretation.

Now, Fodor's observation does not, and should not, show that what Churchill remembers when he remembers giving the speech is different from what he or others remember when they remember his giving it. We shall, however, accumulate evidence that there is indeed something special about Churchill's memorial state.

Before returning to Fodor's and similar cases, I endeavor in the next two sections to establish two points: first, that the gerundive complements in (7)-(10) have an event-like (in a sense to be explained) rather than a proposition-like reference; and second, that those with PRO subjects exhibit the phenomenon of "immunity to error through misidentification" in the sense of Shoemaker (1968), and as refined in Pryor (1999).

3. The Interpretation of (Certain) Gerundive Complements

It is obvious that we remember events, and that we can speak of ourselves as engaged in remembering them. Furthermore, to remember an event --- the Vietnam War, for example, or a birthday party --- one must have had appropriate perceptual and other experience of it. The memory of an event may "fade," as we say, and in particular one may know that one used to remember something one no longer remembers. Using these facts, we can swiftly show that the verbal gerundive complements to *remember* fall on

the event side of things, and in this way are sharply to be distinguished from the finite complements *that so and so*.

(i) My father's father died before I was born. I remember that he was called "Rufus." But I do not remember his being called "Rufus," because I was not alive when he was called that. Thus the complement

my grandfather's being called "Rufus"

does not refer to the proposition (or fact) that he was called "Rufus;" for if it did, then (14) and (15) would be true or false together, whereas in fact (14) is true, and (15) false:

(14) I remember that my grandfather was called "Rufus."

(15) I remember my grandfather's being called "Rufus."

(ii) Expressions involving propositional attitude or epistemic states resist the English progressive. Thus the examples in (16) are simply ungrammatical:

(16) I am believing/knowing/remembering that I walked to school in the 5th grade.

However, if someone comes across me looking pensive, and I am asked, "What are you doing?" I can well respond with (17):

(17) I am remembering walking to school in the 5th grade.

as indeed I could respond with, "I am remembering the birthday party," where the direct object of *remember* obviously refers to an event. We conclude that the complement in (17) so refers.⁷

(iii) Finite complements contrast with the gerundive in that, whereas (18) is reminiscent of the "Moore Paradox," the anomaly of asserting that one does not believe something that one has just asserted, (19) may be an ordinary truth:

(18) I used to remember that I walked to school in the 5th grade, but I no longer

remember it.

(19) I used to remember walking to school in the 5th grade, but I no longer remember it.

The above diagnostics show that the gerundive complements to *remember* speak of memory of events, rather than remembering facts, or remembering, or remembering-true, propositions.

Throughout the above, I have assumed that *remember*, in all uses, is factive: one can only remember true propositions; and when one speaks of remembering, say, walking to school, then there was indeed an event of one's walking to school. Whether factivity can be safely assumed or not, *remember* conspicuously contrasts with the non-factive *imagine*, and indeed with most other contexts in which the analogues of the above diagnostics can be carried out. There must, then, be a general perspective, abstracting away from factivity, from which gerundive and finite complements are put into semantic contrast. Our task in bringing it out is somewhat complicated by the fact that gerundive complements do not always contrast with finite or other complements in the stark way we have just illustrated, using *remember*. Passing over the complexities of the linguistic taxonomy, I outline a way that I would favor of making the semantic contrast where it obtains.

I will assume, as in the familiar story derived from Davidson, and exploited by myself and many others, that the basic argument structure of a natural language predicate P will have, besides some number n of slots for the overt arguments that appear with it, also a slot for a variable e that ranges over events, thus:

$$P(x_1, \dots, x_n, e)$$

(We abstract as usual from considerations of time and tense, as well as from the details of English morphology.) Supposing the arguments apart from the position marked by e filled in, we derive a predicate $\varphi(e)$ true only of events, as it might be $\text{love}(\text{John}, \text{Mary}, e)$.⁸

An assertion of the sentence *John loves Mary* is an assertion:

$$(\exists e) \text{love}(\text{John}, \text{Mary}, e)$$

to the effect that there is at least one event of (as we may say) John loving Mary. Any sentence that may be asserted may occur also as a finite complement (hosted by a complementizer such as *that*), in which case it will refer to a proposition. Following Richard Montague's (1960) notation, and for the purposes of this article his interpretation as well, the reference of the finite complement is

$$\hat{\ }(\exists e) \text{love}(\text{John}, \text{Mary}, e)$$

where the circumflex '^' represents λ -abstraction over possible worlds. If $\psi(p)$ is a factive predicate with respect to a propositional argument p , then its factivity is expressed by adopting the postulate:

$$\psi(p) \rightarrow \check{p}$$

where \check{p} refers to the extension, in this case the truth value, of p in the actual world. If $\psi(p)$ is so to speak anti-factive, implying the falsehood of its propositional argument (as is plausibly the case, for instance, with the propositional argument in *He fancies that p*), then the postulate to be adopted is

$$\psi(p) \rightarrow \check{\neg}p$$

The above familiar apparatus may now be modified so as to apply to the case of event-like reference, as in the gerundive complements. Consider a predicate $\varphi(e)$ of events. This predicate yields a term $(\lambda e)\varphi(e)$ by λ -abstraction, hence a term $\hat{\ }(\lambda e)\varphi(e)$ referring

(in Montague's terminology) to properties of events. We propose that a predicate Γ with respect to an event-like argument (as established through the above diagnostics) takes arguments α of the type of $\lambda e\phi(e)$: thus *John remembers/imagines Mary singing the song* would be

$$(\exists e') [\text{remembers/imagines}(\text{John}, \lambda e \text{ sing}(\text{Mary}, \text{the song}, e), e')]$$

If $\Gamma(\alpha)$ is a factive predicate with respect to arguments α , then we adopt the postulate

$$\Gamma(\alpha) \rightarrow (\exists e^*) [(\forall \alpha) e^*]$$

and if $\Gamma(\alpha)$ is anti-factive (as is plausibly the case for the gerundive complement in *He feigned taking offense at that remark*) then the postulate to be adopted is

$$\Gamma(\alpha) \rightarrow \neg(\exists e^*) [(\forall \alpha) e^*]$$

An easy calculation now confirms that *John remembers Mary singing the song*, together with its factive postulate, implies

$$(\exists e^*) \text{ sing}(\text{Mary}, \text{the song}, e^*)$$

We have thus, following in effect Montague's strategy of ascending to the worst-case scenario, found a uniform way to give the semantics of gerundive complements in our target contexts, while accommodating, via postulates, the intuitions governing the factive and anti-factive cases.

Suppose now that in place of gerundive complements we have ordinary nominal objects, as in (20):

(20) Mary remembered/imagined the party for her 33rd birthday/a tree.

We can suppose that the remembered objects are ordinary things: birthday parties, trees.

But Mary may imagine the party that (as we say) she will never have, and she may imagine a tree without imagining any particular tree. Evidently, we can bring these cases

under our wing by supposing that when Mary imagines, or remembers, the party for her 33rd birthday, she imagines, or remembers $\hat{(\lambda e)}e=\text{the party for her 33}^{\text{rd}} \text{ birthday}$, and that when she imagines a tree (but no particular tree), she imagines $\hat{(\lambda x)(\exists y)\text{tree}(y) \ \& \ y=x}$; that is, the property of being a thing x that is identical to some tree. It does not seem possible to remember a tree without there being a tree that one remembers (even if one does not remember what tree it was). But the factive postulate, applied to this case, has only the consequence that if Mary remembers a tree, then there are trees, and not also the consequence that there is a tree of which her memory is a memory. But I leave this further matter aside for the present discussion.

4. Immunity to Error Through Misidentification: A Characteristic of PRO

The phenomenon of immunity to error through misidentification is widely appreciated, following Shoemaker (1968) and subsequent work. Recent discussions to which I will advert include Pryor (1999) and Campbell (1999). In this section I will sketch the phenomenon, and show that it applies to the contexts that we have been considering, of memory, imagination, and others, with respect to the positions marked by the understood subject PRO; but also, and more tentatively, I will show that it does *not* apply to other subjects, including the reflexive, the emphatic reflexive, and even the first-person pronoun itself. A closer examination of the phenomenon follows in section 5.

Put in terms of language, the phenomenon of immunity to error through misidentification arises for certain circumstantial reports of experience in which one cannot sensibly wonder whether it is oneself that plays a given role in what is reported. Suppose (to use an example of John Campbell's) it seems to me, on the basis of present perceptual experience, that I hear trumpets. I might be mistaken. It may not be trumpets

that I hear. If I am mistaken, I have made an error due to misidentification; misidentification, that is, of the object that is the source of my experience. I might ask myself whether I am so mistaken. But I cannot sensibly ask myself whether I am mistaken *just* in thinking that it is *I* who hear trumpets. In this sense, there is immunity to error through misidentification.

Similarly, as in a famous discussion of Wittgenstein's, if I feel that I am in pain, although I might conceivably ask whether it is really pain that I am in (perhaps others would consider it merely mild discomfort), I can't ask whether it is *I* who am in pain if anybody is, or think correctly that somebody is in pain, and wonder whether it is me.

What I have just put in terms of language may also be put in terms of the properties of occurrent thoughts; indeed there would seem to be nothing to distinguish these descriptions of the phenomenon except for the manner of presentation; or so I shall assume in what follows.

What is the reason for immunity to error through misidentification in the case of thinking, on the basis of a present perception, "I hear trumpets?" I will assume it is this: that when I am in the relevant perceptual state, what I think is *that the subject of that state hears trumpets*. Hence, there can be no question of my *identifying* myself as the subject of the state. I enlarge upon this thesis below.

There are, as Pryor notes, two distinguishable cases of error through misidentification, and correlatively of immunity to error. In the first, or as I shall call it the simple case of immunity, one is immune from the following error in believing $F(a)$: that, knowing that $F(b)$ and believing, mistakenly, that $a=b$, one believes on those grounds that $F(a)$. In the second, or as I shall call it the generalized case, one is immune from the error, in

believing $F(a)$: that, knowing that something is F , and believing, mistakenly, that it is a 's being F that is responsible for its being the case that something is F , one therefore believes that $F(a)$.

Illustration of the simple case: John sees a woman at some distance, and sees that she is smoking a cigarette. Identifying the woman as Mary, he comes to believe that Mary is smoking a cigarette. The identification is mistaken, and so there is an error due to misidentification, even if, unbeknownst to John, Mary is indeed off somewhere smoking a cigarette. Illustration of the generalized case (from Pryor): John smells the smell of skunk in his back yard. Going outside, he sees a skunk, and concludes that it is that skunk that is (wholly or partly) responsible for the smell; but in fact the source is (entirely) another skunk.

The above characterizations of errors of identification, and of immunity to error, are rough in various ways. For one thing, they leave the notion of identification itself obscure. We often pass from the knowledge that $F(b)$ to the belief that $F(a)$ on the basis of the just-acquired belief that $a=b$; but only some such cases are cases of identifying b as a . Knowing as I do that Fischer played the best move in position **P**, and having on the basis of my own analysis concluded that P-K6 is the best move, I come to believe that Fischer played P-K6 in **P**; but I do not thereby identify P-K6 as Fischer's move. Not all occasions in which one comes to believe an identity are identificatory of anything. Furthermore, although it is natural, as in the core cases of perceptual identification, to think of identification as identification of objects pure and simple, and therefore of the belief $F(b)$ that one has prior to identification as being *de re*, as a general rule this is too restrictive, I think: when, looking out of the airplane window, I identify yonder snowy

expanse as Mont Blanc, no doubt I have beliefs about Mont Blanc itself; but it is also crucial to my so identifying it that it was presented to me as a snowy expanse glimpsed from an airplane window.

Passing over the further subtleties that would be required precisely to delimit errors due to misidentification, and even with the limited apparatus sketched to this point, we can now ask, as Shoemaker (1970) asked, whether immunity to error through misidentification arises in the case of ordinary memory, as it does in perception. That memory of events is subject to errors of misidentification is evident. Suppose that it seems to me that I remember Mary walking through the streets of Oxford. I describe the scene: certain building, cobblestones, etc., and Mary walking along. You observe, however, that the buildings and streets I am remembering are obviously in Cambridge, not Oxford, and from my description you recognize that it is Alice I am talking about, not Mary. I stand corrected. As you might say to me: you remember something all right, but it is not Mary walking through the streets of Oxford, but Alice walking through the streets of Cambridge that you remember. I have made a (double) error of misidentification: everything is in place for knowledge by memory of past events, except that certain objects have been misidentified.

Suppose I seem to remember walking through the streets of some city or another. I might again be corrected about which city I am remembering walking through, or informed that the city I am remembering walking through is in fact Cambridge, not Oxford: but can I be mistaken just in thinking that it is *my* walking through the streets of Cambridge that I remember, or remember someone's walking through the streets of Cambridge, and wonder whether it is me?

Such errors do not at first appear credible: but Shoemaker considers the possibility that some sort of “false memory” has been implanted in me, so that it seems to me that I remember walking through the streets of Oxford, whereas the experience that is responsible for the memory is either nonexistent, or the experience of someone else, functional bits of whose brain have been somehow transplanted into mine. In my state, I experience what Shoemaker calls “quasi-memories;” that is, what appear to me to be genuine memories of mine (and might actually have originated in the memories of someone else) but which are not memories of mine anyway, even if they could by a stretch be called memories, deriving as they did from someone else’s experience.

Putting quasi-memories (and quasi-imaginings as well) aside, we appear to have immunity to error through misidentification in the case described, of seeming to remember walking through the streets of Oxford. But I am now going to observe that, when it comes to the gerundive complements that we have been considering, the problem of their first-personal character shows up when we contrast PRO subjects with the first person itself.

5. Ways of Remembering and Imagining.

Suppose that we form a small party, agreeing that we will call on John and encourage him to finish his thesis by July. Having cornered John, we explain how he should really be prudent given his scholastic and financial circumstances, and so forth. After the session, I try to remember whether we merely hinted around the subject, or whether it was explicitly said to John that he should finish his thesis by July. After a time, I might remember someone saying to John that he should finish his thesis by July; but I don’t remember whether it was I who said it. Your memory for the occasion is better than

mine, and you do remember my saying it; and you tell me so. I draw an inference as follows:

(21) I remember someone saying John should finish his thesis by July;

In fact, as I am now assured, it was I who said it; therefore,

I remember my saying John should finish his thesis by July.

The reasoning seems to me impeccable. But (22) does not follow from the premisses of (21), and is indeed obviously false:

(22) I remember saying John should finish his thesis by July

However, the only difference between (22) and the true conclusion of (21) is that the first-person pronoun has been replaced by PRO.

A certain intuition about this case seems clear enough. When I remember (what turns out to be) my saying that John should finish his thesis by July, I remember this through my recollection of the words in the air, which turn out to have been put there by me, though I don't remember that. When I fail to remember saying that John should finish his thesis by July, that is a failure that would be remedied if I remembered my saying that as an act of mine.

But suppose I do come to remember saying that John should finish his thesis by July. It cannot (and should not) be inferred that the event that I then remember is different from the event that I remembered earlier, through remembering hearing the words in the air. There is just one event in question, namely my saying John should finish his thesis by July. I might remember this event (as indeed I remember people, cities, and other things) in any of various ways; but these different ways do not translate into different

remembered events in the complements of the conclusion of (21), on the one hand, and (22) on the other.

And, finally, there is this case: I remember a certain person's (my) saying that John should finish his thesis by July, but not saying it; that is, I remember hearing the words that, given that I recognize my own voice, I know could only have come from me, but I do not remember their having done so. Again there is just one event. But it can be remembered in either of two ways, namely as an action that I performed, or as an event that I witnessed. The latter sits well with pronouns, but the former is required of PRO.

Imagination goes along with memory here. Thus (an example modeled on Peacocke (1998, p. 212)) I can imagine my playing *Three Blind Mice* on the piano, but (since I do not play the piano) I cannot imagine playing *Three Blind Mice*; and that is because I cannot imagine it as an action that I myself perform.

As Michael Martin pointed out at the Sheffield conference, the above characteristic of PRO shows up in many contexts. There is, for instance, a difference in the intentions I may have when I intend to stop smoking (i.e., PRO to stop smoking), and when I intend merely that I should stop smoking. The latter intention might be fulfilled, say, by paying someone forcibly to remove cigarettes from my person whenever I am caught with them; but that is not fulfillment of an intention to stop smoking, which can only be done through willful refusal to put a cigarette to my lips and light up.

I have concentrated on remembering and imagining actions, or events that could be actions; but the peculiarities of PRO appear in matters that befall one, as well as what one does. Thus for me to remember falling downstairs, or crying, is to remember those events as processes that I underwent; whereas I might remember my falling downstairs

by remembering how I felt as I picked myself up, or remember my crying by remembering how the tears felt rolling down my cheeks.

Pryor (1999) alludes to remembering things “from the inside,” as he puts it, setting aside other ways of remembering. A useful label to put on the problem, which might be translated into the present setting by saying that if in reports of remembering events, as indicated by gerundive complements, there is no semantic difference between the first person pronoun and the understood subject PRO, then there can be no semantic difference between (22) and the conclusion of (21), contrary to fact. Indeed, there is more: the conclusion of (21) seems to be subject to errors of misidentification, at least given the right context. Remembering as I do someone saying John should finish his thesis by July, I might, in virtue of remembering the tone of voice that I heard, ascribe authorship of this event to you by mistake, or to myself by mistake; but that scenario seems impossible with (22).

In the connection with immunity to error through misidentification I have spoken thus far of memory; but examples with imagination come to mind as well. There is an intuitive difference between (23) and (24):

(23) Mary imagined herself flying through space (although she didn't realize it was she herself who was flying through space).

(24) Mary imagined [PRO flying through space] (although she didn't realize it was she herself who was flying through space).

(23) is possibly true; (24) is contradictory, or so it would appear.

If the above is correct, then we have isolated a case wherein PRO is, so to speak, more first-personal than the reflexive forms, and even more first-personal than the first-person

pronoun itself. It is more first-personal in two respects: (i) unlike the first-personal forms, it is immune from error through misidentification; and (ii) the use of PRO always brings in an "internal" dimension to the way what is remembered or imagined is apprehended. These phenomena point up the reality of the *de se* phenomenon more radically than the ordinary contexts of propositional attitude, or epistemic states. In the next section, I offer a way of understanding them.

6. The Semantic Contribution of PRO.

A solution to the problems posed thus far in our discussion should have the following properties:

- (a) It should state what is peculiar about the semantic contribution of the understood subject PRO;
- (b) This contribution should be such as immediately to characterize the distinction between *de se* and *de re* interpretations;
- (c) It should likewise imply immunity to error through misidentification;
- (d) It should do this in such a way that immunity to error through misidentification in the classification of ordinary perceptual experience follows along as a special case; and
- (e) It should explain the distinction between the "internal" dimension signaled by PRO, and the possibly only "external" dimension expressed by ordinary pronouns and reflexive forms.

In short, as I see it, the problems of the *de se*, immunity to error through misidentification, and the rest, call for a solution in terms of logical form. The solution, moreover, should explain the *grounds* upon which *de se* interpretation, and immunity to error through misidentification, arise.

Now, it is obvious that immunity to error through misidentification must arise, not because identification is infallible, but because there is no question of identification at all.

If we ask what it is for a position, call it π , in a construction

... π ...

not to require identification within that construction, there is a case that at once presents itself, namely the case where π is a variable, free within the context ... π ..., but bound from outside it; and more generally where π is a functional context $f(\pi')$, π' a variable free inside but bound outside the context, and f interpreted by a function that is given once the context surrounding ... π ... is given.

In Higginbotham (1995) I argued that contexts of the above shape were crucially involved in locating events with respect to one's own position in time. So in saying (or thinking), as in A.N. Prior's famous example, "Well, *that's* over," the event of which one is thinking that it is over is thought of as coming before one's thinking that. Thoughts of this type were called *reflexive*, the general form being

$$\varphi(e) \ \& \ R(e,e')$$

where φ is some condition on e , and R expresses a temporal relation between e and the episode e' of thinking or saying. Suppose, for instance, I think, or say on the basis of perception, that a dog just barked. The logical form is then (25):

$$(25) \ (\exists e') \text{ think/say}[I,e', \wedge(\exists e)(\exists x) (\text{dog}(x) \ \& \ \text{bark}(x,e) \ \& \ e \text{ just before } e')]$$

I cannot then ask whether I am right in having located in the near past any barking that would make my belief or utterance true; for it is given in the thought itself that it belongs to *my* near past; i.e., just before my thinking or saying.⁹

Perhaps my last statement is somewhat too strong. Perhaps, even with the alleged barking still ringing in my ears, I have misestimated the lapse of time between my perception and my thought; and how much before is just before anyway? However this may be, I cannot ask myself whether I am mistaken just in thinking that it was in *my* past that a dog barked.

Following the analogy of reflexive thoughts with respect to time, suppose that we identify as the peculiar semantic contribution of PRO that it presents the subject as *the subject (or experiencer) of the event or state e* as given in the higher clause, or $\sigma(e)$ for short. This conjecture immediately gives for (4), repeated here, the logical form (26):

(4) John/Each man expects PRO to win.

(26) (For John/each man= x) $(\exists e)$ expect $[x, e, \wedge(\exists e') \text{ win}(\sigma(e), e')]$.

The intended contrast is with (27):

(27) (For John/each man= x) $(\exists e)$ expect $[x, e, \wedge(\exists e') \text{ win}(x, e')]$.

The complement of the latter is one interpretation of (1), repeated here:

(1) John/Each man expects (that) he will win.

an interpretation that, on the hypothesis under consideration, is not available for (4). But this complement is not immune to error through misidentification. For suppose that I expect, indeed know, that the contestant who trained hardest will win the competition I have entered; but I do not know, and I know that I do not know, whether I myself am that contestant. Then, for all I know, it is I myself whom I expect to win. But I also know that it is not true that I expect to win. Later, but still before the competition, I come on the basis of certain evidence to identify myself as the person who trained hardest.

Holding fast to my previous knowledge, I now come to believe that I myself will win,

and therefore to expect that I will win, and to expect to win. My first expectation, that I will win, is subject to error through misidentification: if the person who trained hardest turns out, despite my evidence, to be other than me, then I have made such an error. But my second expectation, namely my expectation of winning, is not subject to such an error, despite the fact that I have this expectation only because I expect that I will win, an expectation that involves the possibility of error.

The informal considerations just given are decisive against construing (4) as merely having the logical form (27), provided indeed that we take it that the reflexive pronoun *myself* can only be taken up as a bound variable, which is precisely what the proposal that there will be an intervening mode of presentation denies. However, the premise that the complement of (1) on some construal is not immune to error through misidentification still stands. The situation as I conceive it is the following. When I expect that I, or I myself, will be F , my expectation is subject to error through misidentification if my grounds for it are the possibly mistaken belief that I am the G , and the knowledge merely that the G will be F . Thus the complement of (1) is subject to such error, where taken up as shown in (27). Knowing as I do that I am the experiencer of the state e of that expectation, when I expect that I myself will be F , I also expect to be F ; hence anything that supports or undermines the first expectation will support or undermine the second. However, whereas I can sensibly ask myself whether, after all, I have identified myself correctly in expecting that I will win, I cannot ask myself whether I have identified myself correctly in expecting to win.

The considerations just given are intuitive. Indeed, so long as we hold to the thesis that the subject or experiencer $\sigma(e)$ of a state e cannot fail to be known as the thing x that is

that subject or experiencer, the link between, for instance, (28) and (29) cannot be broken:

(28) I expect that I (myself) will win.

(29) I expect to win

There is, nevertheless, a difference between them. The thought

$$\wedge(\exists e') \text{ win}(\sigma(e), e')$$

is distinct from

$$\wedge(\exists e') \text{ win}(x, e')$$

(for given values of x and e), even if they are not intensionally different; or so I would submit. The peculiar contribution of PRO, an optional contribution of pronouns and (depending upon the context) perhaps of reflexive forms as well, is that it picks up the subject, the "controller" in familiar linguistic terminology, through the role that it plays with respect to the superordinate or matrix predicate, in this case *expect*.

Consider the application of the above view to the classic case, Castañeda's amnesiac war hero, where the data are that (5) but not (6) may be said truly:

(5) The war hero thinks that he is a hero.

(6) The war hero thinks that he himself is a hero.

Suppose we allow that (5) is true, with the logical form (30):

$$(30) \text{ (For the war hero}=x) (\exists e) \text{ think}[x, e, \wedge(\exists e') \text{ hero}(x, e')]$$

but take (6) as false, with the logical form (31):

$$(31) \text{ (For the war hero}=x) (\exists e) \text{ think}[x, e, \wedge(\exists e') \text{ hero}(\sigma(e), e')]$$

Are the thoughts attributed in the two cases intensionally different? In other words, could anyone other than the war hero have been the subject or experiencer of that war hero's

individual state? This is not a trivial question; but I shall proceed here on the assumption that the answer is negative.¹⁰ If that is so, then there are no grounds upon which the intensionally individuated contents of (30) and (31) may be distinguished: they will coincide in truth value in truth value in any actual or counterfactual situation. Even if they do coincide, it does not follow that they are the same thought (but for a construal that would make them intensionally different, see section 8 below.)

Thus far I have concentrated upon one aspect of PRO, namely its immunity to error through misidentification. But there is also the "internal" aspect of PRO to be considered, an aspect that is not brought out on the view given thus far.

In the picture that we have been using throughout this discussion, the bare predicates of a language are to be thought of as classifiers of events, and their arguments are selected as participants in the events so classified. The relation of a given participant to an event -- agent, or patient, or whatever it is exactly --- is, to use a terminology that is common to a variety of perspectives, and even conceptions of semantics, the *thematic role* that the participant bears to that event. Now, without delving into the variety of linguistic and metaphysical questions that have been raised about thematic roles, we may distinguish: (i) the case where an object α in a predication $\varphi(e, \alpha)$ is given through some description external to its participation in e ; and (ii) the case where α is given simply as $\theta(e)$, where θ expresses the thematic role that α bears to e .

I suggest that the "internal" aspect of PRO results from that element's being considered, either exclusively or in conjunction with its value as given by its antecedent, as the bearer of some thematic role that its syntactic position selects for the bare predicate with which it is in construction. On this hypothesis, for example, the interpretation of the subject in

PRO falling downstairs

is as in

falling downstairs($\theta(e), e$)

where $\theta(e)$ expresses the relation, something like *undergoer*, that a thing falling downstairs bears to an event of so falling.

The above suggestion accords with Michael Martin's observation. When I intend to stop smoking, I intend that I am the agent of my stopping smoking; but when I merely intend that I stop smoking there is no implication of my agency in the stopping, any more than there would be on the part of the automobile in *the car stopped moving*.

In Martin's case and others, it will be a primitive property of PRO that it is construed as the bearer of one or another relation to events as classified by the predicate it appears with. Supposing this property given, we can express the difference between (22), repeated here, and the conclusion of (21), here given as (32):

(22) I remember saying that John should finish his thesis by July.

(32) I remember my saying that John should finish his thesis by July.

For (22) to be true I must remember the event in question *as* something of which I was the agent; that is, the property of events that I remember puts me in the picture as *agent of e*. That requirement does not obtain in (32).¹¹ Likewise we can distinguish, I believe, minimal pairs such as those in (33)-(34):

(33) I remember/imagine [PRO/myself falling downstairs].

(34) I remember/imagine [PRO/myself crying].

We can now join Martin's observation together with the suggested explanation for immunity to error through misidentification, as follows. When one remembers or

imagines *PRO being F*, then the linguistic element PRO is distinguished in two ways: (i) by being understood as the thing $\sigma(e)$ that is in the state e of remembering or imagining itself; and (ii) by being at the same time understood as the bearer of the thematic role $\theta(e')$ as determined through the selection for the subject of the predicate $F(e')$ (so in general that $\sigma(e)=\theta(e')$ is presupposed). I shall abbreviate this dual role of PRO by ' $\sigma(e)$ & $\theta(e')$ '. With PRO so understood, (22) for instance comes out as (35):

(35) $(\exists e)$ Remember $\{I, e, \lambda(e') [\text{say}(\sigma(e) \ \& \ \theta(e'), \text{that John should finish his thesis by July}, e')]\}$.

We thus bring out both the fact that, if (22) is true, then the subject cannot make an error of misidentification, and that what is remembered is remembered *as* an action performed.

In concluding this section, I show how the Fodor-Thomson example comes out on the view advanced here. We are given the premisses (36) and (37):

(36) Only Churchill gave the speech.

(37) Churchill remembers PRO giving the speech.

And we aim to derive (38):

(38) Only Churchill remembers PRO giving the speech.

The proper name *Churchill* may be rendered as a quantifier, *for $x=Churchill$* , with obvious semantics, and the subject *only Churchill* of (36) and (38) is understood as quantificational as well, where in general, where *Only x such that A are B* is true if and only if all B are A . We need also to assume, perhaps as a matter of presupposition, that *Only Churchill is F* licenses *Churchill is F* .

With these assumptions, (36) will have a logical form equivalent to (39):

(39) $(\exists e)$ give(Churchill, the speech, e) & $(\forall x \neq \text{Churchill}) \neg(\exists e')$ give(x , the speech, e').

For (37), taking PRO as merely a bound variable, we have (40):

(40) (For $x=Churchill$)($\exists e$) Remember[$x,e,\wedge(\lambda e')$ give(x ,the speech, e')].

and it will be sufficient to derive (41):

(41) ($\forall x\neq Churchill$) $\neg(\exists e)$ Remember[$x,e,\wedge(\lambda e')$ give(x ,the speech, e')].

Supposing on the contrary

$x\neq Churchill$ & Remember[$x,e,\wedge(\lambda e')$ give(x ,the speech, e')]

we have, by the factivity of *remember*, the consequence

$x\neq Churchill$ & ($\exists e'$) give(x ,the speech, e').

contradicting (39).

The assumption crucial to the above derivation of (38) is that PRO be "captured" by the quasi-quantifier *only Churchill*. Because, as remarked above, it is not true that the premisses (36)-(37) lead to the conclusion (42), we must say something about the interpretation of the pronoun *his*, even where it is anaphoric to *Churchill*:

(42) Only Churchill remembers his [Churchill's] giving the speech.

It is sufficient to suggest that the pronoun, unlike PRO, can simply go proxy for the name, that being enough to block the implication.

The simple treatment of PRO as obligatorily a bound variable is sufficient to explain the Fodor-Thomson observation. On the more complex treatment suggested in this section, the premiss (39) is replaced by (43), and the desired conclusion (41) by (44):

(43) (For $x=Churchill$)($\exists e$) Remember[$x,e,\wedge(\lambda e')$ give($\sigma(e)$ & $\theta(e')$,the speech, e')].

(44) ($\forall x\neq Churchill$) $\neg(\exists e)$ Remember[$x,e,\wedge(\lambda e')$ give($\sigma(e)$ & $\theta(e')$,the speech, e')].

The derivation of this conclusion is only marginally more complex, requiring as it does just the added point that, where e is a state of Churchill's remembering, and e' is any event of giving the speech, Churchill himself is indeed $\sigma(e)$ and $\theta(e')$.

But could a person x be in a state e of imagining being F without *recognizing* that $x=\sigma(e)$, the subject of the property of events being imagined? If this can happen then perhaps, as tentatively suggested in Campbell (1999), that person would have thoughts of which he did not seem to himself to be the author. In any case, it seems safe to assume that any such condition would be pathological.

This concludes my discussion of the contribution of PRO, and the perspective that the account given here affords upon the semantic data that suggest it. If the account is correct, then PRO, and I will assume at least some reflexive and emphatic reflexive elements, are distinguished by having special interpretations, as in the third of the three options that I presented in section 1 above. In the next section I take up the question of the other options directly.

7. Alternatives Explored.

As we know from the story surrounding (5), repeated here, in any situation in which the war hero's beliefs are true, he is not reading about himself.

(5) The war hero thinks that he is a hero.

In the world according to the war hero, as we might say, there is on the one hand him, and on the other the person whose exploits he is reading about. It is this reflection that encourages the idea, expressed in the first hypothesis of section 1 above, that the apparent existence of contexts that are simply and strictly *de re*, but not first-personal, is after all

an illusion: in all such cases, there is an intervening "mode of presentation," which would in the present instance be brought out in the logical form (45):

(45) (For the war hero= x) $(\exists e)$ think $[x, e, \wedge(\exists e')$ hero(the person x is reading about, $e')$]

This mode of presentation contains the war hero himself as value of x , and in that sense is *de re*. But if all that is going on in the example is that the thought attributed in the complement of *He thinks he is a hero* has a structure that is not revealed in the linguistic material alone, then it would seem that the first-personal or *de se* thought has no peculiar status to be explained, being simply the *de re* case unadorned.

As I understand it, Stalnaker's (1981) discussion of indexical belief could lead to a similar conclusion, transposed, as it were, into a different key. For Stalnaker, the background story (and its presuppositions, including the presupposition, if it is present, that the war hero is in fact reading about himself) serves, not to fix a mode of presentation, but rather to determine which alternative situations of utterance of *He (the war hero) is a hero* are to be taken into account. So consider all of those situations w such that the war hero is in w reading about a person other than himself. In those situations, the speaker's embedded subject *he* refers, not to the war hero, but to that other person. Moreover, only such situations are consistent with what the war hero believes (assuming that the speaker's presuppositions are satisfied, as they would be if the war hero thinks that the person he is reading about is not himself). The content that the speaker attributes to the war hero's beliefs is then just the content

$\wedge(\exists e')$ hero(the person x is reading about, $e')$]

for the war hero as value of x , with the presuppositions mentioned. This content is constructed, not by first filling in a mode of presentation, and then taking the expression

so filled in around the counterfactual situations w , but rather by taking what Stalnaker calls the *diagonal* proposition, obtained by determining, for each w , what the truth value of *He (the war hero) is a hero* would be as said by the speaker in w .

It is pretty clear that, for any *particular* way of filling in the notion of a situation consistent with what the war hero believes, we can construct a mode of presentation that will, when supplanted for x in the complement clause, deliver along the horizontal just what the diagonal proposition delivers along the diagonal. In the case under consideration, we supplant x with *the person x is reading about*, for instance. We can also proceed in the other direction, supplanting *the person x is reading about* in with the simple variable, but taking along the diagonal the worlds w and the utterances, "He is a hero," where the pronoun refers to whoever it is would have been meant by its utterance in w . There are perhaps differences between these approaches, centering on the point that in the method of supplementation we ascribe a mode of presentation for which, in effect, we regard what is said as elliptical, whereas in Stalnaker's diagonal method we supply instead a conception of what is merely implicated by the speaker; but I shall not dwell on these differences here. In either case, we are taking those *de re* attributions that are not *de se* (that is, those that would sustain a use of the bare pronoun, but not of the emphatic reflexive, or of PRO) as not meant literally. And indeed, how could they be? For, after all, we know that the amnesiac war hero believes he is not the person whom he is reading about, and that he is ready to say as much, and in the first person. Yet his beliefs cannot be internally faulted, or rectified by mere cogitation. So if we are to have a conception of how the world would be if his beliefs were true, we cannot find it upon a

combination of the literal interpretation of what he is reported to believe, on the one hand, and his own avowals on the other.

I suggest that the considerations just rehearsed do show, either through Stalnaker's route or through that suggested by invoking modes of presentation, that so long as we wish to have a conception of "how the world would be" if a believer's beliefs were true (assuming of course that there *is* such a way); or, to put it only slightly differently, but to the same effect, so long as we need to have a conception of what the notation we use to ascribe thoughts to ourselves or to others is a notation *for*, we should regard the true report in (5), as it might be called a report of a true belief in the absence of proper identification (in this case of the person the war hero is reading about with himself) as expressing something other than just what is given on its face (taking the proposition expressed in the usual way, and not "on the diagonal").

The tempting conclusion is that (6), repeated here, is the true *de re* report:

(6) The war hero believes that he himself is a hero.

The difference between (5) and (6) cannot be left just at that, however; for we need also to explain why (6) cannot be asserted under the same circumstances as (5), or why PRO should contrast even with the emphatic reflexive in the contexts of imagining, for example.¹²

For these reasons also in part, the construction of Lewis (1979) and Chierchia (1990) doesn't appear to satisfy the demands of the case. Recall that on Chierchia's view the positions critical for *de se* interpretation are abstracted over, so that the object of the thought that one is oneself a hero is: λx hero(x). There is no evident reason why immunity to error through misidentification should be associated with the abstracted

position, or why, to use Pryor's term, true reports of imagining and remembering with understood, PRO, subjects should always indicate imagining and remembering "from the inside." Furthermore, as suggested by David Kaplan in his Bielefeld lecture, 1995, and subsequently in a Gareth Evans Memorial Lecture at Oxford, Lewis's strategy would generalize, implausibly, to other embedded indexicals, words like *now*, *here*, *today*, and the like. Finally, as noted in Landau (2000: Chapter 2), the account faces difficulties in accounting for the phenomenon he calls *partial control*, where the understood subject is plural, but its antecedent is singular; I give one example of this phenomenon below.

Granting that the understood subject must be a real subject in the logical form, it may be asked whether the unadorned *de re* already exhibits the property of immunity to error through misidentification; that is, whether the recourse to $\sigma(e)$ exploited above is really necessary. The question arises, because the apparent cases of *de re* that were not *de se* all seem upon closer examination, either through the invocation of hidden modes of presentation or through Stalnaker's suggestion, to be understandable as carrying more semantic baggage than the simple *de re* would suggest. This diagnosis applies even to cases of the embedded first person, as we saw above: when I remembered (what, from my point of view, turned out to be) my saying that John should finish his thesis by July, I remembered it through a conception of the subject of that action *as* some producer of words in the air, the bearer of a certain tone of voice, or something similar. Thus when the *de se* is missing, there is always, so to speak, a fallback to some underlying content that is not strictly *de re*.

Now, this last reflection, assuming it correct, of course does not show that the *de re* thought is not there, but only that, if it is there, it is there in virtue of something else,

something that is not purely *de re*. But this phenomenon holds quite generally for embedded indexicals. If I wonder, now, quite out of the blue, whether it was sunny yesterday in Los Angeles, there is no question of my misidentifying yesterday, that day being given as *the day before my wondering this*. But if I wonder whether it was sunny yesterday in Los Angeles because I wonder whether it was sunny on May 27th, believing, mistakenly, that the 27th was two days ago rather than yesterday, then I make an error due to misidentification (I have not properly grasped what I am wondering about). That fact alone does not show that I am not wondering about yesterday, the day itself. Likewise, in the setting due to Castañeda, the war hero's belief that he is a hero can be a belief about himself, purely *de re*, despite his misidentification or lack of identification. That attribution need not be withdrawn just because another stands behind it. Our question therefore remains.

For the purposes of this article, I shall leave that question unanswered. In any case, it seems to me that immunity to error through misidentification calls for a thing's being given through material supplied from a higher context, as in the example just presented of wondering whether yesterday was sunny in Los Angeles. But I must leave this matter for another occasion.

8. Links to Formalization.

Much of the recent literature on the *de se* has taken as its point of departure the account of demonstrative and indexical expressions developed originally in Kaplan (1977). Schlenker (forthcoming) proposes a modification of Kaplan's Logic of Demonstratives that is designed to incorporate the *de se* by admitting context-sensitivities of the type that Kaplan dubs "monsters." Assuming the account of the *de se* advanced above, I examine

the extent to which it can be seen, within the formal perspective of that modification, as a further specification of it.

Recall that in Kaplan (1977), truth in a model is defined relative to pairs (c,w) , where c is a context, and w a possible world, and that the content of a sentence in a context c is the set of possible worlds in which it is true. Contexts, whatever they are, are anyway such that one can pull from them a speaker, a time, and a world of the context, namely the world in which the speaker is conceived of as saying the sentence at the time in question. The system thus involves double-indexing with respect to the possible worlds, as in two-dimensional modal logic, and it therefore permits, in principle, operators that generalize over context worlds as well as worlds of evaluation. Such operators are “monsters,” and Kaplan urged that, although they are definable, they could not exist in our language (or, presumably, human languages generally).

Schlenker (forthcoming) develops a formal account of the *de se* that allows manipulation of the context parameter, specifically the agent of the context (a suggestion to similar effect is found in Israel and Perry (1996)). This allows Schlenker to distinguish, for example, the content of the complement in *x hopes x is elected*, namely the one true in those worlds in which x is elected (independently of the agent), from that of *x hopes [PRO to be elected]*, which comprises those (a,w) in which a is the agent of the context, and a is elected. Assuming as Schlenker does the modal account of the propositional attitudes, *x hopes that p* will be true in a context c in a world w just in case every *context* (not: possible world) compatible with what x hopes is one in which p , where the context comes with a world of evaluation, and also an agent. Thus a context is compatible with what x hopes in hoping that x is elected if and only if it is one in which x is elected; but it

is compatible with what x hopes in hoping to be elected if and only if it is one in which the agent of that context is elected, independently of whether that agent is x .¹³

Assuming the account developed in this article, we can see it as a further specification of Schlenker's method. Supposing as above that the content of the complement subject position in x hopes [*PRO to be elected*] is $\sigma(e) \ \& \ \theta(e)$, where e is the situation of x 's hoping, we may suppose that, although the individual situation e belongs essentially to x , the *kind* of situation that e is could belong to anybody. If, therefore, we allow the instance e of that kind to vary independently, and make it part of the context, we arrive at an account that specifies more precisely, but in the same manner as the parameter of the agent of the context, what it will be for a context to be compatible with what x hopes. Thus, whatever the advantages and uses, or the disadvantages and abuses, of the modal theory of the attitudes, the empirical material given here, if I am right, will supplement the purely formal move that Schlenker suggests.

9. Concluding Examples and Extensions.

The above discussion of PRO has developed a point of view according to which two rather sophisticated concepts are involved in its correct employment: (a) the concept being an agent that thinks thoughts, expressed by ' $\sigma(e)$ ', and (b) the concept of being an undergoer or agent of experience, expressed by ' $\theta(e)$ '. It follows that those who use this form, or the relevant constructions, correctly, or have thoughts that are properly reported using them, must be able to deploy both concepts. But is that really essential to first-person thoughts?

I leave aside the question of reports of the thoughts of creatures for which the question how to take our very expression of the thoughts that we attribute to them, or in what

sense they have thoughts, is itself an issue. The problem is whether there are notions of the self, more primitive or anyway different from those that I have recruited, namely of oneself as a thinker of thoughts and undergoer of experiences, that, whether or not they answer to English PRO or anything else, deserve to be brought into the realm of first-person thoughts. Recall the general suggestion of Peacocke (1981), that first-person thoughts by x contain a constituent [$self_x$] a token of a type $self$ available to all. I am proposing, to use Peacocke's terminology, that one such type is $\sigma(e) \ \& \ \theta(e)$, where e is some particular event, so that the tokens are instantiations of e , where $x=\sigma(e)=\theta(e)$. That does not exclude the existence of other types. At the same time, the phenomena of memory "from the inside," and of complements necessarily interpreted as denoting particular ways of remembering and imagining things, appear to show that these types are not included in our most basic system of speech; unsurprisingly, since we all do conceive of ourselves in the ways required for the account.

In concluding, I consider three points, the first two of which have appeared in the literature, and the third of which will likely have occurred to the reader in passing. The first point is the problem of like belief, due to Lewis, and the second concerns a question raised early on for Castañeda, whether or to what extent first-personal knowledge can be shared. The third point addresses briefly the peculiarities of the distribution of the understood subject, whose semantics has chiefly occupied us.

We owe to Lewis (1979) the case of crazy Heimson, who believes that he --- he himself --- is David Hume. He is ready to say, "I am David Hume," "I wrote the *Treatise*," "I served my King in France, and was denied a chair in Edinburgh," and so forth, and to behave accordingly, whatever that may come to exactly. In Lewis's view, there should be

a sense in which Heimson and Hume believe alike; that is, a sense in which it is not merely true that each conceives himself to be an x such that: x is David Hume (for this would give them different beliefs, a false one for $x=$ Heimson, and a true one for $x=$ Hume), but rather, or anyway also, that each has numerically the same belief as the other. Lewis presented several considerations in favor of this view. Abstracting from these, and from the critical discussion in Stalnaker (1981) and Higginbotham (1991), I consider the question, on the view of beliefs about oneself advanced here, to what extent Lewis's view can be sustained.

It cannot, if I am right, be literally sustained. For the anaphoric possibilities for an element in a complement structure, whether (in the terminology developed here) event-like or proposition-like, are limited to these:

(i) The anaphoric element inherits the reference of its antecedent, giving for

Heimson believes that he (himself) is Hume

the logical form

For Heimson= x , $(\exists e)$ believes $[x, e, \wedge(\exists e')$ identical (x, Hume, e')]

(ii) The anaphoric element takes on the same reference, but now given as $\sigma(e)$,

expressing the relation of the subject to the state that it is in, giving

For Heimson= x , $(\exists e)$ believes $[x, e, \wedge(\exists e')$ identical $(\sigma(e), \text{Hume}, e')$]

(iii) Besides taking on the reference as $\sigma(e)$, the anaphoric element expresses the thematic relation $\theta(e')$ determined by the events classified by the subordinate predicate, in this case identity. Then we have

For Heimson= x , $(\exists e)$ believes $[x, e, \wedge(\exists e')$ identical $(\sigma(e) \& \theta(e'), \text{Hume}, e')$]

In none of these cases will Heimson and Hume believe numerically alike.

On the other hand, I suggest, there is a sense in (ii), and therefore in (iii), in which Heimson and Hume do believe alike. Let α and β be two belief states satisfying

$$\wedge(\exists e') \text{ identical}(\sigma(e), \text{Hume}, e')$$

Does $\alpha=\beta$? Well, we know that $\alpha=\beta$ if and only if $\sigma(\alpha)=\sigma(\beta)$; but there is as it were nothing *within* α and β to distinguish them: they are states of believing that one is David Hume, possibly the same, possibly different, depending upon whether the bearers $\sigma(\alpha)$ and $\sigma(\beta)$ of those states are the same or different. In this sense, they are alike.

For an analogy, consider two collapses of bridges. Are they the same collapse, or different collapses? Ignoring time, we may say that they are the same if the bridges are the same, otherwise different. But insofar as they are given merely as collapses of bridges, there is nothing to tell them apart. It is not just a matter of one-to-one correspondences. It is true that two squares of natural numbers are identical just in case the numbers of which they are the squares are identical; but the numbers are already given as the same or different independently of this consideration. With events and states, I am suggesting, this is not so (and perhaps this counts in some degree against thinking of events as objects). Events of the same kind are discriminated through their participants, rather than intrinsically. I conclude, then, that we have a strong sense in which Heimson and Hume believe alike.

In this article I have been defending a version of Castañeda's original view, that there is something special about first-personal reports of thoughts. An issue for this view that was raised early on, in Kretzmann (1966), was: could anyone other than x know what x knows when x knows something first-personal about x ? Could even God know it?

Reviewing the discussion, Perry (1983) cites Castañeda's response (Castañeda (1968)), which I paraphrase as follows.

Suppose that John is in the hospital, and he knows in the ordinary way that he himself is in the hospital. What he knows is a fact or true proposition, p say. Evidently, many others may know, and God in particular is bound to know, that John knows p . But p itself is a consequence, and a pretty trivial one at that, of what they then know. We would be, to put it mildly, surprised to learn that whereas x knows that John knows p , x does not know p ; and such incapacity is out of the question where God is concerned. So we may assume that x knows p , and it follows in particular that God knows p .

The above reasoning is all very well; but it does not apply to the construction of first-person knowledge given here. Suppose we had taken up *John knows that he is in the hospital* merely as the proposition expressed by (46):

(46) For $\text{John}=x$ $(\exists e)$ knows $[x,e,\wedge(\exists e')$ in-hospital (x,e')]

From (46), there follows, knowledge being knowledge of truth, (47), where α is the person John:

(47) $(\exists e')$ in-hospital (α,e')

Thus Castañeda's response to Kretzmann would be vindicated. But we are supposing that the peculiarly first-personal reading of our target sentence is given by (48):

(48) For $\text{John}=x$ $(\exists e)$ knows $[x,e,\wedge(\exists e')$ in-hospital $(\sigma(e) \& \theta(e'),e')$]

and from knowledge of what (48) expresses we would have to be able to infer (49):

(49) $(\exists e')$ in-hospital $(\sigma(\beta),e')$

where β is the state of knowledge that John is in. But this is too much to ask: from the fact that I know that there is some state e or other such that John knows that its subject

(him, John) is in the hospital, I can't derive the knowledge, with respect to any state β at all, that the subject of that state is in the hospital; to suppose otherwise would be as wrong as to suppose that merely by knowing that someone or other knows that he is in the hospital I could come to know, with respect to a particular person, that that person is in the hospital.¹⁴

I suppose we should concede that it would be different with God. For (49) expresses one of the things that there is to be known; and God knows everything. That case apart, we have a further dividend of the account suggested here: for John, being himself *in* the state β , can be expected to know what (49) expresses; but others in general will not, even if they know what is expressed by (48).

Of course, it does not follow that no one other than John and God can know what is expressed by (49). On the contrary, there is at least so far no intrinsic reason that one cannot have knowledge of John's state of knowledge as one has knowledge of other events and states. Events and states, I have suggested, have their participants essentially, so that no one other than John can *be* in the state β ; but nothing about knowledge of β follows from that.

We have at this point bumped up against some familiar philosophical questions of privacy, even perhaps "privileged access," and so forth, which I will not consider here. However, it may be observed that the construction lends credence both to the thesis that there is a special way in which one is given to oneself when one knows in a first-personal way that one is in the hospital (or driving a car, or looking around the room, etc.), *and* that this special way is, *mutatis mutandis*, available to all.

A final example. In this discussion I have concentrated upon the properties of understood subjects, construed here as PRO. It happens that this particular element occurs only in one place in English and other languages, namely as the subject of a clausal or nominal complex lacking a tense.¹⁵ We can, however, so to speak tease the effects of PRO into another position by taking it as the antecedent of a reciprocal construction, as in (50):

(50) *a* and *b* remembered/imagined [PRO kissing each other].

where we are interested in the interpretation according to which each remembered or imagined the reciprocal kissing (and not the one where each merely remembered or imagined kissing the other).¹⁶ Even if the events remembered by *a* and *b* are the same, they are given to them in different ways. For, on the account suggested here, *a*'s memory is given to *a* as of concurrent events *e'* and *e''*, and through a memorial state *e* such that

$$\text{kiss}(\sigma(e) \ \& \ \theta_1(e'), b, e') \ \& \ \text{kiss}(b, \sigma(e) \ \& \ \theta_2(e''), e'')$$

where θ_1 expresses agency, and θ_2 , say, the "undergoing" relation, whereas not only is *b*'s memorial state different, but also the roles are reversed, as in

$$\text{kiss}(\sigma(e) \ \& \ \theta_1(e'), a, e') \ \& \ \text{kiss}(a, \sigma(e) \ \& \ \theta_2(e''), e'')$$

I believe that these consequences are in accord with intuition. If so, then we can so to speak push the properties of PRO into the direct object position, as we can in verbal passives, such as *a remembers PRO being kissed by b*. That language, or English anyway, provides us with only a small window for the direct expression of the peculiarities of certain first-person thoughts is not, therefore, of itself a reason to doubt their existence.

Notes

* A first draft of this paper was presented at the Sheffield conference *Epistemology of Language*, July 2000, Michael Martin commenting. That draft was itself based upon earlier presentations at the University of Oxford, at the University of Michigan, and at the tenth conference on Semantics and Linguistic Theory, Cornell University, March 2000. Presentations subsequent to the Sheffield draft include talks at the University of Siena, the University of London, and the University of California, Davis. I am grateful to my various audiences for their comments, and especially to Carl Ginet, Allen Gibbard, Alessandra Giorgi, Stephen Parkinson, Christopher Peacocke, Philippe Schlenker, Gabriel Segal, and Orsolya Schreiner. Alex Barber and Christopher Peacocke provided very useful comments on earlier drafts; and discussion with Philippe Schlenker and members of the audience at the European Summer School for Logic, Language, and Information in Trento in 2002 helped me to draw the connections between my discussion and certain revisionist movements in intensional logic.

1. The logophors are special anaphors, said to refer to "centers of consciousness." The term has been extended to cases including some English reflexives. For discussion of the syntax and semantics of logophoric forms, see Huang (2000), Schlenker (1999) and (forthcoming), and references cited there.
2. For recent critical discussion, see Landau (2000), especially Chapter 5.
3. The scenarios are of two basic kinds: (a) those in which a subject knows (or believes, etc.) something about herself that it is questionable whether others know, as when I know in the ordinary way that I am standing; and (b) those where a subject knows something about himself, but does not have the knowledge that would be required in order to justify

the use of locutions with reflexive, emphatic reflexive, or PRO (understood subject) arguments, as when he knows (as we say) that he was elected, because he knows that the candidate with the largest war chest was elected, but does not know that he himself was elected, having failed to realize that he himself was the candidate with the largest war chest. Again, see Castañeda for persuasive examples.

4. The case of (4) was already observed in Fodor (1975), as noted below. Chierchia (1990) provided a more systematic taxonomy, concentrating however on propositional attitudes.

5. In English, the ordinary pronominal *him* as in

John remembered/imagined [him going to the movies]

cannot be anaphoric to *John*, just as it cannot be anaphoric in simple clauses such as *John remembered him*. Hence I use the possessive form *his* in (7). The emphatic reflexive *him, himself* in (8) seems to me acceptable and interpretable, if unnatural in comparison, say, to (9) and (10).

6. Fodor used the validity of (12), in conjunction with the Katz-Postal hypothesis, that optional grammatical transformations did not change meaning, to argue against "pronominalization;" that is, against the view that the understood subject represents deletion of an anaphoric pronominal or reflexive element in the course of the syntactic derivation of (11). His thesis translates into the present syntactic system as: PRO is not semantically equivalent to any anaphoric pronominal or reflexive form. I will argue below that the thesis is correct.

7. The "success verbs," as Ryle called them, do not generally admit the progressive: thus I cannot say that I am seeing my watch. When the same verbs are used without a

“success” connotation, however, the progressive becomes possible, as when I say that I am seeing spots before my eyes. Individual dialects diverge here: thus Christopher Peacocke finds the progressive of *see* acceptable where it involves making something out, or discernment, as in *I am not seeing the third letter*, said during an eye test. In my speech, however, this is a far fetch. It is useful, though not of itself enlightening, to put the divergence down to a distinction between expressions ranging over processes and those ranging over momentary transitions, or *achievements* in a common terminology, with my dialect admitting only the latter for epistemic or “success” expressions, including *see*.

8. In these last sentences and in what follows I suppress some distinctions between use and mention. I use italics for mentioned English words throughout, but do not further mark elements already italicized.

9. Thus the experience *e*’ has been “loaded” into the proposition expressed. But isn’t the truth of the proposition independent of that connection, and indeed of whether there was any experience at all? An analogous problem is discussed in Higginbotham (1995), inconclusively; however, I would maintain that the interpretation of the speaker’s word, or thought-constituent, marked by *just* does carry the temporal anchoring expressed in (25).

10. Plainly, however, a full treatment must take a view on the very general question, what properties of events (taken as objects of some sort, following Davidson) are contingent, and what necessary. (In Higginbotham (1995) I suggested that the spatiotemporal location of at least some events was a contingent matter.)

11. As Stephen Parkinson has made me aware, there is at least one case where contexts such as (22) and (32) are not to be discriminated, namely the case where the question is not merely what one remembers, but more stringently what one remembers out of one's own memory, or of one's own knowledge. Thus suppose I am to testify as to what was said to John. I remember someone's saying that he should finish his thesis, but not that it was me who said it. In that case, even if I have become convinced on other grounds that it must have been me who spoke, I cannot agree with an inquisitor's assertion that I do after all remember my saying it; for in the setting of testimony I am supposed to state what I remember without reliance upon external evidence. Cases of this sort are of course not confined to memory of oneself: if I remember the blue Ford at the scene of the crime, and have only since become convinced (rightly) that it was the getaway car, I still cannot respond, "Yes" to the question put at trial whether I remember the getaway car at the scene of the crime.

12. There are important comparative-linguistic questions here. To take one example, note that in English *believe*, unlike *expect*, does not take controlled complements, so that we do not have **John believes [PRO to be a hero]*. The corresponding construction in Italian, namely *Gianni crede [PRO di essere un eroe]* is, however, fully grammatical, and (I am informed) unambiguously *de se*. Corresponding to English *John believes that he himself is a hero* Italian has the construction *Gianni crede che lei stessa è un eroe*; but this, as again I am told, is not necessarily *de se*. From these facts we might draw the cross-linguistic conclusion that it is the understood subject (which, as pointed out for instance in Chierchia (1990), must for languages like Italian be distinguished from the null subject pronoun of a finite clause) that is crucially involved in the *de se* phenomena.

Schlenker (forthcoming) carries these questions much farther, in ways I won't consider here.

13. Schlenker's formal theory involves an extensionalization of Kaplan's Logic of Demonstratives, in which the relevant parameters occupy quantifiable places.

14. Compare the discussion in Perry (1983), p. 88.

15. There has been a lengthy debate in the linguistics literature about whether PRO occurs as the subject of nominals as well as clauses, a matter from which I abstract in the present discussion.

16. The example is a case of "partial control" in the sense of Landau (2000: Chapter 2), in the sense that the subject position of what each remembers includes the other as well as herself. As remarked briefly above, partial control, which brings the *de se* in its wake, threatens the Lewis-Chierchia picture of the subject as disappearing under abstraction.

References

- Campbell, John (1999). "Immunity to Error Through Misidentification and the Meaning of a Referring Term." *Philosophical Topics* **26**. pp. 89-104.
- Castañeda, Hector-Neri (1966). "'He*': A Study in the Logic of Self-Consciousness." *Ratio* **8**: pp. 130-157.
- Castañeda, Hector-Neri (1967). "Indicators and Quasi-Indicators." *American Philosophical Quarterly* **4**. pp. 85-100.
- Castañeda, Hector-Neri (1968). "On the Logic of Attributions of Self-Knowledge to Others." *The Journal of Philosophy* **65**. pp. 439-456.
- Chierchia, Gennaro (1990). "Anaphora and Attitudes *De Se*." Renate Bartsch *et al.* (eds.), *Language in Action*. Dordrecht, Holland: Foris Publications. pp. 1-31.
- Chomsky, Noam (1981). *Lectures on Government and Binding*. Dordrecht, Holland: Foris Publications.
- Fodor, Jerry A. (1975). *The Language of Thought*. New York: Crowell.
- Frege, Gottlob (1918). "The Thought: A Logical Inquiry." Translated by A.M. and Marcelle Quinton. P.F. Strawson (ed.), *Philosophical Logic*. Oxford: Oxford University Press, 1967. pp. 17-38.
- Higginbotham, James (1989). "Reference and Control." *Rivista di Linguistica* **1, 2** (1989). Reprinted with minor revisions in Richard Larson, Sabine Iatridou, Utpal Lahiri and James Higginbotham (eds.), *Control and Grammar*. Dordrecht, Holland: Kluwer, 1990. pp. 79-108.
- Higginbotham, James (1991). "Belief and Logical Form." *Mind and Language* **6, 4**. pp. 344-369.

- Higginbotham, James (1995). "Tensed Thoughts." *Mind and Language* **10**, 3. pp. 226-249. Reprinted in Wolfgang Kühne, Albert Newen, and Martin Anduschus (eds.), *Direct Reference, Indexicality, and Propositional Attitudes*. Stanford, California: CSLI Publications, 1997. pp. 21-48.
- Higginbotham, James (forthcoming). "Tensed Second Thoughts: Comments on Richard." For Aleksander Jokic (ed.), *Time, Tense, and Reference*. Cambridge, Massachusetts: The MIT Press.
- Huang, Yan (2000). *Anaphora: A Cross-Linguistic Study*. Oxford: Oxford University Press.
- Israel, David and Perry, John (1996). "Where Monsters Dwell." In Jerry Seligman and Dag Westerståhl (eds.), *Logic, Language, and Computation*, volume 1. Stanford: CSLI Publications. pp. 1-14.
- Kaplan, David (1977). "Demonstratives." Published with Afterthoughts in J. Almog *et al.* (eds.), *Themes From Kaplan*. Oxford: Oxford University Press, 1989. pp. 481-614.
- Kaplan, David (1995). "Meaning and Use." Gareth Evans Memorial Lecture, University of Oxford.
- Kretzmann, Norman (1966). "Omniscience and Immutability." *The Journal of Philosophy* **63**. pp. 409-421.
- Landau, Idan (2000). *Elements of Control: Structure and Meaning in Infinitival Constructions*. Dordrecht, Holland: Kluwer.
- Lewis, David (1979). "Attitudes *De Dicto* and *De Se*." *The Philosophical Review* **88**. pp. 513-543.

- Lewis, David (1980). "Index, Context, and Content." In Stig Kanger and Sven Öhman (eds.), *Philosophy and Grammar*. Dordrecht, Holland: D. Reidel. Reprinted in Lewis (1998), pp. 21-44.
- Lewis, David (1998). *Papers in Philosophical Logic*. Cambridge: Cambridge University Press.
- Montague, Richard (1960). "On the Nature of Certain Philosophical Entities." *The Monist* **53**. Reprinted in Montague (1974), pp. 148-187.
- Montague, Richard (1974). *Formal Philosophy*. Richmond Thomason (ed.). New Haven: Yale University Press.
- Peacocke, Christopher (1981). "Demonstrative Thought and Psychological Explanation." *Synthese* **49**, 2. pp. 187-217.
- Peacocke, Christopher (1998). *Being Known*. Oxford: Oxford University Press.
- Perry, John (1983). "Castañeda on *He* and *I*." In James E. Tomberlin (ed.), *Agent, Language, and the Structure of the World: Essays Presented to Hector-Neri Castañeda With His Replies*. Indianapolis: Hackett Publishing Company. Reprinted in Perry (2000), pp. 77-100.
- Perry, John (2000). *The Problem of the Essential Indexical and Other Essays*. Expanded Edition. Stanford: CSLI Publications.
- Pryor, James (1999). "Immunity to Error Through Misidentification." *Philosophical Topics* **26**. pp. 271-304.
- Schlenker, Philippe (1999). *Propositional Attitudes and Indexicality: A Cross-Categorical Approach*. Unpublished Doctoral Dissertation, MIT.
- Schlenker, Philippe (forthcoming). "A Plea for Monsters." *Linguistics and Philosophy*.

Shoemaker, Sydney (1968). "Self-Reference and Self-Awareness." *The Journal of Philosophy* **65**. pp. 555-567.

Shoemaker, Sydney (1970). "Persons and Their Pasts." *American Philosophical Quarterly* **VII**, 4. pp. 269-285.

Stalnaker, Robert C. (1981). "Indexical Belief." *Synthese* **49**. Reprinted in Stalnaker (1999), pp. 130-149.

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