ABSTRACT. Roderick Chisholm contrasts semantic theories that presuppose "the primacy of the intentional" with those that presuppose "the primacy of the linguistic". In The First Person he attempts to develop an analysis of first person singular reference that presupposes the primacy of the intentional. In this paper I attempt to develop a semantics of first person singular reference (what I call 'I-reference') that presupposes the primacy of the linguistic. I do three things in the paper. First, I criticize Chisholm's (and Frege's) account. Second, I attempt to answer the general criticism that is commonly leveled against an analysis of 'I' that presupposes the primacy of the linguistic. Third and finally, building upon insights of David Kaplan, I present an interpretation of meaning-rule under which 'I' operates in its first person use.

It is customary in the philosophy of language to contrast semantic theories that provide a truth-functional analysis of meaning with those that provide a communicational analysis. In The First Person: A Study in Intentionality and Reference (Chisholm, 1983) Roderick Chisholm cuts up this territory differently. He advises us to distinguish between theories that "presuppose the primacy of the intentional" and those that "presuppose the primacy of the linguistic". These are broad categories, but as I shall use the terms in this paper, a theory which presupposes the primacy of the intentional (hereinafter IT or ITR) takes thought or some other intentional entity to be conceptually prior to language in the mechanism of reference. Conversely, a theory that presupposes the primacy of the linguistic (hereinafter LT or LTR) takes language to be logically prior to thought. In contrast to the traditional classification, Chisholm's revised scheme groups Frege and, for example, a philosopher such as Grice into the same camp, for each takes an intentional entity (for Frege, a "sense"; for Grice, a "speaker's intention") to operate as the primary vehicle of reference.

Chisholm's program is to use first person singular reference—what we may call 'I-reference'—as a test against which both ITR and LTR may be evaluated. His thesis contains both a metaphysical and semantical component. The metaphysical component, largely a holdover from the views developed in Person and Object (Chisholm, 1976), consists of the theory that the structure of reality is built upon the irreducibly sub-
jective points of view of unique consciousnesses or selves. The semantic component consists of the view that since the primary purpose of 'I' is to express this subjective point of view, only a theory which presupposes the primacy of the intentional can do justice to the peculiar semantics of that term. The *First Person* attempts to work out a semantics of first person singular reference that adheres strictly to IT assumptions.

This paper shares with Chisholm the assumption that the correct treatment of I-reference presents a challenge to both ITR and LTR. Its thesis, nevertheless, is more modest than Chisholm's, and it runs counter to his approach. My argument is that LTR offers far greater promise than does ITR in accounting for what is semantically interesting about first person singular reference.

To make out this thesis I want to do three things. First, I want to demonstrate that intentionalist analyses of 'I'—specifically the accounts of Frege and Chisholm—are conceptually flawed. Frege's classical treatment of 'I' cannot accommodate some of our most basic intuitions about that pronoun, and Chisholm's recent variation on Frege is simply wrong about reference in general. Second, I want to show how one recent non-intentionalist approach, namely David Kaplan's Direct Reference Theory, can answer the general objection that intentionalists raise against LTR. Third and finally, I want to offer a refined treatment of first person singular reference that, building on Kaplan's insights, adheres strictly to LT assumptions.

It will be seen that while my thesis is formulated with the help of certain categories borrowed from Chisholm, this thesis should not be interpreted as confronting Chisholm on every front. While I do not accept Chisholm's metaphysical theory that the structure of reality is built up on the subjective point of view of unique selves, nothing in this paper speaks directly to the issues raised by this theory. Furthermore, the primary role of 'I' in language is not in my opinion to express the subjective point of view—a problematic notion if there ever was one. I do not, however, directly criticize the arguments for this claim that Chisholm produces in Chapter 1 of *Person and Object*.

Let us begin with Frege's classical treatment of I-reference. In Fregean semantics designators determine objects indirectly. A speaker will associate a certain *sinn* or sense with a given designator, and the individual uniquely falling under the concept associated by that speaker with that designator is said to be the denotation of the term. Hence, the primary vehicle of reference for Frege is always a nonlinguistic entity. There was an unresolved tension in Frege's thought over the ontological status of senses. Wanting them to be intersubjectively accessible, he declined to characterize them as "mental". But he never clarified how senses could be both conceptual and nonmental. Since Frege, most theorist have followed Church in giving a Platonic interpretation to senses. On that realist interpretation senses become "individual concepts".

This tension shows up in Frege's application of his logic to 'I'. He claims that 'I' has both a public sense and one that is "presented to no one else". (Strawson, 1967, 20) In its public and communicable guise, the
individual concept associated with the pronoun, according to Frege, is "He who is speaking to you at this moment". If I say, "I own that cat", I indirectly single out a unique individual as owner of the cat because that individual, namely me, falls under the concept expressed by the description 'He who is speaking to you now'. 'He who' can be interpreted as potentially expressing a definite description encapsulating my self-concept.

This is not, however, a successful analysis. The perspective provided by contemporary possible worlds semantics, though not of course available to Frege, shows how the analysis leads to unwanted consequences. Imagine a Putnam-like Twin Earth example. (Putnam, 1975) Jones-1 is an inhabitant of Earth, and Jones-2 is Jones-1's doppelganger on Twin Earth. By hypothesis, Jones-1 and Jones-2 mirror each other exactly—both physically and psychologically. They associate identical individual concepts with the pronoun 'I'.

If 'I' accomplishes reference by means of the mediating sense associated with it, that is, if the denotation of 'I' is determined by the self-concept under which it falls, then 'I' expressed from the lips of either Jones-1 or Jones-2 succeeds in referring to both of them. On the Fregean analysis, therefore, 'I' is sometimes subject to referential ambiguity. Of course a determined Fregean, in light of this, may be tempted to throw out possible world semantics or he may be tempted to accept the result. Neither of these alternatives is very attractive, however. On standard methodological grounds, a theory which avoids these alternatives, assuming that it is not itself subject to decisive objections, is surely preferable to one that requires such a major reevaluation of our other beliefs. We will, however, return to the issue of referential ambiguity in the final section of this paper.

A distinct yet similar difficulty emerges when the case is changed slightly. Imagine that Jones-1 and Jones-2 are subject to self-deception and that the self-concept held in common by both of them fits neither of them. By chance of fate it is Smith who possesses the self-concept that both Jones-1 and 2 associate with their respective uses of 'I'. In this case 'I' expressed from the lips of either Jones-1 or Jones-2 refers to neither one or other of them, nor to both of them. It refers to Smith. On the Fregean analysis, therefore, the reference of 'I' is sometimes open to misfiring; that is, on occasion it may inadvertently refer to someone other than the one to whom its user intended to refer. This, too, would seem to be a counterintuitive result and one that we should avoid if possible. Again, we return to the issue of referential misfiring in the final section of this paper.

We have tested the public sense of 'I' against the recherché cases made available by a possible worlds perspective. Our results, though, will not deter a contemporary defender of Fregean semantics who believes that reference is only secondarily a matter involving communication. Does the private sense of 'I' fare any better? In the passage of "The Thought" where Frege introduces the communicable sense of 'I' he asserts, as mentioned earlier, that it also expresses an incommunicable sense, one anchored in the way that "everyone is presented to himself in a particular and primitive way, in which he is presented to no one else."
In *The Subjective View* McGinn argues that this claim is ambiguous between two readings. (McGinn, 1983, 58) Under the "AE" interpretation wide scope is assigned to "everyone". Thus:

$$\forall x \exists m [ x \text{ is presented with } x \text{ in } m \land \neg \exists y (y \text{ is presented with } x \text{ in } m \land y=x)].$$

Under the "EA" interpretation wide scope is assigned to "a particular and primitive way". Hence:

$$\exists m \forall x [ x \text{ is presented with } x \text{ in } m \land \neg \exists y (y \text{ is presented in } m \land y=x)].$$

The AE reading allows that different persons are presented to themselves by means of distinct particular and primitive modes of presentation. The EA reading claims that there is but a single manner of self-presentation common to all agents—they all represent themselves in the same way when they use 'I'. McGinn does not find anything in Frege to settle decisively the exegetical dispute arising from this ambiguity, but he is persuaded that the EA thesis is semantically correct.

I am persuaded otherwise; but more to the point I am persuaded that McGinn is mistaken on the exegetical issue. McGinn is fully aware that EA violates the principle most commonly associated with Frege, namely, that sense determines reference. If everyone conceives of him or herself under the same cognitive representation when he or she uses 'I', then that concept will not be fine-grained enough to individualize the separate thinkers of it. Something extra-conceptual—perhaps context—will have to be added to the mode of presentation to guarantee individuation of referents. McGinn suggests that Frege might have been arguing his way to the realization that indexicals in general cannot be adequately handled by the Logic of Sense and Denotation.

McGinn's reasoning, it seems to me, goes against the grain of the textual evidence. The purpose of "The Thought" is precisely to show how the Logic of Sense and Denotation can be extended to include indexical expressions. Frege's claim is that a "complete thought" will often contain more than its "verbal expression". In the case of a present-tense statement, for instance, "one must know when the sentence was uttered to apprehend the thought correctly. Therefore the time of utterance is part of the expression of the thought." (Strawson, 1967, 24) This is by no means clear on the surface, so why did Frege maintain it? The answer is that he did not wish to abandon his foundational principle that sense (in this case "thought") determines reference. If the index of time is not included as part of the expression of the thought, then its reference (for Frege, its truth-value) cannot be determined.

It is significant that Frege mentions 'I' in the textual context where the above point is being pressed. He says, "the same utterance containing the word 'I' will express different thoughts in the mouths of different men. . . ." (Strawson, 1967, 24) It would appear, therefore, that each individual has to present himself in a "particular" mode when he presents himself to himself from the inside by means of 'I'. Hence, AE would seem to more accurately represent Frege's thinking.

The point I wish to make, however, is that Frege's account is just as problematic if it is construed as an AE thesis as it is if it is con-
strued as an EA thesis. Regardless of whether each I-presentation is unique or not, the crucial characteristic of Fregean self-presentation is that they are private.

The notion of privacy may be given a strong and weak interpretation. An experience is strongly private if it is characterized by both (a) uniqueness of possession (i.e., "only I can have just this experience") and (b) inexpressibility. An experience is weakly private if it is characterized by this first property but not the second. Someone might claim that having a pain is an example of a weakly private experience. If I were to mention my headache, for instance, you can understand what I mean, but you cannot have this very experience of headache.

If the AE thesis is correct, then my subjective I-presentation would presumably be private in the strong sense—it would be "primitive" in the sense that it is strictly incommunicable. On the other hand, if the EA thesis is correct, then my subjective I-presentation is only weakly private. While others cannot have just this presentation, they may, nevertheless, be in a position to understand the sort of thing determined by this presentation. In advancing the EA thesis McGinn is in the awkward position of having to communicate to us the sort of thing we are to grasp when we grasp the thing determined by a private I-presentation. But he becomes vague just when details are expected. In respect of indexical expressions in general he tells us that "... to think of something indexically is to think of it in relation to me, as I am presented to myself in self-consciousness". (McGinn, 1983, 17) But this just begs the question against the skeptic who would hold that there is no sense in which I am presented to myself in consciousness.

Frege is in an even worse position if, as I have argued, he holds the AE thesis. On that view the experience that I have when I present myself from the inside by means of an I-thought is different from the one you have in the same circumstance. It is fundamentally incommunicable. This makes it particularly difficult for Fregean semantics. For if my knowledge of my self turns out to be inexpressible, this will not aid me in explaining the idea of self-presentation to someone who fails to find it obtaining in his own case. In the absence of this idea it is difficult to imagine providing content to the idea of a private I-sense.

Since, then, it is the notion of privacy that gets us in trouble here, could the essentials of the Fregean account be preserved in some other way? Some philosophers have thought the notion of a "haecceity" could prove helpful here. (Chisholm, 1976) Of any x it can be said that x is identical with x. If the ability to express an open sentence such as "_ is identical with itself" is sufficient to provide a genuine property—and not, as Shoemaker would say (following Geach), a "mere Cambridge" property (Shoemaker, 1980)—then x's "being identical with itself" is a property essential to x and essentially unique to x; i.e., it is a property which x has and x has in every possible world in which x exists.

Now if 'I' were to express my haecceity, then I and only I can fall under the individual concept so interpreted. Thus, while retaining the primitive and particular character of Frege's incommunicable sense of 'I', this way of unpacking that sense would eliminate the unfortunate privacy affecting the plausibility of Frege's account.
The difficulty with this interpretation is this. Whatever manner of presentation Frege thought he had in view when he spoke of a particular and primitive sense of 'I', it would not seem to be captured by the thin and strictly formal notion of one's haecceity. A haecceity does not constitute the kind of entity seemingly required by IRT. The notion of intentionality as traditionally understood involves the directedness of thought upon an object. My haecceity, understood simply as my formal property of being identical with myself, does not give me a concept, a vehicle, by means of which to direct my thought upon myself. Frege's incommunicable senses, if they were to exist, would have the capacity to be grasped by the mind and to be directive of thought. By contrast, as Chisholm correctly remarks, "The property, if there is one, which is intended by the expression 'being identical with me' would seem to be extraordinarily empty." (Chisholm, 1981) A theory of first person reference which avoids securing reference by means of haecceities would, on these grounds alone, be preferable to one that utilizes this strictly formal device.

II

Chisholm's recent work represents an attempt to surmount the difficulties of the Fregean analysis while remaining faithful to intentionalist intuitions. For Frege the pronoun 'I', along with its indexical brothers, lies at the periphery of investigation. It constitutes an exception, to be incorporated, if possible, into a more general treatment of singular terms. But for Chisholm I-reference lies at the center of a metaphysical-semantic structure. We use 'I', according to him, to express a subjective point of view, and, as mentioned earlier, this point of view allegedly represents something metaphysically basic. Hence, in The First Person Chisholm does not just reject the Fregean concept of sinn, he adopts an entirely different way of looking at these matters. Fundamental to this orientation is a distinction between what he calls "direct attribution" and what he calls "indirect attribution".

The intentionalist act of directly attributing a property to oneself is a primitive act within Chisholm's system. Two minimal principles may be affirmed about the operation of this act: (P1) one can only directly attribute to oneself, and (P2) what one directly attributes is some property F. Thus, when I think to my self (or publicly utter) "I am fortunate", I thereby, first, take myself as an "intentional object" and, second, in so doing "grasp or conceive a certain property" which I attribute to myself. (Chisholm, 1981, 22) The precise sense in which I "take" myself as an intentional object is indefinable.

The importance of this idea is that indirect attribution or attribution of properties to objects and persons other than oneself is conceptually parasitic upon direct attribution. Thought and talk about others always depends upon more basic thought about oneself.

Suppose it is you to whom I ascribe fortune. According to Chisholm, I make you my reference (or "intentional object") by attributing a certain property to myself. That property, which is termed the 'content' of direct attribution, is one that "singles you out and thus makes you the object of an indirect attribution. (Chisholm, 1981, 29-30) The property is two-fold: it consists of an "identifying relation", a certain relation R such that you are the only thing to which I bear R, and, secondly, an
implication to the effect that the only thing to which I bear R has property F. In the present case the property I indirectly ascribe to you is that of being fortunate; the identifying relation R that I bear only to you is, for example, that of talking to you (or living across the street from you, or reading a book that you wrote, etc.): and the property that I directly attribute to myself is that of talking with exactly one person and with a person who is fortunate. Indirect attribution is more formally introduced by means of the definition:

\[ D_2: y \text{ is such that, as the thing that } x \text{ bears } R \text{ to, } x \text{ indirectly attributes to it the property of being } F = \omega x \text{ bears } R \text{ to } y \text{ and only to } y; \text{ and } x \text{ directly attributes to } x \text{ a property which entails the property of bearing } R \text{ to just one thing and to a thing that is } F. \] (Chisholm, 1981, 31)

Chisholm takes the standard intentionalist line that the first person personal pronouns are merely the normal, but by no means essential, devices for expressing the act of direct attribution. This act, as noted, involves a primitive relation between thought and object and, in keeping with the intentionalist motivation, can presumably take place without the use of language at all. Indirect attribution, which is expressed in the normal case not only by second and third-person indexicals but also by proper names and definite descriptions, is accomplished by means of the more basic act of direct attribution.

Someone might object to Chisholm's analysis on the grounds that what is desired from it, an explanation of how I-thoughts get directed to particular I's, is just what is made impossible by the claim that direct attribution is indefinable. In some sense I think this is the most serious problem with this account. But Chisholm himself is not impressed with this objection (arguing that all theories have their primitives), so I wish to focus here upon a more particular difficulty with the account. This is the problematic status of the first conjunct of D2.

My ability to indirectly attribute F to you depends upon the existence of some identifying relation R that I bear to you and only to you. This first conjunct:

Chis 1: \( x \text{ bears } R \text{ to } y \text{ and only to } y \ldots \)

must be satisfied if reference to you is to be guaranteed. What happens, however, if this condition fails to obtain on some occasion? Chisholm asserts that such a contingency need never arise:

\[ \ldots \text{ whenever a person indirectly attributes a property to a thing, he can specify a certain identifying relation } R \text{ which is such that the thing in question is the thing to which he bears } R. \text{ This is not difficult to do. Consider } \ldots \text{ what identifying relation is such that W.V. Quine is now the thing to which I bear that relation? It could be that of looking at my copy of a certain book that he wrote. } \ldots \text{ (emphasis mine)} \] (Chisholm, 1981, 31)

The modal term 'can' presents the problem. From "possibly p" we are not justified in inferring 'p'. It may be possible for me, in attributing a property to a, to specify a certain R which singles a out; but it plainly does not follow that R has anything to do with the actual
mechanism by means of which a was referred to. For instance, it is possible to construct a relation such that W.V. Quine is now the person to which I bear that relation; it is that of looking at my copy of a certain book he wrote, say, *Word and Object*. But if one is to argue that it is impossible to refer to Quine and to attribute a property to him without conceiving (at least tacitly) some such relation, then a principle is needed to supplement Chis 1.

This principle (left unformulated by Chisholm) is:

Chis 2: there is an identifying relation R that the speaker can specify and which is, in fact, the relation by means of which that speaker did single out the object of his indirect attribution.

The explicit formulation of Chis 2 makes it clear that when I say "y is F", I actually have not referred to y, as the object of my attribution, unless the identifying relation R—which I not only can conceive but actually do conceive in referring to y—is such that it singles out y and only y. The objection to this condition is that it is too strong. Cases are imaginable in which one successfully manages to attribute a property to another even though the identifying relation R at issue is not uniquely identifying.

Consider the following case. I believe that the identifying relation R that I bear to y is uniquely identifying, but it turns out not to be. I say, "Quine is a famous philosopher", believing that the relation R that I bear to Quine is that of now looking at a copy of a certain book, *The Web of Belief*, of which I think he is the sole author. This belief is false; in reality I bear R not only to W.V. Quine but also to Joseph Ullian. Does it follow that, in my utterance, I have failed to attribute a property to Quine? On Chisholm's analysis it does; I have not succeeded in referring to Quine because the relation by means of which I attempt to single Quine out is not a relation I bear to Quine and only to Quine; that is to say, Chis 1 fails of satisfaction. But this seems to me to be counterintuitive. I would seem to have referred to Quine if the proper name I use bears the correct connection to the man at issue. The misimpression contained in my belief is entirely irrelevant to the success of my reference or to my capacity to ascribe a certain property to the thing to which I have referred.

One is reminded of Donnellan-like cases. (Donnellan, 1966) Jones says, "The man in the closet reading Shakespeare is behaving oddly", believing of Smith that he is behaving oddly. Jones is mistaken. It is Baker rather than Smith who is in the closet. Now, what is to be concluded if the ostensibly identifying relation R, conceived of by Jones in order to single out Smith, is: that of thinking about the person in the closet named Smith? Since Chisholm's analysis places no restriction on the formulation of identifying relations, there is nothing to rule out this conception. Nevertheless, there is no one to whom Baker bears the relation R; his attempt at indirect attribution has presumably failed. Hence, on Chisholm's account Jones has not referred to anyone, neither Smith nor Baker.

Again, this seems counterintuitive. One might say, making use of Donnellan's famous distinction, that what Jones expressed attributively is that [Baker is behaving oddly] whereas what he expressed referentially
is that [Smith is behaving oddly]. Or one might say, appropriating the Gricean distinction between word meaning and speaker's meaning, that what Jones "expressed" propositionally is that [Baker is in the closet behaving oddly] whereas what he meant to "convey" is that [Smith is in the closet behaving oddly]. (Grice, 1957) Neither of these ways of putting the matter, however, sanction Chisholm's conclusion of reference failure.

Chisholm's account requires some such principle as Chis 2. But it is precisely this principle that proves counterintuitive. The role identifying relations have to play with regard to attribution is not that of picking out the object of attribution. It is plausible to assert that we can always conceive of some relation R that uniquely individuates the object of attribution. It is a mistake to assert that it is necessarily R, or some similarly conjured up relation, that is used to accomplish reference to that object.

The conclusion we may draw from this objection is that either reference is accomplished in some nonintentionalist manner or that there exist some identifying relations by means of which reference is accomplished that are not uniquely identifying. The first disjunct entails concurring the failure of any account presupposing the primacy of the intentional, and the second disjunct implies that Chisholm's specific analysis has little to recommend it.

I find little appealing in Chisholm's basic intuition; that is, the idea that thought and talk about others is to be explained finally in terms of thought and talk about oneself. My specific objection, however, to his theory is that he fails to deliver on his promise to show how this explication can be executed without distorting our intuitions about reference.

III

It is time to examine the alternative to ITR. As I mentioned before, the most fully developed example of a nonintentionalist theory (or at any rate one that can be given a nonintentionalist interpretation) is Kaplan's Direct Reference Theory. (Hereinafter, we substitute DRT for our earlier initials, LTR.)

The fundamental idea of DRT in its application to indexical expressions is that the mechanism of singular reference can be completely explained without making use of intentional entities. Suppose I say, "That doesn't belong here". On the version of this theory proposed by Kaplan (Kaplan, 1977), the indexical terms 'that' and 'here' in this sentence possess both a "character" and a "content". The character of an indexical is the implicitly understood linguistic rule(s) governing the use of the word. The content is the referent of the term. Thus, understanding the meaning-rule or character under which 'that' and 'here' operate, I express the above sentence in a given context, and in that context certain references (say, the ordered pair [dirty shirt, bathroom floor]) are automatically assigned to the indexical expressions at issue. Hence,

KT 1. Reference is function from particular, actual or possible contexts to individuals.
Note that DRT makes use of the notion of "understanding" a meaning-rule. One may be inclined to see understanding as an irreducibly intentionalist notion and thus to see DRT as ultimately a version of IRT. This is a mistake. DRT rejects the assumption that some intentional entity--be it a "sense," an "intention" or a "direct attribution"--must be brought into the analysis of the mechanism of reference. Here we follow McGinn (1981) in distinguishing between a semantic analysis of the mechanism of reference and a psychological analysis of what is involved in understanding a semantic rule. A meaning-rule will not of course be effectual unless it is known. But from this it does not follow that a knowledge state is constitutive of that rule. To borrow an analogy from McGinn: a necessary condition of something being a combustion engine is that it is made of metal. "But grasping the mechanism by which a combustion engine works does not involve this fact." (1981, 161)

In any case, the sentence in question is Russellian in nature. Hence,

KT 2. Propositions consist, inter alia, of the contents of the indexicals used in the sentences that express them.

If DRT is to serve as a challenge to intentionalist approaches to reference, however, these two theses are not sufficient. We need an additional claim. Consider the case of the ignorant heiress. Kidnapped, locked in the trunk of a car, and not knowing where she is or what time it is, she nevertheless utters, "Well, here I am now". What singular proposition is expressed by this sentence? According to Kaplan, the actual place and time of the heiress' entombment function as the contents of the expressed proposition even though the heiress herself is ignorant of these referents. Hence,

KT 3. Ignorance of the referents does not defeat the referential character of indexical expressions.

For Kaplan propositions must be detachable from the actual or possible contexts that produced them. It is only in this way that they may be evaluated from the perspective of different circumstances (i.e., possible worlds). It is this requirement that justifies KT 3.

KT 3 has the anti-intentionalist effect of driving a wedge between the semantic proposition and what's going on in the heiress' head. It is for precisely this reason that defenders of intentionalist semantics will find the account unsatisfactory. Consider, for example, H-N Castañeda's response. He points out that on the standard view propositions are the bearers of truth-value. But while Kaplan's Russellian proposition may be what is truth-valued from a purely external point of view, it omits the "truth" that is experienced by the heiress. Castañeda comments:

The truth she [the heiress] proclaims seems indeed to lie right there in the middle of her experience; it is the truth that there is quiet at the time of her experience as she experiences it and at the place where she experiences herself to be. She seems to be asserting a purely experiential and experienced truth, which must, therefore, be thoroughly differentiated from the Russellian-Kaplanian proposition. . . .
(Castañeda, 1983)
Castañeda believes that we must admit of an "internal accusative of thought" to function as the truth that the heiress experiences. It is this internal accusative, omitted in DRT, that constitutes the basic structure of reference for intentionalist theory. (Castañeda omits mention in this passage of any alleged internal representation for the 'I' in 'Here I am now'. For purposes of simplicity, I will follow him in this.)

The correct reply, I believe, to Castañeda's problem can best be expressed as a dilemma. Either the internal accusative to which he refers is truth-valued or it is not. If this structure is truth-valued, then it turns out to be irrelevant to semantics. If it is not truth-valued, then it is irrelevant to semantics for another reason. By either horn, DRT is warranted in neglecting the internal accusative, if there is such a thing.

Suppose we assume that the accusative is a bearer of truth-value. In what does its structure consist? If the heiress were not suffering from a condition of complete ignorance, we might picture this internal accusative as a structured entity containing certain mental representations of what, inter alia, 'here' and 'now' single out. The referent of 'here' would be a mental representation of the actual place of the utterance, and the referent of 'now' would be a representation of the actual time of the utterance.

This picture is no help to Castañeda, however. These representations, were they to exist, would simply be modeled on—that is to say, their analysis would be dependent on—the primary referents contained in the K-R proposition. The referent of 'here' would presumably be a phenomenological representation of the car trunk and its spatial context, and the referent of 'now' would be whatever phenomenological representation the heiress could have of the actual time. The point, though, is that the internal accusative of thought would merely be a shadowy, mentalese imitation of the external proposition. There is no reason for a specifically semantic analysis to elevate this structure to relevance. Thought would be "directed" in virtue of its function as internal speech.

But of course the heiress is in no position to display these representations to her consciousness. What, then, could function as the content of the mentalese counterparts to 'here' and 'now'? What Castañeda says is "that there is quiet at the time of her experience as she experiences it and at the place where she experiences herself to be". This is by no means clear, but two interpretations suggest themselves.

According to the first interpretation we simply abandon the claim that the internal accusative or 'truth' is singular in structure. Perhaps the general proposition, "There is an x and there is a y such that x is an experience of place for S and y is an experience of time for S and x and y have property P", more accurately represents the correct construction of the internal accusative. If this is the case, however, the burden of proof would seem to be entirely on Castañeda. For even if we assume this construction represents the general features of the phenomenological experience the heiress is undergoing, we are still left with the task of explaining what that construction has to do with the terms 'here' and 'now' as employed in the sentence, "Well, here I am now". Those terms do not appear to be operating semantically in the way this interpretation requires, i.e., as propositional functions. Instead they operate as singular designators. In this case and in the more general case
the proper semantic treatment of sentences containing indexicals would seem to demand existential commitment to singular propositions.

The second interpretation involves construing the internal accusative as a singular truth but taking 'here' and 'now' to be operating as definite descriptions for something on the order of 'the quiet and dark place . . . ' and 'the eerie time . . . ', respectively. On no construction, though, does this solution work.

If a Russellian interpretation is placed upon these definite descriptions, singular propositions are surreptitiously abandoned in favor of general ones, and we are back to the preceding solution. If the definite descriptions are interpreted along Donnellan's lines, then they are in fact not definite enough. They can accomplish neither referential nor attributive reference. 'Here' and 'now' cannot latch on to something referentially because there is no antecedent something in the mind of the heiress to which these terms are to aid us (or her) in singling out. Furthermore, 'here' and 'now' fail to refer attributively—i.e., to refer to whatever happens to satisfy their disguised description—because there is no reason to assume that these descriptions or their mental representations are sufficiently definite to admit of unique satisfaction. Finally, if these descriptions are interpreted as elliptical for 'the quiet and dark place where I am' and 'the eerie time that it is,' unique satisfaction is ensured only at the expense of affirming Kaplan's general point. For if the expressions 'the place where I am' and 'the time that it is' succeed in referring to the actual place and time of the heiress' captivity, they do so even though she is entirely ignorant of them; even though, as it might be put, she is not de re en rapport with them.

We are thus led to the second disjunct of our original dilemma. I assume most proponents of DRT would find this the most plausible alternative. Whatever else is said of the internal accusative, it would seem not to be truth-valued. If we could project the stream of consciousness that characterizes the heiress' mind as she contemplates her problem and utters "Well, here I am now", there is no reason to suppose this stream possesses propositional structure. The heiress is, as Castañeda says, "experiencing" quiet (and darkness) at the time and place of her cramped captivity. But this clearly does not warrant the thesis that "she seems to be asserting a purely experiential and experienced truth". What she is asserting is the K-R proposition; what she is experiencing is no truth at all.

Therefore, the internal accusative insofar as it is an object of experience rather than a bearer of truth, would seem to be irrelevant to the project of semantic analysis. Alternatively, one might say that if the heiress' expression marks a "truth" which functions as an experiential accompaniment to the semantic proposition, this sense of 'truth' has no referential cash value. On this horn, in other words, the idea that thought has "direction" independent of language is simply not accepted.

IV

I do not contemplate that this answer will serve as a comprehensive response to every objection that might be leveled against the very idea of DRT; the relation between thought and language is after all the fundamental issue in post-Wittgensteinian philosophy of language. I
would, however, claim that this argument should block the checkmate that intentionalist prejudices sometimes put in the way of even proceeding along Direct Reference lines. If we can liberate ourselves from the picture of 'I' as essentially representing or expressing something internal, we are in a position to see how 'I' actually operates externally, that is, in natural language.

However, the puzzle we are left with is this. I have a number of devices at my disposal for referring to myself. Consider just two of these—'I' and my proper name 'KA'. If what we said in the preceding section is correct, then I-propositions share in part an identical content with singular propositions that get produced by means of my proper name. A certain person, me, is the subject constituent of any token of this type. It might seem to follow, therefore, that the first person pronoun 'I' and proper name 'KA' demand the same semantic treatment. So how can DRT account for I-reference without assimilating it to a proper name?

On the standard DR view of proper names 'KA' refers to the person who bears the correct causal or historical connection to the person who was baptized 'KA' in a certain time, at a certain place, in the actual world. In their recent book *Knowing Who* (Boer and Lycan, 1986) Boer and Lycan mistakenly assume that DRT will treat 'I' as a proper name too. They write, "Note also that for some singular terms, such as indexical pronouns like 'you' and 'I', the 'appropriately shaped causal chains' are so short and direct as to be degenerate cases." (Boer and Lycan, 128) But with 'I' there are no causal chains, appropriately shaped or otherwise. To think of a use of 'I' as a degenerate case of tagging something with a name is to endorse an entirely misleading picture of both the Causal Theory of Reference (one variation on DRT) and first person reference by means of 'I' and 'me'.

This is not to say that 'I' (or even 'me') cannot be used as a proper name. It can even be used as a definite description. Compare two cases. In the first case you wear a moustache. One day while shaving you notice splotchy bald spots, alopecia areata, in your moustache. You utter, "I am losing my moustache". Contrast this with a second case. You are an extremely famous scientist, well known for your public stands on controversial social problems. You are watching a play, and the play is about you. Midway through the play you are amused to notice that the false moustache on the actor portraying you is beginning to slip. You turn to your companion and utter, "Look. I am losing my moustache".

The same sentence is uttered in both of these cases and that sentence contains a use of 'I' in the subject position. However, the first example discloses a "first person" use of 'I' whereas the latter example does not. Why is this? The answer is that the indirect discourse proxy for the first person use of 'I' would necessarily make use of the indirect reflexive locution 'he himself'. By contrast, the use of 'I' that does not constitute a first person use would not nor could not utilize this indirect reflexive.

Consider. In the first case it is appropriate to say of you, "He knows that he himself is losing his moustache". The reflexive pronoun "himself" is essential to this transformation because it grammatically represents the fact that in the first case your utterance of 'I' is directed upon yourself qua performer of the utterance. In the second case
the indirect use of the reflexive is entirely inappropriate. In this case we would not be tempted to say "He knows that he himself is losing his moustache". We know from the description of the case that you are in no danger of mistaking the actor for yourself. Hence we would use a nonreflexive, singular term for the indirect discourse transform, (e.g.) "He knows that the actor portraying him is losing his moustache".

This suggests the following criterion. An employment of 'I' in a sentence of direct discourse is relevant to the semantic analysis of first person singular reference if and only if the transformation of that sentence into indirect discourse necessarily involves (after the relevant that-clause) the use of a grammatically indirect reflexive expression. This criterion filters out those conceivable uses of 'I' that have little to do with the study of first person semantics. If a parent were to name her child 'I', we would feel sorry for the child but would not confuse this idiosyncratic employment with the use of the pronoun 'I' in natural language. I think we should assume a similar attitude toward the use of 'I' as a definite description in the second case above. Clearly it is a possible use, but this does not imply that it will prove relevant to the analysis of first person, singular reference.

To return, then, to our puzzle: we want to put me in the K-R proposition whether I say first personally, "I intend to write this paper," or whether I say "KA intends to write this paper," (or even "This man intends to write this paper,"). But importantly we want the referential mechanism by means of which this is accomplished to differ in the two cases. The difference is to be sought in the fact that these two kinds of expression exhibit divergent characters.

We found the 'he himself' locution to be essential to the formulation of a criterion for first person uses of 'I'. Grammatically, as noted, the term 'himself' is a reflexive pronoun. It is the concept of reflexivity, I believe, only now applied to semantics, that can help us to see how a first person use of 'I' differs in character from other devices of singular reference.

The idea of reflexivity analytically entails the notion of "turning back upon". Consider another reflexive expression, 'this very phrase'. Semantically we might, with Nozick, describe this expression as a case of reflexive "self-reference". (Nozick, 1986) What the phrase refers to of course is itself; on each occasion that 'this very phrase' is tokened or replicated it refers to the phrase so tokened or replicated. It is important to note that to say a term self-refers in this sense is just to mean, neutrally, that it refers to itself. In this example the reflexivity is provided by the introduction of the term 'very' into the phrase 'this phrase'. Although one might use 'this phrase' to single out itself, one can use it to refer to the same phrase on different occasions of use. The addition of 'very' makes the phrase sensitive to its own tokening.

It would be a mistake to suppose that 'I' reflexively self-refers in precisely this way. Plainly 'I' does not refer to itself. Instead, what 'I' refers to on each occasion of its replication or tokening is the producer of that token. I use the term 'producer' here advisedly. Since the extension of the concept "producer" is broader than the extension of the concept "self", the use of that concept allows us to remain noncommittal about whether 'I' refers to a "self," in the sense of a nonempirical agent which may somehow take itself as an object of its own reflection.
In this way we avoid confusing the semantical question "How does 'I' refer?" with the metaphysical question "To what does it refer?". Of course it is this former question that interests us, not the latter.

Initially one might believe that the explanation of the reflexivity characterizing I-reference could be piggybacked upon the explanation of the reflexivity of 'very', as in 'this very phrase'. On a view once held by Reichenbach 'I' just means "the person who utters this token"; and it is clear that he intended to say 'the person who utters this very token'. A refined version of this, incorporating the point made above, might be: 'I' just means 'the producer of this very token'. The objection to this meaning-postulate is that it would make inexplicable the difference between the analytic and necessary statement:

The producer of this very token is the producer of this very token,

and the synthetic and contingent statement:

I am the producer of this very token.

(We must assume here, of course, that the definite description 'the producer of this very token' does not fail to denote.)

The advent of DRT shows why meaning-postulates of the type suggested by Reichenbach cannot function in a semantic analysis in the way that was once supposed. The expression 'the producer of this very token' is not a synonym of 'I'. By KT 1, this expression acts as a character which, together with a context, fixes the reference of 'I'.

The idea of using a character to fix a reference rather than to supply a synonym is borrowed by Kaplan from Kripke. (Kripke, 1980) What sets Kaplan's application of this idea apart from Kripke's, however, is that Kripke is concerned with the use of an accidental property of a thing (e.g. being the length of the standard meter stick in Paris) to fix the reference of a singular term naming that thing (e.g., 'one meter'). While rules of the type envisioned by Kaplan have nothing to do with accidental properties, if the idea can be extended in the way proposed, Kaplan argues that indexical expressions fall into either one of two categories.

These categories are defined by the manner in which the expressions subsumed get their respective references fixed. Demonstratives, terms such as 'this', 'that', 'he' or 'she' require two elements to accomplish reference. They need, first, some appropriate character--e.g., "the reference of 'she' must be of the female gender". They need, second, an appropriate demonstration of the exact thing which meets the condition(s) specified by the character--e.g., "She (pointing to the person with the book in her hand) is a fine lecturer". Pure indexicals, on the other hand, terms such as 'today', 'tomorrow', 'here', and 'now' require only an appropriate character--e.g., "'today' refers to the day in progress". In their case no associated demonstration is required to fix a reference.10

Where does 'I' fit? According to Kaplan, it is not a demonstrative; for, aside from emphasis, any demonstration of oneself accompanying an utterance of 'I' is redundant. It therefore is to be classified with the
pure indexicals. On his view the rule in the sense of character under which 'I' operates is formulated by what Perry and Castañeda have called the K-rule: In each of its utterances, 'I' refers to the speaker who utters it. (Tomberlin, 1983)

The advantage offered by Kaplan's account is immediately clear. It provides us with a mechanism for distinguishing I-reference from other types of singular reference. As just noted, 'I' is not a demonstrative. Further, the manner in which the reference of a proper name is fixed—even one's own—is causal rather than indexical in character. Finally, definite descriptions, at least in their basic use, accomplish reference by means of satisfaction; in the expected situation one and only one person satisfies the description 'the conqueror of the North Pole'. Hence, on Kaplan's view we can register the semantic differences among reference by means of 'I', demonstratives, proper names, and definite descriptions.

The difficulty with Kaplan's account is that the distinction between demonstratives and pure indexicals is more problematic than Kaplan himself supposes. Two types of demonstration are sanctioned by Kaplan—explicit and opportune. (Kaplan, 1977, 100) An opportune demonstration takes advantage of the opportunity afforded by the circumstances. Thus if a fly were buzzing about my head, the 'that' in "That is annoying" would be taken by most of us to demonstrate the fly, even though I do not point at the fly directly. But the acknowledgment of opportune demonstration—and it is difficult to see how that category could be denied—makes pure indexicals look very much like demonstratives. Consider 'today'. There is a way of seeing the reference of 'today' as always being fixed opportune. The reference of 'today', we might say, is fixed by the fact that a particular day happens to be in progress on the occasion of my utterance. Similarly with, for example, 'now': the reference of that term is fixed by the fact that my utterance happens to occur during a particular temporal passage. In other words, one might argue that the referents of pure indexicals are always demonstrated opportune. But if the notion of pure indexicality does not subsume a genuine category, we do not make conceptual progress by describing 'I' as a pure indexical.

Furthermore, even if we waive this difficulty, Kaplan's analysis proves anemic. To say that 'I' gets its reference fixed by a meaning-rule alone is not to say very much. If what we want is an explanation of how 'I' accomplishes reflexive reference, that is, an explanation of how 'I' turns back and singles out the producer of its token, we will want to supplement Kaplan's classification with an interpretation that shows what is peculiar about I-reference. Here, I think it is crucial to recognize that there is a "demonstrative-like" aspect to reference via 'I'. For, in contrast to the other pure indexicals (see, endnote 10), a demonstration of oneself may accompany a use of 'I'. In fact, a demonstration of oneself may take the place of a use of 'I'.

To explain this demonstrative-like feature I want to introduce the idea of a surrogate demonstration. Think of a case where you raise your hand in response to a "Who?" question, (e.g.) "Who wants the last piece of cake?". The act of raising your hand takes the place of uttering an expression that singles you out. The demonstrative act is a surrogate for an act of singular reference. With 'I', it seems to me, the surrogate relation is reversed. The producer of 'I' is performing the symbolic surrogate of overtly pointing to him or herself. To put the point metaphori-
cally, it is as if 'I' in language were an arrow, and the arrow always points to the agent who produced it.

If this idea were to prove analytically useful, we would have a DRT treatment of 'I' that accomplishes what we are seeking. We want (a) a rule that does not assimilate 'I' to other devices of singular reference and (b) an interpretation of this rule that explains how a first person use of 'I' accomplishes reflexive reference. On this hypothesis, both constraints are satisfied. In the first place, 'I' will not be assimilated semantically to a proper name. According to the standard DRT account of proper names, such terms refer causally, not reflexively. Secondly, 'I' will not be classified with the definite descriptions. Definite descriptions get their references fixed in the normal case by means of unique satisfaction. As we saw in our remarks on both Frege and Reichenbach, 'I' does not express any descriptive sense. Thirdly, 'I' will be distinguishable from the demonstratives, for no demonstration need accompany it. As it might be expressed, a demonstration need not accompany 'I' because the demonstration has already been accomplished by use of reflexive language. Fourth and finally, 'I' is seen to possess a demonstrative-like aspect. To interpret the rule under which 'I' operates we require, or so I have suggested, the notion of a surrogate demonstration.

The interpretive idea that 'I' behaves, on an occasion of utterance, like a surrogate demonstration will no doubt be viewed as naive in several respects. I want therefore to consider three objections to it. Naturally, I do not find these objections decisive against this idea, but I do find them helpful in developing my interpretation of Kaplan's analysis.

The semantic account now on the table, modeled upon the K-rule, is: In each of its first person uses, 'I' refers to the producer of that very token. Let us call this the Revised K-rule. But suppose the wicked witch, saying nothing, thinks to herself, "I am the fairest of them all." Assume that this use of 'I' serves in thought to distinguish the witch from them all. Surely it would be wrong to suggest, wouldn't it, that 'I' accomplishes this task in virtue of a surrogate of self-pointing?

This is the counterexample that will be presented against us by a proponent of ITR. The objection is that the proposed explanation is insufficiently generalizable. It presumably cannot apply to one form of I-reference, namely the use of 'I' in an unspoken thought.

Our earlier response to Castañeda indicates how we should answer this objection. If, as it seems to me most plausible to hold, the thought, qua internal accusative, is experiential but not propositional, then it lacks truth-value. If it lacks truth-value, then it is not relevant to semantics. If, on the other hand, it is assumed that truth-value is to be predicated of the witch's internal accusative, then the mental image she presents to herself of the content of 'I' is simply a representation of what she would have demonstrated in a surrogate way had she spoken 'I'. (Note that there is no problem about ignorance in this case.) The internal accusative in this counterfactual way will be modeled on the external proposition.

Consider an analogous case: the written instance of 'I'. Since no act of utterance occurs in this case, it may seem that our interpretation is inapplicable. Now, written uses of 'I' fall into one of two categories:
the fictional or the autobiographical. When 'I' is employed by a char-
acter in a novel, it refers reflexively in virtue of a surrogate self--
pointing to the fictional individual who produced that very token. Hence,
the Revised K-rule, interpreted according to the referential mechanism
we have accorded it, and qualified by a proviso notifying the reader of
the fictional character of the discourse, is an appropriate formulation of
the semantic rule controlling the fictional 'I'.

The autobiographical use is more problematic. One could adopt a
simple amendment to the Revised K-rule: In each of its first person
written instances, 'I' refers to the producer of that very written in-
scription. However, now the idea of explaining the notion of broad re-
flexivity in terms of surrogate self-demonstration, if not wrong, begins
to look less perspicuous. The temptation to make up an autobiographical
structure of reference, with different propositions and different con-
tents, should be resisted. The more reasonable approach is to adopt a
counterfactual application of our antecedent analysis:

If a writer S uses 'I' as an element in a written sentence of
autobiography A, then S has performed what would have
been an act of surrogate demonstration had 'I' been used
by S to communicate A verbally.

Of course the pragmatic circumstances of the case may make it im-
possible to know to whom 'I' refers in any given autobiographical use.
The point to be underlined, though, is that even in a first person use
of 'I' where no utterance-act occurs, the idea of surrogate demonstra-
tion, applied counterfactually, is the correct explanatory notion.

The use of 'I' in written autobiography provides the model for in-
tegrating the I-thought into our DRT analysis:

If a speaker S uses 'I' as an element in an I-thought T, he
has performed what would have been an act of surrogate
demonstration had 'I' been used by S to communicate T
verbally.

This postulate can be viewed as a transformation rule. It gives
the mechanism for transforming the use of 'I' in an internal accusative
into a use of 'I' suitable for producing a proposition with which seman-
tics can deal. As with the case of the autobiographical use of 'I' we re-
ject the intentionalist maneuver in favor of a counterfactual application
of the Revised K-rule. This reemphasizes the adherence of DRT to a the-
ory that holds consistently to the primacy of the linguistic.

The explanatory device we have employed in our account will seem
naive for a second reason. A demonstration would seem to be subject to
mishaps that are inapplicable to a use of 'I'. Any given demonstration
may fail to designate any object at all. It may fail to designate a unique
object. And it may inadvertently designate an object other than the one
intended by its performer. Are we to suppose that the semantic rule
governing 'I' allows for similar possibilities?

This objection forces the recognition that the Revised K-rule is
not sufficiently revised. The type of demonstration of which a first per-
son use of 'I' is an instance is, I believe, subject to severe semantic re-
stricitions. In this connection I wish to introduce the notion of a referential accident.  

To introduce the concept we require a distinction between pragmatics and semantics, and, within this framework, some Kripke-like distinction between the speaker's referent, on a given occasion of utterance, and the semantic referent of the linguistic expression used to make that utterance on that occasion. (Kripke, 1977) Roughly this distinction (drawn from Grice, 1957) holds that there may be a difference between the object to which a speaker intends to refer by means of an expression, on an occasion of utterance, and the object to which his words actually refer in language. A referential accident thus occurs when a singular term is unable to successfully accomplish either the reference that semantic convention demands or the reference that the speaker of the term pragmatically intends (or both).

Reference failure consists of course in the failure of an expression to designate any referent whatsoever. Some philosophers, influenced by Descartes, take 'I' to be immune to reference failure. The proposition expressed by the sentence "I do not exist" might be thought necessarily false. In my opinion one of the advances of contemporary possible world semantics is the insight that even first person, existential statements are contingent. According to that semantics, the sentence "I do not exist" is possibly true. It is true "in," "of" or "at" all those worlds in which I do not exist. Hence our semantics of 'I' need not build in an immunity to reference failure.

The issue of referential ambiguity is more complex. Many context-relative indexicals, even sortally disambiguated ones such as 'this phrase', fail (absent an accompanying demonstration) to designate a unique individual. If 'I' referred reflexively on the model of 'this very phrase', then its reference could not fail to be unique. But 'I' always refers to its producer, and one might begin to wonder whether there might be, in some instances, more than one producer of 'I'. Imagine a case of speaking-through-the-mouth-of-another. The speaker, qua mouthpiece says, "I want you to come to me", but the person responsible for the sentence is the mad psychiatrist. Could 'I' be convicted of referring to both the mouthpiece-producer and the psychiatrist-producer?

I would think not. If you were on the receiving end of this communication, you may not know whom to identify as the person wanting you. But this would seem to be a feature of the pragmatic context and not be salient to the semantics of 'I'. Depending on how the details of the story are filled in we may want to identify the psychiatrist as the referent of 'I' or we may want to identify the mouthpiece as the referent, but from this it does not follow that 'I' could refer to both producers. 'I' would seem not to be subject to that type of possible ambiguity.

Consider, finally, the issue of referential misfiring. I intend to refer to Smith by means of the expression 'the man in the closet reading Shakespeare'. But since it is Jones rather than Smith who satisfies this description, the reference has misfired. Referential misfiring can thus occur in cases where the speaker's referent (the object the speaker has in mind) may diverge from the semantic referent (the object to which the words in language actually refer). Is it possible for the
speaker of 'I' to refer inadvertently to some producer other than the
producer to whom he intends to refer?

Consider this case. Ernst Mach looks into the mirror and says, "I
am a shabbily dressed academician". Unbeknownst to Mach the mirror is
angled in such a way that he is in fact viewing the image of someone
other than himself. That person is the shabbily dressed academician.
Someone might claim, on the basis of such a description, that Mach in­
tends to be making a statement about the person in the mirror. If so,
the argument continues, the semantic reference of 'I' is Mach, and the
speaker's reference is the person reflected in the mirror. On the basis
of such a possibility it might be claimed that 'I' is indeed subject to
misfiring.

This conclusion betrays a confusion. We can agree that Mach's
statement expresses a false belief. We would say of him, "He mistakenly
believes that he himself is the shabbily dressed academician". We can
agree, moreover, that if the statement were to be true, then it would
have to be about the person reflected in the mirror. But this counter­
factual formulation of the conditions under which Mach's I-statement is
ture does not introduce a speaker's referent possibly divergent from
the semantic referent. Mach's use of 'I' signaled his intent to use that
term in accord with the appropriate semantic rule, and I am prepared to
argue that this rule is sufficiently definite to rule out the possibility
that he could intend that word to refer to someone other than himself.

The argument I have in mind is best seen as a response to an ob­
ject that may seem irresistible at this point. Surely, a defender of
ITR will say, the speaker of 'I', in this case Ernst Mach, must "intend"
or "mean" to be referring to himself by use of that pronoun. Isn't that
"intention" or "meaning," the objector will continue, the salient semantic
feature—a feature that your analysis entirely omits? And isn't it possi­
ble for that "intended" reference of 'I' to diverge from what you call
the "semantic" reference?

The defender of a DR approach to the first person will point out
that the concept of an intention is ambiguous in this objection. We must
distinguish between a general or first-order intention and a more spe­
cific or second-order intention. When a speaker utters (e.g.) 'This __',
he must have the first-order intention to employ that term in conformity
to a certain rule ("'this' refers to things in one's near spatial and tem­
poral range"), but this intention is not sufficient to fix a reference. The
speaker must have the additional, second-order intention to refer to
some given item among those things in that range.

The predication of reflexivity to I-reference amounts to the claim
that this second-order sense of 'intention' is inapplicable. Intention only
comes in at the first-order level. That is to say, if the speaker of 'I'
has the first-order intention to use that term first-personally, then that
speaker, qua producer of that very token, is singled out automatically.
In the case at hand, Mach indeed "means" or "intends" to be using 'I'
first-personally—not as a proper name or definite description. But given
this, there is no further second-order intention controlling his first
person use of 'I'. The character governing 'I' assigns the reference to
that term independently of what is going on in Mach's head.
It is for precisely this reason that with the word 'I' (in its first person employment) there is only semantic reference. This is why no sense can be given to the thought that 'I' could fix a speaker's reference divergent from its semantic reference. There simply is no speaker's intended reference. Hence, the conditions requisite for the possibility of referential misfiring would seem to be absent in the case of 'I'.

Our interpretation of the Revised K-rule is based upon the idea of a symbolic demonstration. But, as just noted, 'I' operates under certain semantic restrictions inappropriate to the nonlinguistic act that informs its analysis, i.e., demonstration. In particular we must build into that analysis the apparent immunity of 'I' to the accidents of referential ambiguity and referential misfiring. A more satisfactory formulation is:

**Final Revised K-rule:** In each of its first person uses, 'I' refers to one and only one agent, the producer of that very token.

Finally, our characterization of first person reflexivity in terms of surrogate demonstration may seem objectionable on the grounds that it fails to recognize the use of 'I' in oratio obliqua. Here I would argue that we require not so much a revision of or amendment to the Revised K-rule, but rather a qualification. In the sentence, "Phillip believes that I ate the cherries," nothing significant turns on the fact that 'I' falls within the scope of the prefix 'Phillip believes that'. The 'I' still reflexively refers to the unique producer of that token.

Here Castañeda provides a distinction that is conceptually helpful. (Castañeda, 1981, 290) It is the difference between the external and internal construction of a singular term. Consider Chisholm's example, "Columbus believed that Castro's island was China". If the singular expression 'Castro's island' is construed as internal to the psychological prefix, then it is being suggested, anachronistically, that Columbus himself could have referred in some language to Cuba as Castro's island. To avoid this interpretation the phrase 'Castro's island' should be construed externally: "Columbus believed of Castro's island that it was China". 'Castro's Island', now falling outside the scope of the psychological prefix, shows how the speaker of the sentence referred to Cuba, but implies nothing about how Columbus himself referred to Cuba. Castañeda claims that all indexical occurrences must be construed externally; certainly the occurrence of 'I' demands that construction. Thus in the case of Phillip and the cherries we get, "Phillip believes of me that I ate the cherries". The *oratio recta* occurrence of 'me' in this sentence is a first person use of that pronoun, and, as such, 'me' reflexively refers to the producer of it. We are then free to construe the *oratio obliqua* occurrence of 'I' as an anaphoric use, referring via the reference of 'me'.

What have we accomplished? Chisholm holds that since 'I' primarily functions in language to express the subjective point of view, only a semantics that presupposes the primacy of the intentional will do justice to this term. Castañeda identifies himself with this view by arguing that the Direct Reference Theory omits an essential "internal accusative of thought". This internal accusative of thought, though having no neces-
sary connection with language, must nevertheless be central to semantic analysis.

In opposition I have argued that if we look at the specific semantic analyses of 'I' proposed by defenders of the intentionalist approach, we find them conceptually flawed. Second, the alleged internal accusative of thought, far from being essential to semantics, is irrelevant to that enterprise. Third and finally, a straightforward description of how 'I' operates in natural language, as against the prejudice that it expresses the subjective point of view, reveals that our intuitions about the pronoun can be accommodated nicely on nonintentionalist assumptions. I argued that it is perspicuous to picture 'I' as a surrogate demonstration, and then to qualify the resulting semantics by building immunity to referential "ambiguity" and "misfiring" into the meaning-rule under which I operates. Though there are no doubt problems with the alternative picture I have offered, in my opinion this linguistic account of 'I' is considerably more successful than its intentionalist alternatives.

ENDNOTES

1 I wish to thank the following individuals for valuable help in formulating the positions taken in this paper: R. Albritton, P. Foot, J. Kupperman, A. Casullo, R. Coburn, and D. Kaplan. I am also indebted to the anonymous referee for the Philosophy Research Archives. I am of course responsible for the views expressed in the paper and for whatever mistakes it contains.

2 This is a reconstruction of an argument originally given by Kaplan. See, Kaplan, 1977.

3 Perhaps McGinn's insistence on the AE/EA ambiguity here is in part a function of his consistent substitution of the word "special" for the word "particular" in the Frege quotation. This is odd because he footnotes the standard Strawson anthology (which uses a translation by A.M. and Marcelle Quinton), and in that translation the passage reads, "Now everyone is presented to himself in a particular and primitive way. . . ." (emphasis mine). See McGinn, 1983, p. 58.

4 I do not, however, want to be identified with the claim that sensations can be private—in any sense. In order to claim that only I could have just this headache we need to be able to apply the distinction between numerical identity and qualitative indistinguishability to sensations. I agree with Malcolm that this region of discourse seems not to tolerate that distinction. See, Malcolm (1977), pp. 104-132.

5 I use brackets here to distinguish propositions from sentences.


7 Those familiar with Castañeda's "'He': A Study in the Logic of Self-Consciousness," 1966, will recognize this case as the one originally suggested by Norman Kretzmann in footnote 12.

8 Castañeda is primarily responsible for bringing the attention of the American philosophical community to the importance of the indirect reflexive in the analysis of 'I'. But in attempting to show that 'I' has an
irreducible sense he did not think to use the "he*" device as a way of distinguishing between first person and non-first person uses of 'I'. For that, Geach's original note is more helpful. My use of the term 'proxy' comes from him. See, P.T. Geach, 1957; H-N Castañeda, 1966.


10 Strictly speaking, the semantics of 'here' and 'now' are slightly more complicated. Kaplan recognizes both a demonstrative and a pure indexical use for those two terms. If I were to say "We were here" while pointing to Tahiti on a map, this would constitute a demonstrative use of 'here'. On the other hand, if the odd circumstance were to arise where someone says "I am here" allowing the expression 'the place where I am' (interpreted attributively) to fix the reference of 'here', this would constitute a pure indexical use of 'here'. Kaplan does not mention a third case. You are lost, in a fog, on a mountain. You yell "I am here!". There are two interpretations of this case. The first construes the use of 'here' as a nonreferential. Under this interpretation while the term 'here' directs the attention of others to your spatial location, it does not specifically pick or single out that location. According to the second interpretation, 'here' operates in a demonstrative-like fashion. It is as if the utterance 'here' were "pointing to" a specific spatial location, the location where the voice uttering 'here!' originates. It is the second interpretation of this third case that we will exploit to provide a model for how 'I' reflexively refers.

11 This concept was first suggested to me by Rogers Albritton.

12 This may be the point Sidney Shoemaker was trying to express some time ago when he noted that the speaker of 'I' had "no latitude" in the determination of the reference of that pronoun. See, Sidney Shoemaker, 1968, 559.

REFERENCES


Chisholm, Roderick: 1976, Person and Object (La Salle, Illinois: Open Court Publishing Co.).


