

# 10

## The First Person\*

### I

I will concentrate here on the perplexities some philosophers have felt concerning the simple first person pronoun 'I'. The genesis of these reflections is a fairly recent invitation to Barcelona<sup>1</sup> to give a talk about my views on David Kaplan's manuscript 'What Is Meaning? Explorations in the Theory of *Meaning as Use*' (n.d.),<sup>2</sup> as well as on his classic publication, 'Demonstratives' (1989).

Were I to be making a general discussion of Kaplan's recent material, I would emphasize my enthusiasm for his general approach.<sup>3</sup> There is one aspect of his

\* The present paper was delivered at the conference 'Saul Kripke: Philosophy, Language, and Logic' held at the City University of New York, Graduate Center on January 25, 2006. There has been substantial revision of the original talk, which was delivered without a written text, but I have not entirely eliminated the conversational tone. I give special thanks to Gilbert Harman and Robert Stalnaker for showing me that one of my criticisms of David Lewis in the original version was too strong.

My paper 'Frege's Theory of Sense and Reference: Some Exegetical Notes' (chapter 9) was unpublished and unknown to the audience when I gave the present talk—thus my presentation included much overlapping material, since it was highly relevant to the topic. In the current version I have cut down on the overlap but certainly not eliminated it. There are some significant considerations raising problems for Frege that should have been in Kripke (2008) (included here as chapter 9), but were not known to me when I wrote its final version; nor were they known when I gave the original presentation of this talk. See pp. 303–04.

<sup>1</sup> I was invited by the Logos Group to give three lectures at the University of Barcelona, Spain, in December 2005. Joseph Macià suggested Kaplan's papers as a possible topic for one of the lectures. For further remarks on the background of this topic in the philosophy of language, especially due to the influence of Castañeda, see note 12 below.

<sup>2</sup> From now on I will call this manuscript 'Meaning as Use'. The version that I have is subtitled 'Brief Version—Draft #1'. Of course, the slogan 'meaning is use' is derived from Wittgenstein, and certainly Kaplan's manuscript to some extent is influenced by Wittgenstein's later work and even some of his most famous examples, but the manuscript should probably not be regarded (nor was it so regarded by its author) as adopting a 'Wittgensteinian' approach.

<sup>3</sup> In particular, I share (and have always shared) his broad conception, as I understand it, of what should be included in semantics (and not relegated to pragmatics). Anything that a language teacher should regard as part of the teaching of the meanings of a particular language, as opposed to customs and sociological facts about speakers of the language at a particular time, should be included in semantics. I have never understood why some people wish to restrict semantics so as to include only what is clearly to be given by truth-conditions, excluding, among other things, the study of indexical expressions. One may think that the issue is purely terminological, but I have a strong feeling that

approach that I would also applaud in general terms, but that I would warn may lead one astray in its particular application to the main theme of this talk. One should not, he argues, think of the task of the linguist or the semanticist—as, for example, Quine does in some of his writings—as analogous to that of translating utterances into one's own language. That presupposes the semantics of one's own language and doesn't get us very far. Rather, the linguist or the semanticist gives a description 'from above' of the uses in the community.

Kaplan refers to what some philosophers have called 'scientific language',<sup>4</sup> and assumes that the description 'from above' is formulated in such a language. The

this is not entirely so, that the opposite conception will lead one astray. For example, Ludlow and Segal (2004) think that on Gricean principles 'but' and 'and' 'literally mean the same thing' (424), though they differ in conventional implicature. Similarly, they think that 'a' and 'the' are synonyms in English (424), even though they state how they are used differently in English (in this view, they have surprised many philosophers who might have followed them thus far). Following Kaplan, as I understand him, 'tu' and 'vous' (as the polite second person singular) are not synonymous in French. In contrast, changes in French attitudes as to when it is appropriate to use 'tu' are matters of the changing sociology of the French, not of changes in the language. (Perhaps one can imagine cases where the distinction is not sharp.) Distinctions of Gricean 'conversational implicature', as in my own suggested treatment in Kripke (1977) of the referential-attributive distinction for definites, are not distinctions in the language. (At the end of Kripke 1977, I suggested that the same strategy might apply to indefinites, as was carried out by Ludlow and Neale 1991.) Ludlow and Segal (2004) should not have considered their own strategy to be a case of the same one that was used in the two papers just mentioned. Note that the issue has nothing to do with whether Ludlow and Segal are correct in their views about 'a' and 'the'. However, they are not entitled to say that, on their view 'a' and 'the' are two expressions with different spellings, but the same meanings—synonyms, rather like "gray" and "grizzled" or "grisly" and "gruesome" (424). In contrast, 'and' and 'but', on my view, though not theirs, are not differently spelled expressions with the same meanings.

I hope I understand Kaplan correctly when I express agreement with him about this. He has a great deal of illuminating material, in particular, about a novel conception of logical validity that accords with this conception of semantics. His remarks about 'oops' and 'goodbye', and about pejoratives, looked at in terms of a use theory of meaning, are also very illuminating.

<sup>4</sup> Kaplan mentions Quine as an originator of the conception that scientific language should not contain indexicals, tense, and the like, even though Kaplan himself disagrees with the view (which he states Strawson got from Quine) that such devices are not susceptible to logical study. I believe that in conversation he also mentioned Russell in connection with this conception of 'scientific language'. Even though he also (see below) clearly would recognize that this conception has little to do with what is allowed in actual scientific papers, it nevertheless influences his idea of how to describe a language 'from above'.

Quine is a philosopher who might be described as 'pro-scientific' (by an admirer of the orientation), or 'scientistic' (by a detractor). On the other hand, the later Wittgenstein was probably 'anti-scientistic'. However, he has a similar conception of 'scientific language'. In *Philosophical Investigations* he writes: "I" is not the name of a person, nor "here" of a place, and "this" is not a name. But they are connected with names. Names are explained by means of them. It is also true that physics is characterized by the fact that it does not use these words' (Wittgenstein 1953: §410) (I have altered the translation of the last sentence, which is weaker and less puzzling in the printed version: '... it is characteristic of physics not to use these words'.) What does he mean here? Certainly, as I have said, not that such indexical terminology never appears in physics papers. It might be claimed that such terminology never appears in physical laws, but once it is granted that many terms are explained by them, this strikes me as dubious. Moreover, physics may say that such-and-such a physical quantity has a certain value now, but it is decreasing. Even more puzzling to me (if my translation is right) is the claim that this restriction (the absence of indexicals and demonstratives) tells us what physics is.

so-called scientific language itself would contain neither indexicals nor tense, but be generally stated as applicable to arbitrary speakers, places, times, and the like; if modality is involved, to arbitrary possible worlds as well. Moreover, all this must be done coolly. For example, as Kaplan says, one must be able to describe words expressing anger without getting angry at the same time. This is relevant to what others have written on historiography. We have at one extreme the view of, for example, C. G. Hempel (1942), which takes historical writing to be little different in principle from writing in physics, involving general laws, confirmations, refutations, and so on. The other extreme says that history employs a particular method of *Verstehen*. I think the second view certainly has something to it: the historian is trying to put himself in the position of his subjects to see what they themselves might have thought. Some have given a strong formulation: if you write about any historical character, you should literally try to *become* that character.<sup>5</sup> Whatever one might say about history, surely the ‘coolness’ requirement as Kaplan states it is correct for the description of a language ‘from above’.<sup>6</sup>

What is description ‘from above’? (I perhaps add something here to Kaplan’s formulation.) The description, first and foremost, is a description of how the language is used, but it also has an instructional aspect. If language can be described completely and correctly ‘from above’, in a neutral indexical-free language, the description (of, say, English) should be usable as an *instruction manual*, a set of imperatives for a foreigner wishing to learn English. The instructions themselves should, if given for this purpose, be stated in the foreign language. Thus they will tell the foreigner that ‘goodbye’ is conventionally used in English when taking leave. Even when only truth-conditional semantics is in question (or truth-conditions with respect to indices—such as speaker, time, possible world described, and the like), the description from above should be usable not just as a description but also as an instruction manual for a language learner.

Now, in ‘Meaning as Use’ (in the section ‘Meaning vs. Uses’), Kaplan writes:

Consider the indexical ‘I’. What does it mean? An initial answer might be that it is the first person pronoun. But this is a kind of functional description. What does the first person pronoun *mean*?

As I said, I was invited to speak on Kaplan, and under such circumstances one naturally emphasizes those points where one disagrees. (Having been invited to

<sup>5</sup> When I was in college, I wrote a paper on this topic called ‘History and Idealism: The Theory of R. G. Collingwood’, which I never published, though an expert in the field did recommend that I do so. In the paper, I remarked that if you write about Hitler, you should not try to become Hitler; this would be a very dangerous idea. Some of you have probably heard of the writer David Irving—he originally started with very respectable publishers, and he is at any rate *something* of an embodiment of this approach (see, e.g., Irving 1977, where he explicitly says that he will attempt to describe the war through Hitler’s eyes). Unfortunately, from this book onward he became increasingly successful at fulfilling my youthful fear that it was a dangerous idea.

<sup>6</sup> However, even in that case there is something to be said for the method of *Verstehen*. One might have to experience anger to understand descriptions of words expressing anger.

give a paper on Kaplan's material, my audience would not have been so enthusiastic if I could only remark, 'Yes, I agree, and this says it all'.) So here I will comment on one point on which I have some agreement but on which I ultimately diverge significantly from Kaplan—namely, his views on the proper treatment of the first person pronoun. And I will talk about some other authors as well.

Kaplan gives the following example to show that in the case of 'I' a proper semantical treatment is not provided by a definition but rather by an account of how the term is used:

For example, my *Webster's Third* provides, 'the one who is speaking or writing' and they quote the Psalm 'I shall not want' in order to drive the point home. (This caused me to imagine sitting in the back of the auditorium at a lottery award ceremony, and whispering to the psalmist, 'which of the people here won the ten million dollars?', and he whispers back 'the one who is speaking' or, equivalently, according to *Webster's Third*, 'I did; I shall not want'.) (Ibid.)

Two or three comments here. I, of course, am in agreement<sup>7</sup> with the semantical point Kaplan is trying to make about how to explain the word 'I'. But does 'the psalmist' say 'I shall not want'? Many educated Americans appear to think that the psalmist (like other biblical authors) wrote, or completed his work, in 1611.<sup>8</sup> In fact, 'the psalmist' who 'wrote' the 23rd Psalm must have consulted the published standard Hebrew *Urtext*. It does not contain a Hebrew equivalent of the separate word 'I' at all: as in many languages, 'I' is used only for emphasis and is, in most cases, simply a suffix or prefix to the verb.<sup>9</sup>)

Now, in some respects this observation may even support Kaplan, because his point is that you should not look for the meaning of 'I'. Kaplan's basic thought is that the search, as in Frege, for what the word 'I' means, or the *sense* of 'I', is obviously a mistake. The correct semantical account of 'I' is wholly given by the 'scientific language' in which its truth-conditions are neutrally expressed from above: when a speaker *S* says 'I . . .' what he says is true (or true of the possible world he is thinking of at the time, if that is relevant) if and only if *S* . . . This is a purely general statement, and it wholly determines the semantics of 'I'.<sup>10</sup> The point (that a definition of 'I' as a term denoting the speaker is not really in

<sup>7</sup> However, Kaplan is not completely right. If I were writing a paper I could say 'the present writer thinks there is a mistake'. 'The present writer' may or may not be synonymous with 'I', but it is standard in some academic writing to use it to replace 'I', perhaps to be a little more formal or impersonal. 'The present speaker' in the same sense is rarer, certainly not a stock phrase like 'the present writer', but maybe on some occasions it can be used in that way.

Or suppose someone has written, 'all Americans support such-and-such'. One could object, 'not *this* American', meaning *not me*. One could no doubt imagine many other such cases.

<sup>8</sup> I believe, if my memory is correct, that I read in the *New York Times* something like this: 'The Bible says . . . [or 'the Old Testament says'], as opposed to the more recent Bible versions.' My father heard a Christian fundamentalist radio preacher say 'until, or rather 'til, as scripture says, . . .'

<sup>9</sup> In the perfect and imperfect respectively (in Biblical Hebrew).

<sup>10</sup> This point is independent of Kaplan's emphasis on meaning as use, as in the later Wittgenstein. It would be compatible with a truth-conditional (or truth-conditional with

question) can only be strengthened when we consider the existence of languages in which the first person is expressed exclusively by a prefix or suffix (or where this is usual and 'I' or its equivalent occurs only in cases of special emphasis).<sup>11</sup>

Kaplan notes that *Webster's* itself realizes that its attempt to define 'I' 'won't do' and that it goes on to say 'used . . . by one speaking or writing to refer to himself.' He comments 'Now here they have finally given us what we need to know, how the expression is *used*'.

Consider Kaplan's distinction between character and content. The character gives a general rule for the use of 'I', and the content will depend on one's view of content. If one takes the simple propositional view, it will be about the speaker; or, as Kaplan states, it doesn't have to be a speaker, it could just as well be a writer, or a thinker, thinking to herself. We will return to this.

Kaplan's treatment of the first person in 'Demonstratives' (1989) is rather strongly influenced by Perry's criticisms of Frege (Perry 1977) on first person statements (and other demonstratives).<sup>12</sup> I myself have dealt with Frege's views on these issues, including the problems Perry raises for Frege, and their relation to Frege's much-discussed views on indirect quotation and his less-discussed views on direct quotation (see chapter 9). But here I not only want to talk about what is *true according to Frege* but also about what is *true according to the truth*, or, that is—to use a predicate that I should like to think is coextensive—true according to *me*.

In conversation, Kaplan has acknowledged that 'scientific language', in the sense that he has used the term, is obviously not satisfied by the language of scientists in a lab, who use tenses and indexicals all the time. It is not satisfied in

respect to various indices, such as speaker, time, possible world described, etc.) conception of semantics.

At the end of Anscombe (1975), discussed below at some length, she attributes to J. Altham the remark that such a rule about 'I', viewed truth-conditionally, has a problem of sufficiency: 'How is one to extract the *predicate* for purposes of this rule in "I think John loves me"? The rule needs supplementation: where "I" or "me" occurs within an oblique context, the predicate is to be specified by replacing "I" or "me" by the indirect reflexive pronoun' (65). The grammatical notion 'indirect reflexive' is explicated in Anscombe's paper.

In Kaplan (1989:505) the two main rules are that "'I" refers to the speaker or writer' and that 'I' directly refers. Perhaps Kaplan thinks that a direct reference account of 'I' and 'me' gives an adequate treatment of examples such as Altham's. I am sympathetic to such a viewpoint myself. It would have been good if this had been spelled out.

<sup>11</sup> I spoke with Kaplan about this, and he said he was familiar with the example of Latin.

<sup>12</sup> One should mention, whenever one talks of the first person as a special subject in contemporary philosophy, the papers of Hector Neri Castañeda (1966, 1968, and others), who more than anyone else made this a special topic for the philosophy of language (both first person sentences and their relation to the indirect discourse locution exemplified by 'Betty believes that she herself . . .'). See also Geach (1957a) and Prior (1967), cited by Lewis (1983:139). I think also of Wittgenstein (1953), as mentioned below.

There is a mutual influence between Kaplan and Perry. Perry himself mentions an earlier version of Kaplan (1989) in Perry (1977).

Of course, contemporary philosophy of mind, as well as philosophy of language, has emphasized the difference between first and third person points of view, and this is also related.

scientific papers either. No scientific journal would reject a paper for failing to use exclusively 'scientific language', especially tense. I think it is relatively hard to give genuine examples of tenseless sentences about particular ordinary objects (though not about mathematical objects or the like) in natural language. Some examples that I have seen in the literature are not really tenseless.<sup>13</sup> Scientific language in the sense in question is a philosophers' invention, spoken by no one. In spite of Kaplan's recognition of these facts, this conception of what can be stated in a 'scientific language' is important to his own account.

Now, Kaplan calls some statements Frege makes about the first person 'tortured' (1989:501), though later (533) he says that reinterpreted in the light of his own theory, Frege could be thought of as talking about the character of 'I', and that under such an interpretation this passage (which supposedly 'has provoked few endorsements and much skepticism') could be defended as essentially correct. Kaplan goes on to say how 'a sloppy thinker' might misinterpret the situation. Given his earlier characterization of the passage, and taking into consideration the influence of Perry (1977), I think that Kaplan really thinks that the sloppy thinker is Frege himself.<sup>14</sup>

Frege writes:

Now everyone is presented to himself in a special and primitive way, in which he is presented to no one else. So, when Dr Lauben has the thought that he was wounded, he will probably be basing it on this primitive way in which he is presented to himself. And only Dr Lauben himself can grasp thoughts specified in this way. But now he may want to communicate with others. He cannot communicate a thought he alone can grasp. Therefore, if he now says 'I was wounded', he must use 'I' in a sense which can be grasped

<sup>13</sup> For example, Sider's impressive book (2001) gives as examples 'World War I occurred after the American Civil War' and 'There existed dinosaurs before the appearance of this book'. Neither of these sentences can change their truth values if uttered at different times, but to me it is obvious that both are past tense; I don't know exactly what someone would have in mind imagining them uttered before World War I or before the appearance of 'this book'.

Another example Sider gives, 'It is raining on 28 June 2000', is dubious English, unless uttered on 28 June 2000, in which case 'It is raining today, 28 June 2000' is much better. After 28 June 2000 one must say 'It was raining on 28 June 2000', and before that date 'will be'. (Obviously, a particular place in which the raining occurs is presupposed.) Someone might of course be sure of the rain on the date in question, but not be sure of the date now, or not wish to commit herself. But in that case she should say, 'It either was, is, or will be raining on June 28 2000'. This is not a tenseless statement, but amounts to a disjunction of tensed statements (or, alternatively, applies a disjunction of tensed predicates). Probably something amounting to this disjunction is what Sider has in mind as the interpretation for his 'tenseless' statement, but his attempt at expressing it in English seems defective. His example was supposed to contrast with 'It is *now* raining,' which he gives earlier as an example of a tensed statement but which does not appear to be a proper English sentence. While genuine examples of such tenseless statements may exist in English, they do not occur nearly as much as has been suggested.

In spite of these remarks, I am *not* suggesting that we have no conception of a language giving the entire history of the world tenselessly. I think that we can imagine such a language. However, most of it will be a philosophical invention. It is an intelligible language but does not overlap with natural language to any significant degree.

<sup>14</sup> See chapter 9 (page 284).

by others, perhaps in the sense of 'he who is speaking to you at this moment'; by doing this he makes the conditions accompanying his utterance serve towards the expression of a thought. (1918–19:333, note omitted)<sup>15</sup>

Not only has the passage been discussed critically by Perry and by Kaplan (under Perry's influence), it has also been defended against Perry's criticisms in Evans (1981). What Frege is saying about the way everyone is presented to himself seems to me not to be at all unfamiliar. It is the familiar view, going back at least to Descartes, that I am aware of myself in a special first person way. However, Perry, and following him, Kaplan, both argue that for his view of the first person to go through, 'what is needed is a primitive aspect of me, which is not simply one that only I am aware of myself as having, but that I alone have' (Perry 1977:490).

Why does Perry think that this is needed? Well, the special first person Cartesian sense would have to be something like *the subject*, or *the thinker*. But who is that? Is there only one thinker, only one subject? If one reformulates it as *the subject for me*, *the subject that I am aware of*, by *being aware of my own thinking*, the formulation obviously runs into a circle. How can one avoid the circle? Only by there being a special quality, a primitive aspect of me, that I alone have. This is Perry's argument for his conclusion, and Kaplan follows him.

Following Perry, Kaplan makes two objections to Frege. First, he says:

I sincerely doubt that there is, for each of us on each occasion of the use of 'I', a particular, primitive, and incommunicable Fregean self-concept which we tacitly express to ourselves. (1989:534)

So far, Kaplan might just seem to be doubting the neo-Cartesian doctrine of a particular first person perspective (except to the extent that it is given by his theory of the 'character' of 'I'). However, he immediately goes on to assume that the theory must involve Perry's stronger conclusion that the self-concept in question would have to characterize its subject uniquely in a neutral language, and objects:

[E]ven if Castor were sufficiently narcissistic to associate such self-concepts<sup>16</sup> with his every use of 'I', his twin, Pollux, whose mental life is qualitatively identical with Castor's, would associate the *same* self-concept with *his* every (matching) use of 'I'. (Kaplan 1989:534; italics in the original)

One of Kaplan's basic points in 'Demonstratives' is the distinction between demonstratives and indexicals. Demonstratives (such as 'this') require some gesture or something else (such as pointing) to determine their reference, whereas

<sup>15</sup> All references to 'Der Gedanke' (Frege 1918–19) are to the Geach and Stoothoff translation, titled 'Thoughts', as reprinted in Beaney (1997) with the title 'Thought'. Note that the translation of the passage that Kaplan himself uses (1989:533) is the earlier Quinton and Quinton translation. As far as I can see, the differences do not affect the discussion.

<sup>16</sup> Kaplan plainly means to write 'such a self-concept'.

indexicals (such as ‘I’ and ‘now’) require only a general linguistic rule to determine their reference.<sup>17</sup> For example, ‘I’, when used by a given speaker, always refers to the speaker (or thinker, or writer, etc.); ‘now’, when said at a given time, refers to that time; and so on.<sup>18</sup>

Kaplan accuses ‘the sloppy thinker’ (Frege in the naïve interpretation) of holding a ‘demonstrative theory of indexicals’. It is as if one needed something other than the semantical rule for ‘I’, a subject somehow pointing to himself in a special inner way, to determine the reference of ‘I’, and similarly for ‘now’ and ‘here’ (Kaplan 1989:534–35). Descartes, and many who have followed him, might be accused of this mistake, except that it is hard to see that Descartes’s point was particularly semantical.<sup>19</sup>

Returning to the earlier view that Frege requires some unique qualitative description of the subject that the subject alone is uniquely aware of and that, in fact, uniquely characterizes ‘the subject’, it may be tempting to conclude that there can be only one subject in existence. Indeed, there may be some philosophers who have drawn such a conclusion, a special form of solipsism about minds, but I am not sure who.<sup>20, 21</sup>

<sup>17</sup> Perry later called them ‘automatic indexicals’ (1997). The speaker need not indicate specific demonstrative intentions on the occasion of utterance.

<sup>18</sup> Kaplan (1989:491) attributes to Michael Bennett the point that ‘here’ is usually an indexical, but sometimes is a demonstrative, as when one says ‘she lives here’ and points to a location on a map, and so on. Similarly, in a footnote on the same page, Kaplan concedes that the rule given for ‘now’ is too simple. If someone leaves a message on the answering machine ‘I am not at home now’, ‘now’ refers to the time when the message was heard, not the time when it was recorded. The opposite can be true: ‘I am in Italy now but will be in Belgium by the time you get this letter’ makes good sense. (My own example; I also changed the answering machine example a bit.)

As far as I can see, no such problems arise for ‘I’, as, indeed, in the first clause of this very sentence.

<sup>19</sup> Perry does in fact concede that some philosophers ‘have come to hold somewhat similar views about the self, beliefs about oneself, and “I”’ without being motivated by any semantical problems (1977:489). He thinks it is possible that Frege was simply writing under the influence of these views, but he thinks it more likely that it was the pressures of an attempt to find a theory of demonstratives compatible with his overall semantical framework that is responsible for Frege’s views.

<sup>20</sup> In the talk I suggested, though I wasn’t sure, that perhaps the early Wittgenstein (1961) could be an example. Since then it has been suggested that L. E. J. Brouwer might be an example; once again this is uncertain (see, e.g., Brouwer 1948).

<sup>21</sup> Let me here make the following digression concerning ‘other minds’. People argue that there *have* to be other minds: after all, everyone behaves similarly to me and since I have a mind they must, too. But some other people say one shouldn’t generalize from only one case. A reply to this objection might perhaps be that minds form a natural kind, so that an examination of one instance is sufficient to determine the basic features of the entire natural kind. But what’s really the trouble here is this: there’s lots of evidence that we are *not* members of the same kind, because various philosophers—or so-called philosophers—of mind state theories that would seem to me to imply that they *themselves* have no inner states (or if they do use expressions that purport to say that they have inner states, they give analyses which I know perfectly well are *not* compatible with genuine inner states). So, what explanation of their behavior can there be? Otherwise, they seem to satisfy criteria of sincerity, honesty, and intelligence (if one assumes that they have minds). So, they are obviously just what they themselves claim to be, that is, very cleverly programmed robots similar to a genuine human subject like *me*. In philosophy departments, at least, there seem to be many more of them than there are genuine human subjects. So, when I look at a random person who is not even a



So what's wrong with the argument that *either* each subject must be psychologically unique or, otherwise, any definite description of the subject must itself use an egocentric term, and thus run into a circle? Well, again, the people who argue this way are thinking of a language spoken by no one, the so-called scientific language. Since Dr. Lauben is the one speaking the language, by 'the subject' he of course means himself. If Rudolph Lingens<sup>22</sup> speaks of 'the subject', he means himself. There is no difficulty for Frege (nor indeed for Descartes), once we rid ourselves of the idea of a 'scientific language' spoken by no one, in supposing that the reference is determined in this way. Nor do we have to worry about the supposed problem of Castor and Pollux. None of these people speaks an impersonal 'scientific language' where the problem would arise. So each of them could determine the referent in the Cartesian–Fregean way, by his own acquaintance with himself.

But perhaps this is not the whole answer. Aren't all these people speaking German, a language in which 'I' (actually '*ich*') should mean the same thing for *anyone*? And isn't Kaplan right to say that the whole use of the word 'I' can be captured in a neutral way by saying that a sentence containing 'I' expresses a truth if and only if the rest is actually true of the subject—the thinker, or the speaker? Or if one doesn't wish to restrict oneself to truth-conditional utterances, even with respect to indices,<sup>23</sup> at least that 'I' in any sentence refers to the speaker (writer, thinker)? So, doesn't Kaplan's characterization (that is, the description of the 'character' in his technical sense) suffice for everything? And doesn't it give the 'content' in each particular case, which indeed is different depending on whom is being referred to by 'I'?

At first this may seem quite conclusive. Doesn't Kaplan's rule give a complete description of the matter? What else could be needed? Well, recall my remarks that the 'description from above' ought to be usable as an instruction manual for someone wishing to learn the language. Though Kaplan's explanation is all very well for some sort of descriptive anthropologist who may in fact have the concept of 'I', it would be very difficult to get it across to Frege (or anyone else who is presumed to lack this concept). So, for example, let Kaplan say to Frege or to anyone else (but if it is Frege, one should use German): 'If any person *S* speaking German attributes a property using the word "*ich*," then what *S* says or thinks is

philosopher, the chances seem much greater that that person is one of these cleverly programmed robots, since from the only sample I have I can be sure of only one genuine human subject (with a mind). So, the robots are right in saying that most of the humanlike people are robots—and I must be a very rare case with genuine inner states!

<sup>22</sup> In Frege's paper 'Der Gedanke', some characters with various interrelations are discussed. In addition to Dr. Gustav Lauben, one person discussed is Rudolph Lingens. Frege considers alternative cases in which Lingens knows Dr. Lauben personally or has only heard of him.

<sup>23</sup> Such philosophers as Donald Davidson and David Lewis have attempted to reduce the semantics of non-indicative utterances (or sentences) to cases where truth-conditional semantics do apply. In stark contrast to this picture, see Wittgenstein (1953: §23). In his later paper 'Meaning as Use', Kaplan intends no particular reduction.

true if and only if *S* has that property.’ But how can Frege use the word ‘*ich*’ on the basis of these instructions? Should he think, ‘*Hmm*, so how am *I* going to use the word “*ich*” on the basis of this general statement? Well, any German should attribute, say, being in pain or being a logician to *himself* if and only if the German is in pain or is a logician, as Kaplan says. So *I* should do this.’ Alternatively, Frege might remark, ‘So Frege, or Dr. Gustav Lauben, should attribute a property to Frege, or respectively to Dr. Lauben, using “*ich*” if and only if Frege (or Dr. Lauben) has the property. But *I* am Frege, so I suppose that I should use the word “*ich*” if and only if Frege has the property.’ Either formulation would presuppose that Frege already has the concept of *himself*; the concept he expresses using ‘*ich*,’ so here we really are going in a circle.

The point is that each one of us speaks a language that he himself has learned. Each one of us can fix the reference of the word ‘*I*’ by means of acquaintance with oneself, self-acquaintance. There is no requirement that this type of acquaintance is given to us by a qualitative description expressible in a ‘scientific language’ spoken by no one. This is so even if the language each of us uses is a common one—English, German, and so on. No one can grasp the rule for ‘*I*’ stated in the common language except by means of one’s own self-acquaintance. Otherwise, there would be no way of learning how that rule tells us to refer. This is what Frege means when he says that Dr. Lauben uses the word ‘*I*,’ thinking to himself, ‘he will probably be basing it on this primitive way in which he is presented to himself (Frege 1918–19:333). Frege also says, a bit before that, ‘The same utterance containing the word “*I*” in the mouths of different men will express different thoughts of which some may be true, others false’ (332). To put this matter in Kaplanian terms, the utterance<sup>24</sup> has the same character in the mouths of all speakers of the language but has different contents in the mouths of different speakers. I have explained in some detail in my paper on Frege how to put the matter in Frege’s own terms (chapter 9, pages 284ff). Put either way, this is possible because of one’s own self-awareness when one is speaking.

However, Frege’s discussion, as quoted above (298–99), of how Dr. Lauben communicates to others using ‘*I*’ (or ‘*ich*’) does confuse the issue. Frege says that when he wishes to communicate, he can hardly use ‘*I*’ in a sense he alone can grasp. But if that is so, it is easy to see how someone would respond that this alleged special and incommunicable sense of ‘*I*’ must be a chimera. Why should the primary sense of ‘*I*’ be something that one never uses in interpersonal communication? One might after all doubt that ordinary language is used in thought at all. Surely, its primary purpose is for communication.

Matters become even more problematic when Frege discusses what Dr. Lauben means by ‘*I*’ when he wishes to communicate with others. He conjectures that it is in the sense of ‘he who is speaking to you at this moment’

<sup>24</sup> Notice that by ‘utterance’ Frege here means a type, not a token. He is not following current technical philosophical terminology.

(Frege 1918–19:333). This can be understood by the hearer in a way that the primary sense of ‘I’ cannot. Kaplan, as I have already quoted him, wittily ridicules those (such as the writers of the definition in *Webster’s Third*) who wish to analyze the ordinary use of ‘I’ in such a way. Moreover, as I wrote in chapter 9 (pp. 288–98), the proffered definition of ‘I’ may not work. For example, perhaps the person I am addressing is at the same time being addressed by someone else. Then the description will not uniquely determine its object. In the same chapter I gave other objections of a similar kind. I add some objections that I had not thought of in the earlier paper (nor in the original version of this talk).

Suppose that the definition does correctly determine its object, and we don’t entertain Kaplan’s worries about its artificiality. There is yet another problem. What, after all, is the Fregean sense of ‘you’ in the proffered definition? Shouldn’t it be ‘the person I am addressing at the present moment’? But then the proffered sense of ‘I’ plainly goes in a circle.

Moreover, Dr. Lauben may think to himself, ‘Leo Peter realizes that I am wounded’, or, alternatively, ‘Is Leo Peter aware that I am wounded?’<sup>25</sup> Since Dr. Lauben is thinking to himself, surely (following Frege) he uses ‘I’ in the special sense that only he can understand. But how can he wonder whether Leo Peter has a thought that Peter cannot understand? Something is going wrong here.<sup>26</sup>

Surely, one must give an analysis of first person sentences where ‘I’ is univocal, whether used in talking to oneself (discouraged in our society, anyway), or in

<sup>25</sup> Remember that, for Frege, asking a question is a paradigmatic way of entertaining a thought without asserting it.

<sup>26</sup> See also my discussion soon below of Frege’s remarks on ‘yesterday’ and ‘today’, and my more elaborate discussion in chapter 9 (pages 277ff), and especially my remarks in note 74 on the objection of Gunnar Björnsson, and related objections by John Perry, concerning indirect discourse. Björnsson phrased his objection in terms of tense, but it could just as well have been phrased in terms of persons, as indeed has been emphasized by Castañeda (and others, see below). (Perry does mention the interpersonal case.)

The present problem is a sort of converse form of the same problem. Someone can use ‘I’ in an indirect discourse attribution to someone else’s thought about herself, even though the other person would not use ‘I’ or an equivalent expression, nor could she understand the expression as used by the subject in question—in this case Dr. Lauben. The paragraph following in the present paper is only a partial answer to the problem. Applying the principles of that paragraph to the present version, one must recognize that it is legitimate for the subject (say, Dr. Lauben) to attribute a thought using ‘I’ to someone else (say, Leo Peter), provided that the other person has the appropriate belief *about the subject*.

However, there is a reason specific to Frege that makes me say that for him these principles give only a partial answer to the problem. Frege’s theories, with or without the ‘I’ problem form of them, are in danger of running into a problem related to my own problem about exportation (chapter 11, this volume). The problem is with Frege’s apparent view that it is sufficient for a name (or pronoun) to designate a given person (and for its user to have a thought about that person) that it be defined for its user by a definite description designating the person (similarly for entities that are not persons). Frege (1892:153) appears to express this view in his well-known footnote on Aristotle and in his later discussion of the way various people may think of Dr Lauben (1918–19). Something must be done to fix the matter up, and if this can be done, the objection raised here could also be addressed.

diary entries (not so discouraged), or in communicating with others. If it is the sense determined by its subject's first person acquaintance with herself, how can it be used to communicate to someone else? Here is one possibility. The hearer is aware that each person, including the hearer herself, uses 'I' to refer to herself by direct self-acquaintance. Hence, knowing what this is in one's own case and taking it to be the same way for others, one understands what the first person statement is, even though it has a sense that is, strictly speaking, incommunicable to the hearer.<sup>27</sup>

Similarly, according to my own understanding of Frege, at no later time can I have the thought I expressed with 'now', and at no later date can I have the thought that I express with 'today'. Nevertheless, I can understand a piece of writing written in the past using 'now' or 'today', similarly to the way I can understand someone else's utterance of 'I'.<sup>28</sup> Therefore, on my view, Frege was wrong on his own theory to say that one could express the same thought using 'yesterday' as one previously expressed using 'today'. Ironically, Kaplan, who is critical of Frege's discussion of 'I', commends his remarks on 'yesterday' and 'today'. My own view is that, from a Fregean standpoint, Frege's remarks on 'yesterday' and 'today' cannot be defended, while his remarks on the first person and the present are correct from a Fregean standpoint (with the exception of his discussion of the ambiguity of 'I', which I have criticized above).

In fact, however, when I discussed these issues with Kaplan, he said that he had come to accept my point that someone must have a concept of the self to follow the general direction for the use of 'I', and attributes his stronger original statements to 'irrational exuberance'. And probably Kaplan did not really

<sup>27</sup> The reader should be warned that I have not given a full presentation of my exegesis of Frege's view, in particular, of the fact that the verbal expression does not express a complete thought. For a more complete account, see my discussion in chapter 9. An important conclusion from the Fregean point of view not mentioned in the present discussion is that 'I', 'now', 'today', and the like have to be viewed as unsaturated expressions according to Frege, strictly speaking standing for functions. This has been omitted, and perhaps even distorted, in the present version, where one would think, as Kaplan says, of 'I' as a singular term denoting its user. The reason is, as I said, that I am only marginally concerned with Frege and Fregean exegesis in this sense in the present paper.

As I mentioned in chapter 9, the later Wittgenstein, and those following him, might object to any idea that one understands 'I' in the mouth of someone else by analogy to one's own case, but I am scouting this issue here (actually, in the earlier paper I had the excuse that Frege was unlikely to have been worried about such an objection; here I don't have it).

<sup>28</sup> Perry (1977:491) in fact objects that Frege would have to be committed by analogy to what he says about 'I' to the view that a thought containing 'now' is inexpressible at any later time. I think that this is indeed the correct consequence of Frege's theory, and say so in chapter 9. However, in my more complete Fregean exegesis, I take the verbally expressed part (in English) to be the same at all times. However, this does not express a complete thought. What does so in the case of 'now' is the verbally expressed part together with the supplementation whose sense is given by autonomous reference and acquaintance with time of utterance. It is this that is unrepeatable, since the acquaintance is preserved at no later time. See chapter 9 for the details.

mean, when he commended Frege for his treatment of ‘yesterday’ and ‘today’, that this treatment is correct on a Fregean approach.<sup>29</sup>

## II

Enough about Frege—what about according to me? Well, according to me, the first person use of ‘I’ of course does *not* have a Fregean sense, at least if this means that it has a definition. But it might be a paradigmatic case, one that I did not mention in *Naming and Necessity* (1980), of fixing a reference by means of a description: it is a rule of the common language that each of us fixes the reference of ‘I’ by the description ‘the subject’. However, since each of us speaks a natural language, and not an imaginary ‘scientific language’ spoken by no one, for each of us the referent can be different. This is the moral that I wish to stress.

A long time ago, in conversation, Harry Frankfurt suggested to me that the Cartesian cogito might be an example of the contingent a priori.<sup>30</sup> At the time I thought that whatever may be said about this case, it has a very different flavor from the examples in *Naming and Necessity*. It is certainly contingent because I (or whichever subject is involved in the relevant cogito) might never have been born, and it is a priori at least in the sense of not requiring any specific experience for its verification. But it now seems to me that it does indeed have some of the flavor of my own examples, and perhaps lacks some of their more problematic features. For it follows from the way I fix the reference, as the subject of my own thought, that I must exist. (I will discuss the famous Humean objection to this conclusion later, but here I am assuming that Descartes is right.) In both the cases of the meter stick (‘stick S’) and Neptune, I must grant that the object might not exist. In the meter stick case, the stick I think I am looking at might be illusory (I was tacitly assuming in *Naming and Necessity* that the reference is being fixed

<sup>29</sup> See chapter 9 (page 284, note 81). Kaplan himself mentions one of the objections to the ‘yesterday’ and ‘today’ case from a Fregean point of view, and as I said, probably thinks of them as directly referential demonstratives, yielding a single ‘content’ (in Kaplan’s terminology). See also my own distinction between fixing a reference and giving a meaning, as spelled out immediately below. In fairness to Kaplan, I should add that his original theory in Kaplan (1989) was not simply that ‘I’ is a term that, when used by any speaker, directly refers to that speaker, but also that it is directly referential.

I might mention that Buber’s *Ich und Du* (1923) (translated as *I and Thou*, or *I and You*) may be thought of, as among other things, giving an alternative account of the semantics of ‘you’ to the one I have, in my discussion above, claimed that Frege must give. I am familiar with this work only in part.

<sup>30</sup> I don’t remember when I had this conversation with Frankfurt. I am now uncertain about the history of my own thoughts on this matter. Such examples of the contingent a priori have been widely accepted even by those who doubt my own examples of Neptune and the meter stick. Kaplan’s example ‘I am here now’, with ‘I exist’ as an obvious corollary, is well known as an example of the contingent a priori (see Kaplan 1989, pp. 508-509). Even ‘I exist’ (or strictly speaking, its negation) is explicitly mentioned by Kaplan on p. 495. Plantinga also suggests that ‘I exist’ is contingent a priori (see Plantinga 1974, p. 8).

by someone who has the stick in front of her), and in the Neptune case the astronomical deduction might have been wrong, with no such planet existing, as turned out to be the case with Vulcan. Thus, if I wish to express a priori truths, I must say 'if there is a stick before me as I see it, then . . . ' (In the Neptune case I must say 'if some planet causes the perturbations in Uranus in the appropriate way, then . . .').<sup>31</sup> The whole point of the cogito is that no such existence problem arises, epistemically speaking. Yet another difference with the meter stick and Neptune cases is this: in both cases there is a closely related statement that is necessary and trivial given the way the reference is fixed, such as 'the planet, if any, that causes this perturbations, does cause them', and 'stick S, if there is such a stick, has as its length the length of stick S'. Thus someone might argue (but see my accompanying note here) that these examples of the contingent a priori are really cases where one has no information beyond that provided by the related and trivially analytic necessary truth.<sup>32</sup> The cogito does not seem to be involved in this problem.

I remember when I was very young, about twelve or thirteen, reading Descartes and finding the cogito very convincing. Some time later, reading Hume, I found this:

There are some philosophers, who imagine we are every moment intimately conscious of what we call our SELF; that we feel its existence and its continuance in existence; and are certain, beyond the evidence of a demonstration, both of its perfect identity and simplicity. The strongest sensation, the most violent passion, say they, instead of distracting us from this view, only fix it the more intensely, and make us consider their influence on *self* either by their pain or pleasure. To attempt a farther proof of this were to weaken its evidence; since no proof can be deriv'd from any fact, of which we are so intimately conscious; nor is there any thing, of which we can be certain, if we doubt of this.

Unluckily all these positive assertions are contrary to that very experience, which is pleaded for them, nor have we any idea of *self*; after the manner it is here explain'd. . . .

For my part, when I enter most intimately into what I call *myself*, I always stumble on some particular impression or other, of heat or cold, light or shade, love or hatred, pain or pleasure. I never can catch *myself* at any time without a perception. . . . If any one, upon serious and unprejudic'd reflexion, thinks he has a different notion of *himself*; I must confess I can reason no longer with him.<sup>33</sup> All I can allow him is, that he may be in the right as well as I, and that we are essentially different in this particular. He may, perhaps,

<sup>31</sup> I discuss these cases from *Naming and Necessity* in an unpublished manuscript 'Rigid Designation and the Contingent *A Priori*: The Meter Stick Revisited' (Kripke 1986). For the examples in question see Kripke (1980:54–57; 79, note 33; 96, note 42).

<sup>32</sup> See my note on this matter in Kripke (1980:63, note 26). I now have more to say about the issue and think that some such stipulations may significantly affect the way one thinks about the world, in particular, in the case of the meter. I have discussed these issues in two unpublished manuscripts, the one mentioned in the previous note, and 'Logicism, Wittgenstein, and *De Re* Beliefs about Natural Numbers' (Kripke 1992).

<sup>33</sup> That's just how I feel about the robot philosophers. See note 21.

perceive something simple and continu'd, which he calls *himself*; tho' I am certain there is no such principle in me. (2000:164-65 [Book I, Part IV, Section VI])

The concluding sentences are, of course, a sarcasm. Now, after I had read this passage with shock, I thought that philosophy was a very confusing subject. One philosopher is very convincing, and then another one comes along and gives a very decisive refutation. Who knows what will happen if I read a third one?

Some years later, when I was in my twenties, I talked about it to a non-philosopher, a friend of mine at the time (my attempt when I first read it to talk about it to a classmate was completely unsuccessful). She said, 'Well, Hume must never have looked in a mirror'. At that time I probably thought that her remark simply showed how uncomprehending non-philosophers could be. For of course Hume was objecting to the notion of a Cartesian ego, a pure mind that is the subject of thoughts and impressions. But at the present time I see some justice in her remark, since the idea of a subject, one that I am aware of through self-awareness, as described above, need not in itself imply that the reference is sometimes other than a person, the same person one is aware of, with or without the help of a mirror, or even a sense of sight.<sup>34</sup>

Moreover, the last sentences of the quoted paragraph, witty and effective as they may be, are strangely near to contradictory: 'If any one, upon serious and unprejudic'd reflexion, thinks he has a different notion of *himself*, I must confess I can reason no longer with him. All I can allow him is, that he may well be in the right as well as I, and we are essentially different in this particular. He may, perhaps, perceive something simple and continu'd, which he calls *himself*; tho' I am certain there is no such principle in me'. *Me? He? P?* The beginning of the paragraph is similarly confusing. It appears to presuppose entirely the very notions it attempts to deny.

As is well known, Hume regarded the self as a notion constructed by relating various impressions through resemblance, contiguity, or causation.<sup>35</sup> All we really have is a bundle of perceptions, unified by these relations. Many problems beset this idea. Why should my own impression not equally resemble that of someone else, or be equally contiguous with that of someone else?<sup>36</sup> And similarly, couldn't an impression of mine have a causal relation to that of

<sup>34</sup> Of course, recognizing yourself in a mirror presupposes some concept of self-awareness already. Also, I have been warned that my reference to a mirror in the conversation I reported might suggest something about the use by psychologists of mirror self-recognition tests to see whether a subject has a concept of her/himself. Nothing like that was in my interlocutor's mind (nor mine), and the example could have been given without a mirror (though Hume would have seen less of himself). I believe the test had not yet been developed when the conversation took place.

<sup>35</sup> Since he is in this section relating the notion of the self to that of a pure Cartesian ego, completely independent of a body, the question of bodily identity is not discussed. But what Hume has to say about physical bodies elsewhere shows that he thought that, for similar reasons, they would be of no help.

<sup>36</sup> By 'contiguity' Hume is referring to the succession of impressions in time, rapidly after each other.

someone else? In fact, all these things do happen.<sup>37</sup> It is not fair to say that only the impressions that *I* am aware of count. And, as I have already emphasized, Hume says that we confusedly form the notion of a single, persisting entity because of the close relations between the various impressions. But who is this 'we' who do this?<sup>38</sup> (The discussion here and following might be compared with my discussion of the issues in the last chapter of my book on Wittgenstein, Kripke 1982).

A more basic problem for Hume is that he seems to think that there could be impressions, mental acts, and so on with no bearer.<sup>39</sup> They can simply 'float'—that is, each impression could simply exist in and of itself, and it is only an accident that they are connected by the relations that he mentions. I myself find the notion of an impression or idea without any subject who has it hard to understand. I must admit that even more recently there are those who seem to agree with the Humean picture. In Kripke (1982:123), I quote Moore's account of Wittgenstein's Cambridge lectures in 1930–33:

[Wittgenstein said that] 'a [physical] eye doesn't enter into the description of what is seen . . .';<sup>40</sup> and he said that similarly 'the idea of a person' doesn't enter into the description of 'having [a] toothache'. . . . And he quoted, with apparent approval, Lichtenberg's saying 'Instead of "I think" we ought to say "It thinks"' ('it' being used, as he said, as 'Es' is used in 'Es blitzt').<sup>41</sup>

As I've indicated, there are others, even more recently, who appear to agree. For example, Peter Geach argues that if Descartes is merely solipsistically thinking to himself, instead of saying 'I am getting into a muddle', he might as well simply

<sup>37</sup> Given Hume's well-known skepticism concerning the notion of causation, it is interesting how he invokes it here and elsewhere. But this is probably no inconsistency and is rather a statement of how one dubious notion depends on another.

<sup>38</sup> I find that Chisholm (1976:39–41) has a discussion of Hume closely related to this one. He himself mentions some similar comments by Price.

<sup>39</sup> I emphasize this problem in Kripke (1982:130–31, note 12).

<sup>40</sup> Probably Moore's quotation is somewhat inaccurate here. Of course an eye is involved in seeing, but what Wittgenstein points out is that it is not part of the visual field, that we do not see the eye. See *Tractatus* 5.633, which also states the analogy to the nonexistence of the subject. See also *Tractatus* 5.631: 'There is no such thing as the subject that thinks or entertains ideas.' He adds that no such subject could be found if I wrote a report on 'the world as I found it'. The remainder of the paragraph elaborates on the point. Moore's quotation shows that Wittgenstein still holds this view in his lectures in the 30s.

Apropos of my friend's comment on Hume: if Wittgenstein looked in a mirror, wouldn't he see his eye? Or should one argue that all he sees is a mirror image? If so, the same would apply to Hume in a mirror, even if the self in question were not some pure Cartesian ego, but rather a person with a body. However, Hume can look at some of himself without a mirror. One thing he cannot see without one is his own eye, so perhaps the issue of the eye is directly related.

<sup>41</sup> Moore (1954/1955). The quotation is on page 309 in the (1959) reprint. As to Lichtenberg, 'Es regnet' in German would mean 'It is raining' in English. 'Es blitzt' correspondingly would mean 'It is lightning,' used in an analogous sense to 'It is raining'. But for lightning there is no analogous usage in English. The point is that the 'It' (or 'Es') in the subject does not refer to any entity.



have said 'that is a muddle'.<sup>42</sup> But what would a muddle be with no one to be in the muddle?<sup>43</sup>

In spite of Wittgenstein (in the periods in question, anyway), Lichtenberg, and so on, Hume's view, for the reasons that I have given, seems to me to be quite unintelligible. He must have gone wrong somewhere. (Though Descartes might have gone wrong somewhere, too.) I mean, what would a floating impression not belonging to anyone be? Yet, according to Hume, the supposed self is simply constructed from a bundle of such impressions that in principle could each 'float' independently of any other impressions, let alone a bearer.

So, basically, I think my friend had a point. So far there is nothing in the notion of a subject, as I have defended it above and as Frege presupposed it, that in and of itself excludes the first person pronoun as referring to the whole person in the ordinary sense.<sup>44</sup>

Thomas Nagel once delivered a paper, unpublished as far as I know,<sup>45</sup> in which he proposed that a person be identified with her brain. Rogers Albritton was the commentator, and remarked that if that were so, he shouldn't have been so worried about his weight—it is much less than he had thought.<sup>46</sup>

What did Nagel have in mind when he proposed to identify each of us with her or his brain? What he meant was that if I lose my arm, this is very unfortunate, but I have remained. As a matter of practical medicine, not too much could be stripped away if I am to remain alive, but assuming medicine to have conquered the problem, I could survive the loss of a great deal more. But my brain, from a point of view that is not immaterialist in this respect, is essential as the seat of my thought. As long as it is still functioning, I have not disappeared; but if it goes, I have gone. While this is so, Albritton's comment means that it is still true that my arm is a part of me, as long as it is there. And if I lose it, I have lost one of my parts.

<sup>42</sup> Geach (1957b:117–21, §26 ["The Fallacy of "Cogito Ergo Sum""]).

<sup>43</sup> In fairness to Geach, what he is arguing is that the cogito does not directly give us the idea of an immaterial subject. Perhaps not, but not because there might be no subject at all. There is also, in the passage in question, some discussion of how 'I' is used in ordinary communication, supposedly distinct from the Cartesian case.

<sup>44</sup> I don't know why John Perry, in talking about this, wanted to emphasize someone who erroneously thought he was Hume, rather than a more standard example like Napoleon or Christ. A while ago, I was having dinner, and someone who wanted to criticize philosophers said: 'Philosophers think one is not certain of anything, that you are not certain whether you are Napoleon or not'. I replied: 'Napoleon must have been the greatest philosopher of all time, because only *he* was right when he thought he was Napoleon'.

<sup>45</sup> Someone looked it up and reported that it was not in his bibliography, so maybe it hasn't seen the light of day, though I'm pretty sure it was read before an audience.

<sup>46</sup> I was not present at the exchange between Nagel and Albritton, but when I delivered this talk someone who heard the exchange told me afterward that he remembered it vividly.

Nagel has, of course, written important material on the nature of the self, which I do not discuss here.

Descartes held that my essence is thinking. The only thing that is really indubitable, and therefore constitutes me, is the thinker—what I am aware of whenever I think and feel. It is easy to conclude that he holds that I really am simply the Cartesian ego that Hume (and others following him) have found hard to comprehend, or to locate in their own self-consciousness. And perhaps he is usually read in this way. But, of course, there is a well-known passage that appears to be to the contrary. Once Descartes has proved to his own satisfaction (what initially he doubted) that his own body is real, he states:

Now there is no more explicit lesson of nature than that I have a body; that it is being injured when I feel pain; that it needs food, or drink, when I suffer from hunger, or thirst, and so on. So I must not doubt that there is some truth in this. Nature also teaches by these sensations of pain, hunger, thirst, etc., that I am not present in my body merely as a pilot is present in a ship; I am most tightly bound to it, and as it were mixed up with it, so that I and it form a unit. (1971:117, translated by Anscombe and Geach)<sup>47</sup>

Descartes goes on to argue in the same vein that the way I feel sensation shows that I am a unity, including my body. The case involved in the Nagel–Albritton exchange is similar, though more materialistically expressed. Am I my brain alone, directing my body like a pilot in a vessel? Or am I a unity, including all of my physical body, even though the brain could be the only part that remains, and as long as it is functioning, I have not disappeared? For Descartes, although I may still exist if my body were stripped away and I were still thinking, as long as that has not occurred, I am a unity including my whole body.

It would be very far from my competence or intent to be giving a historical talk on Cartesian exegesis, though I am drawing attention to a suggestive and well-known passage from Descartes. But the philosophical point is that though I might believe that I can doubt the existence of my left hand (and it might be only a contingent fact that I have it), that does not mean that the left hand is not, in fact, a part of me. The commonsense view (and there is no reason to think that Frege, for one, is opposed to it) is that when Lauben says ‘I’, he means to refer to the person himself, including all bodily parts as genuine parts of himself. Yet he could also think that such parts are subject to Cartesian doubt, that they are only perceived by him as subject, and so on. One could certainly agree with Frege that each of us is aware of himself in a special way. And we have seen that even Descartes ultimately concludes that he is an entity including his own body, even if he believes that he might survive without a body, as long as there is a subject for the cogito.

One notable instance of those who refuse to identify Descartes with the reference of ‘I’ is Elizabeth Anscombe. In the following passage I am her direct target:

<sup>47</sup> Though Cottingham may have become the standard translation (Descartes 1996), the Anscombe and Geach translation was better suited to my purposes—namely, discussing Anscombe’s interpretation of Descartes (see below).

Saul Kripke has tried to reinstate Descartes' argument for his dualism. But he neglects its essentially first-person character, making it an argument about the nonidentity of *Descartes* with his own body. Whatever else is said, it seems clear that the argument in Descartes depends on results of applying the method of doubt. (Anscombe 1975:45; emphasis in text)

I certainly don't think Descartes is identical to his body. This point in itself I don't regard as particularly deep, or even incompatible with views that are, broadly speaking, materialist. For Descartes, I say, was not his body when the body was a corpse. 'Descartes had a serious accident, did he survive?' 'Yes, of course—take a look in this coffin.' The response is absurd; rather, we have to say, 'I am afraid Descartes is no longer with us'.<sup>48</sup> But in and of itself that simply might mean that mere nonidentity is not so important. Perhaps a person is nothing 'over and above' her body, even if they are, strictly speaking, not identical.<sup>49</sup>

However, this hardly touches the main points that Anscombe wishes to make, either substantively or as criticism of my own version of the Cartesian argument in *Naming and Necessity*, that it neglected the first person character of what Descartes says. She mentions Castañeda's well-known discussion (Castañeda 1967; see also his 1966 and 1968) as noting the existence of the indirect reflexive in English, which is grammatically a special form in Greek (the form is exemplified by 'Betty believes that she herself . . .'; see my note 12 above). However, she regards Castañeda's discussion as excessively complicated, a point that Castañeda himself appears to concede. Castañeda does not, as far as I know, mention the technical grammatical term from Greek, but the very simple and clear presentation in Geach (1957b), cited in Lewis (1983:139), does so (see my note 12 again). One would think she would have been aware of this particular author.

Although Castañeda's papers are indeed complicated, Anscombe's is itself none too easy. (I wish I had had the opportunity to talk with her about the paper.) She makes various claims. One is that 'I am not Descartes' follows from the logic of Descartes's argument. She also even argues, from a historical perspective, that Descartes would have accepted this conclusion (see 1975: 55–56). In so arguing, she holds that Descartes uses 'I' to designate the pure Cartesian ego whose nature Hume and others claim not to understand. The idea is supposed to be that, by his very nature, Descartes is a human being, a member

<sup>48</sup> But don't we say, when asked 'Who is that?' (pointing to a tomb), 'That's Napoleon'? Isn't it even better if the corpse is embalmed? For example, 'That's Lenin'. But we could say the same kind of thing at Madame Tussauds, pointing to a wax figure. I have something to say about related examples (see my discussion of the toy duck fallacy in chapter 11, this volume), but no more detail need be given here.

I am aware that some people will reject the simple argument for nonidentity on various grounds. I don't wish to deal with them, but feel obligated to mention such things when I am writing about Anscombe, since, after all, there is Peter Geach and the notion of relative identity.

<sup>49</sup> See my discussion of this point in Kripke (1980:145, note 74).

of an animal species, with a body, and so on, whereas ‘I’ as used in the *Meditations* does not designate such an entity.<sup>50</sup>

Now, we have seen that Descartes’s ultimate conclusion—‘I’ and my body form a unit—is not so simple. There is nothing here to distinguish ‘I’ from ‘Descartes’ in the sense Anscombe is talking about. (Of the several translations of Descartes, I have used the one she is most likely to endorse; but everyone agrees that Descartes makes this point.) However, perhaps Descartes believes in an incorporeal entity whose essence is thinking and is the most indubitable part of this unity, one that would remain if everything else were stripped away. But even granting that this is so, I do not see that he would say that such an entity would not be Descartes—nothing I am aware of supports such a dramatic and paradoxical conclusion.<sup>51</sup> Had he accepted it, I would have expected some explicit assertion of such a claim.

Anscombe believes, however, that ‘if “I” is a “referring expression”, then Descartes was right about what the referent was’ (59). And this is her Descartes, where the relevant use of ‘I’ refers to a Cartesian ego that must be distinct from the man, *Descartes*. She performs certain thought experiments, involving a person being anesthetized, to support her conclusion. Although I am not exactly certain what the argument from these thought experiments is, the conclusion is at least clear: ‘“I” is neither a name nor another kind of expression whose logical role is to make a reference, *at all*’ (60).<sup>52</sup> She also states, ‘“I am E. A.” is not an identity proposition’ (63).

Very likely Anscombe is writing under the influence of passages in the *Investigations*,<sup>53</sup> where Wittgenstein says that ‘I’ is not the name of a person, and distinguishes between ‘I’ and ‘L. W.’ The influence goes down to the use of initials to make the contrast—‘E. A.’ in the one case and ‘L. W.’ in the other.<sup>54</sup>

<sup>50</sup> Very likely she is also influenced by the view, advocated by both Peter Geach and Noam Chomsky, that a proper name such as ‘Descartes’ has a semantical requirement that it name a human being. This might or might not already imply the possession of a body, depending on other views. For my own discussion of Geach’s view (in Geach 1957a: §16), see Kripke (1980:115, note 58).

<sup>51</sup> Part of Anscombe’s argument, also to be found in other authors (see below), is that the indirect reflexive allows someone to be mistaken about who he is, so the conclusion is not self-contradictory. But Descartes is not in this sense mistaken about who he is.

<sup>52</sup> Her discussion has an elaborate contrast of the use in an imaginary society of some name that everyone uses for him/herself (‘A’-users) with our use of the first person pronoun, which is not entirely clear to me. She also remarks that calling it a personal pronoun, and the like, is simply a trivial restatement of first person usage and, in and of itself, gives no information.

However, her general conclusion is clearly stated in what I have just quoted.

I confess that in both the sentence I am footnoting and the present note, I would be willing to replace ‘I’ by ‘Saul Kripke’—only the resulting awkwardness and pomposity would stand in the way. And I have noted the usage of ‘the present writer’ above (see note 7). Surely ‘the present writer’ does make a reference, or at any rate, is a definite description.

<sup>53</sup> Wittgenstein (1953: §§405–6).

<sup>54</sup> Geach (1957a: §26), in the anti-Cartesian passage referred to above, similarly contrasts the Cartesian ‘I’ with ‘P. T. G.’, probably under the same influence. At least, so I conjecture. In my own discussion, I probably would not be inclined to use initials. See note 52.

No doubt when Wittgenstein makes the relevant distinction in the *Investigations*, his thought is continuous with his earlier worries about the metaphysical subject, as quoted above (see note 40 and related discussion). The *Investigations* passages are about many things, most of which can't be discussed here, such as the difference between first and third person attributions of sensations, the alleged connection of first person 'avowals' of sensations with more primitive expressions such as groaning or crying, and so on. Obviously, we can't talk about all these things here (they are to some extent in Anscombe's paper too).<sup>55</sup>

I won't attempt to give an exegesis of Wittgenstein on these topics. But one should note the following. Remember that in §410 he says: ' "I" is not the name of a person, nor "here" of a place, and "this" is not a name' (see note 4). In one sense this is obvious, because 'I' is not a name at all, but maybe something deeper is meant. See §405, where he says:

'But at any rate when you say "I am in pain", you want to draw the attention of others to a particular person.'—The answer might be: No, I want to draw their attention to *myself*.— (§405; emphasis in original)

It is unfortunate that in §405 and the following paragraphs there is a concentration on the case of 'I am in pain'. Perhaps one might wish to be thinking of a 'Cartesian' case of describing one's own inner states. Obviously, however, this is not the general case of first person usage (see what I say in note 55).

A common criticism of Anscombe's position is that if 'I' is not a referring expression, why should we be confident in the inference pattern from, say, 'I live in North Carolina' to 'Someone lives in North Carolina' (in other words, existential generalization), or in the fact that any inference pattern where 'I' is treated as if it refers is valid. Indeed, such an objection is only a technical expression of one's natural reaction that Anscombe's thesis is in and of itself incredible, difficult to understand at all.<sup>56</sup>

As I said, I was never able to talk to Anscombe about these matters, but I do recall a report from someone else as to what she said when queried as to why 'I' behaves as if it refers in inference patterns. Her answer as reported was 'I don't know'.

<sup>55</sup> See Wittgenstein (1953), §398, *passim*. In §406 there is Wittgenstein's well-known derivation of 'avowals' from more primitive expressions. He says: ' "But surely what you want to do with the words "I am . . ." is to distinguish between *yourself* and *other* people.'—Can this be said in every case? Even when I merely groan?' Of course, the one who groans does not use the word 'I' or any equivalent at all; and even if one somehow regards 'I am in pain' as simply derived from a groan, this is hardly an obvious illumination of the general case of all first person usage. But I don't say that Wittgenstein thinks that it is.

<sup>56</sup> See, for example, Peacocke (2008:80). He calls Anscombe's conclusion 'barely credible', and makes the criticism about inference patterns. He also uses the argument about inference patterns to show that 'I' cannot be compared with 'it' in 'it is raining', where the grammatical pattern superficially suggests a subject of reference but there is none. (There are plenty of such cases; 'nobody' is a famous one.) Peacocke's comments are a common reaction to Anscombe's view.

One might here remark that Anscombe is a very special case, coming from a special background, that of Wittgenstein. And though, clearly, some of the argumentation is her own, as is the formulation of her thesis, and though Castañeda's well-known discussions of the matter are plainly an influence, the predominant influence may be the special background.

### III

Let me turn to a philosopher with quite a different background and orientation. David Lewis was worried about the problem of belief *de se* (Lewis 1979). According to Lewis, there is general agreement that the objects of all the so-called propositional attitudes are propositions, though he acknowledges that not all authors agree as to what propositions are. For him, propositions are sets of possible worlds. But though sometimes I have heard Lewis say that here he simply is giving a stipulative definition—and remember, in addition, that Lewis has his own very special sense of the term 'possible worlds'—presumably, in one sense, it is supposed to represent a thesis or discovery having something to do with traditional uses of the term 'proposition'. I myself am one of the originators of this proposal (in connection with modal logic, anyway), though I wasn't thinking of possible worlds in Lewis's sense, and now might modify it. Today, I even feel some reservations about the idea that *all* the phrases traditionally called propositional attitude constructions have the same sort of entities as objects, though this issue is not for the present paper.<sup>57</sup> Lewis also correctly remarks that the general agreement that the objects are propositions is "to some extent phony" (1983:134), since some authors take propositions to be more highly structured entities, more like an abstract, nonlinguistic, analogue of sentences.

In any event, Lewis proposes a theory that was formulated independently by Roderick Chisholm (1981),<sup>58</sup> and actually was anticipated by Quine (1968).<sup>59</sup>

<sup>57</sup> In the first section of Lewis (1979), he gives some reasons for hoping for uniform objects. As far as the arguments of this section go, I agree.

<sup>58</sup> I am familiar with Chisholm's version only superficially. Like Lewis, he proposes that self-ascription be taken as primitive. However, he doesn't put matters in terms of Lewis's special apparatus of possible worlds. In fact, at the end of his book, he states his own view on the nature of possible worlds, which he thinks is the traditional one.

<sup>59</sup> My thanks to Gilbert Harman for emphasizing that Quine's paper anticipates Lewis's theory. However, as Lewis says, Quine uses a different notion of possible world. Other than that, the theories, though in my own opinion somewhat differently motivated, can be interpreted as the same. Perhaps I should add that Quine's paper somewhat surprised the present writer, given what he had said about possible worlds elsewhere. There is even a paper entitled 'Worlds Away' (Quine 1976) expressing his repudiation of the notion, and not mentioning that he had written his own version of a kind of possible world semantics earlier. One might also add that the original journal version in Quine (1976) and the one printed in Quine (1981) are significantly different, representing an unacknowledged change of argument.

Also, Lewis acknowledges Brian Loar (1976) as making the same proposal (Lewis 1983:519, note 4).<sup>60</sup> He also makes a comparison of his views with a proposal by John Perry (1977) (Lewis 1983:150),<sup>61</sup> and plainly Lewis is considerably influenced by Perry's paper, as Lewis himself says (139).

Neither Lewis nor Chisholm actually gives an analysis of statements with 'I'. Rather, as Chisholm says explicitly (1979), what is explained is the indirect reflexive. For example, one takes as primitive 'A self-ascribes such and such a property', where self-ascription is a primitive notion, corresponding to 'A believes that he himself has such and such a property'. Since self-ascription, like any reflexive property  $\lambda x (xRx)$ , is derived from a relation  $R$ , Lewis rightly concludes his paper (sections XIII and XIV) with a brief discussion of belief *de re*. He regards beliefs *de re* as beliefs based on acquaintance, and adds that belief *de se* is a par excellence case of belief *de re*, based on the highest form of acquaintance, namely, one's acquaintance with oneself.<sup>62</sup>

I should mention that Lewis also wishes to account for 'now' and the present tense, and for this reason regards it as a person stage that has an attitude *de se*. But this is a separable part of the view, drawn from the rest of Lewis's philosophy. As far as the present issue is concerned, he could just as well say that a person (not a stage, but an enduring object) self-ascribes not a property (corresponding to a set of worlds), but a two-place relation between a person (himself) and a time (the time of the ascription, corresponding to the same set).

Note also that Lewis, in contrast with some of those to whom he compares himself, regards all attitudes as *de se*, even those that seem to make no reference to the subject (say, 'lightning is an electrical discharge', or 'Australia is a large island', etc.). For this Lewis cites the advantages of uniformity, so that there

Lewis himself (1983:147–48) discusses some of the differences in motivation that he sees between himself and Quine. He remarks (147) that Quine differs from him in considering a divided theory: the objects of some primitive attitudes may be stimulation patterns rather than sets of worlds, making his theory highly nonuniform, something Lewis wishes to avoid. Moreover, Quine is concerned with attitudes and desires of animals too, an issue that Lewis does not consider (nor shall I). Lewis mentions other differences in motivation, and Quine does not share Lewis's modal realism (as Lewis calls it).

I myself would treat possible worlds differently, even given Quine's project. In particular, I do not agree that either classical or relativistic physics has shown the notion of a (physical) geometrical point to be absurd, as Quine thinks, even though he informally uses this notion to motivate his definitions. The important thing for the present purposes is that to explain *de se* attitudes, Quine introduces the notion of 'centered possible worlds', where the centering is on the subject (or its physical location). In and of itself a possible world is not centered, so that a centered possible world might be thought of as an ordered pair of a world and a center.

<sup>60</sup> Lewis states that Loar formulates his view for 'certain exceptional beliefs'.

<sup>61</sup> I have discussed Perry's paper above and in chapter 9.

<sup>62</sup> Elsewhere, Chisholm advocates the theory, rightly (in my opinion) rejected by Lewis in his paper, that one has a belief *de re* about an object as long as one has any description designating the object under which one has the appropriate belief (what I have called 'universal exportation', chapter 11, this volume). So he cannot regard belief *de se* as a special case of belief *de re*. I have not researched whether Chisholm says anything about the issue.

are not two kinds of objects of beliefs, and so on, sometimes propositions, and sometimes properties, but only one. But really, other reasons can be given in Lewis's own framework. First, in the old theory, the object of a propositional attitude is the same for logically equivalent sentences, so that, for any  $p$ , ' $p$ ' and ' $p$  and David Lewis is self-identical' express the same attitudinal object (set of possible worlds), even where  $p$  makes no reference to David Lewis. It would be natural to preserve this feature with 'I' in place of 'David Lewis'.<sup>63</sup> Second, and most important, for Lewis the actual world is distinguished as the one the subject inhabits, and hence any ordinary belief is really a belief that I inhabit a world with certain properties; the uniformity is not really artificial at all, given Lewis's view of the nature of possible worlds as vast concrete worlds and his theory of actuality. For example, my belief that *actually* Australia has kangaroos is a belief that I inhabit a world where Australia has kangaroos.

All this, however, seems to me to be an odd reversal of matters.<sup>64</sup> As I say in *Naming and Necessity*, 'a possible world is *given by the descriptive conditions we associate with it*' (1980:44; emphasis in original). By this I did not mean to identify possible worlds with sets of descriptive conditions, for example, as Carnapian state descriptions (probably impossible in a countable language anyway). But I wish to emphasize the legitimacy of setting up possible worlds by any description we understand, and in particular that it need not be purely qualitative. For example, I emphasize, as against those who worry about a problem of 'transworld identification', that a counterfactual possibility could be stipulated as being about Nixon, using the name 'Nixon', and not worrying about any reduction to a qualitative description.

Now, what language does a person use when describing counterfactual possibilities? Not a 'scientific', indexical-free language—one that none of us speaks. Each of us has a notion of the self and often a word for it (in English, 'I'). Why can we not use such a language to describe a counterfactual situation, and hence a possible world (even if we normally do so only in part)? I use the word 'I' to designate myself, and to designate myself in a particular way, as has been

<sup>63</sup> Pardon the abuse of quotation here. I suppose I should have used Quine's corners. Everyone knows what I mean, and I wish to be excused.

<sup>64</sup> Evans (1977) suggests that perhaps the indirect reflexive could be thought of as deriving from the *oratio recta* construction with the first person pronoun (98). For example, he sees 'John thinks that he is under suspicion' as 'somehow derived from' 'John thinks "I am under suspicion"' in *oratio recta*, and refers to Anscombe's paper. I am sympathetic with something Evans is trying to say, namely, that the form with the first person pronoun is primary. However, one cannot make this point by deriving the indirect reflexive from an *oratio recta* construction. Nothing in the indirect reflexive form implies that John is speaking or thinking in English. The trouble is that if we wish to ascribe to John an 'I'-thought using a 'that-clause', we use the indirect reflexive itself, and thus get into a tangle expressing the derivation, which I do think goes in the direction indicated by Evans.

I confess that I have not taken Evans's other writings on *de se* attitudes into account.



discussed above. Quine speaks of ‘centered possible worlds’.<sup>65</sup> I would rather speak of possible worlds *simpliciter*; the ‘centering’ comes when a particular person describes a counterfactual situation.

Lewis himself rightly (as I have said) describes *de se* belief as a special case of *de re* belief, and bases the latter on a notion of acquaintance. Self-acquaintance is such acquaintance in the highest degree. Why then can someone not use such acquaintance to formulate a word ‘I’, and use it to designate an object in a special way? It will follow from this means of designation that it is rigid, or more strongly, as Kaplan puts it, ‘directly referential’ (if we are putting the matter that way, ‘directly referential in a special way’).

One should not think that any situation, actual or counterfactual, really should be described in a ‘scientific’ language (see note 4), one free of person and tense, and so on, and then tack on an identification of the people and times in it, as ‘I’, ‘now’, and so on. Lewis makes much of a situation with two gods (1983:139), one of whom does one thing and one of whom does another. The gods, he says, could be omniscient, as far as propositional knowledge is concerned, without either one knowing which one of the gods he is. Robert Stalnaker has observed that it is difficult to imagine such a situation as intelligible. How can someone be doing something without realizing that it is he himself that is doing it? (Stalnaker 2008:56ff).<sup>66</sup>

My own view is that to describe a possible world is to give a ‘possible history of the world’ (Kripke 1980:48, note 15). Such a history, or a portion of it, might well be describable in an indexical free (or ‘scientific’) language, but nothing says that it need be so described. In particular, when I am speaking the language, I am entitled to use the first person singular in describing such a history. Thus, both in an informal, intuitive sense, as well as in the technical sense of determining a set of possible worlds, statements about myself do express propositions, only they determine them in a special way.

Let me speak of possibilities in the way that I spoke of ‘metaphysical possibility’ in *Naming and Necessity*. I can wonder what will happen to me, and how things might have come out otherwise, even in the past, had only I done such-and-such. Here the picture should not be as if I might be thinking about possible

<sup>65</sup> See note 59 above. Stalnaker (2008:49) ascribes the notion of ‘centered possible worlds’ to David Lewis. In fact, the terminology, and the associated picture, is due to Quine, even though Lewis’s view can be considered equivalent.

<sup>66</sup> Can two people get into an argument without each participant knowing which side she is taking? I won’t discuss Stalnaker’s analysis, and some of his other objections to Lewis’s version, further. It appeared subsequent to the original version of this paper. Nor do I discuss some possibilities for reinstating Lewis’s example.

worlds and then ‘tacking on’ which person is me. The situation is the opposite; I am determining possibilities by reference to myself.<sup>67, 68</sup>

One might also consider Lewis’s earlier paper ‘Anselm and Actuality’ (1970) in relation to this one (i.e., Lewis 1979). In one sense, which both Lewis and I would acknowledge, each world is the unique one that is actual with respect to itself. However, we think of one world as the actual one. For me, that is the unique world  $w$  such that a proposition  $p$  is true with respect to  $w$  if and only if  $p$ , for all  $p$ .<sup>69</sup> For Lewis, it is the unique world that I inhabit. Then for Lewis the proponent of the ontological argument has little plausibility in arguing that God must have actuality in this sense in order to have every perfection. Why is this a

<sup>67</sup> Lewis (see his 1983:135), aside from other differences in the conception of possible worlds, is not concerned with metaphysical necessity and possibility in my sense, but rather in connection with propositional attitudes, predominantly doxastic attitudes. He discusses the problem of whether attitudes can be attitudes toward sets of possible worlds, since one can apparently have different attitudes toward propositions that hold in the same sets. He says that ‘believing that  $2+2=4$  is clearly not the same as believing that  $123+456=579$ ’, because the latter calculation is not as trivial as the first. He goes on, ‘I know perfectly well that there is such a thing as ignorance of noncontingent matters’. Wishing to leave this issue aside, he says that, if one wishes, one could imagine that we are talking about ‘the attitudes of imaginary hyper-rational creatures’. So, clearly, examples of metaphysical but not epistemic necessity that I have advocated, such as ‘water is  $H_2O$ ’, are not contemplated. One can suppose that another notion of epistemically possible worlds, or of what I called ‘epistemic counterparts’ of metaphysically possible worlds, is involved. However, from my own perspective, as stated in this very paragraph, real metaphysical possibilities are an important notion when considering ‘I’-sentences, and indeed this is an important reason that they should be considered in and of themselves, not simply in terms of indirect reflexive constructions. And it is what I call metaphysical possibility that is relevant to counterfactuals as Lewis and Stalnaker analyze them (see Lewis 1973).

<sup>68</sup> In fairness to Lewis, he certainly considers that situation (by reference to an amnesiac, Rudolph Lingens, lost in the Stanford Library, an example discussed by Perry). (For Lingens, see note 22. Evidently, Perry has transported Frege’s character from Germany to the Stanford Library. No wonder he has become confused. He is supposed to be reading things in the library, but no matter how much he reads, he won’t know where he is. I hope his English is good. Otherwise, it is not surprising the library is not so helpful.)

But the real points are these. First, of course my conception of possibility is different from Lewis’s. Second, I don’t see why Lingens isn’t expressing definite propositional knowledge using the word ‘*ich*’, or fails to be contemplating genuine possibilities using it. This is true even in the unusual case of amnesia, where although in some sense the subject doesn’t know who he is, there would appear to be another sense in which even he could be said to know this, and in which it would be impossible not to.

Suppose someone wonders what time it is now (a case Lewis considers [1983:143–44]). So, in some sense, he is wondering what time it is, and the answer is given by the clock. Or he may be wondering when it will be noon, and the answer may be ‘now’, or ‘two minutes from now’. ‘When did she die?’ ‘Just now’. Both forms of question are legitimate, and equally so. In the first case, the very same situation is regarded in two ways. In my own opinion, the relativity and indeterminacy of ‘*wh*-questions’ like this is exaggerated in the philosophical literature, but it exists and the present instance is a strong case.

<sup>69</sup> And  $p$  is true with respect to  $w$  could be defined as: if  $w$  were the case, then  $p$  (where in spite of the linguistic form we might as well interpret the conditional as strict). (I won’t go into complications for someone who doesn’t believe in S5 or even S4.)

perfection, inhabiting the same world that I do?<sup>70</sup> Here, of course, much depends on our different conceptions of what a world is. I don't think of myself as inhabiting a world in Lewis's sense, nor do I think much of his answer as a solution to the 'problem' of how we know that we are inhabiting the actual world, rather than a merely possible one. But note that, given Lewis's later view, to say that the world is actual is not to state any proposition about it, it is to attribute a special sort of property to the speaker.

What then are the differences between me and Lewis? First, there is the well-known difference in our conception of possible worlds. For Lewis, if I wonder what possible world I am in, it is as if I am wondering whether I am in Pennsylvania or West Virginia, as if I were traveling by car (though, of course, according to Lewis, at least in his ultimate view, such a trip between worlds is impossible). But second, I do not think that genuine propositions must be described in some neutral scientific language. Anyone can describe them in any language he himself speaks. In some sense, one might call this a merely terminological difference, since (and in spite of the considerable influence on him of Perry's paper) Lewis agrees with me that everyone has a special *de se* acquaintanceship with himself. And, of course, he does not deny that special contents using the first person and involving tense are asserted, entertained, and the like. But he does think that these are not genuine propositions, in some intuitive sense. How can a proposition depend on who expresses it? The room around me (and this would be valid for worlds even if one had Lewis's conception of them) is not dependent on who looks at it, but everyone is permitted to describe it using person and tense.<sup>71, 72</sup>

<sup>70</sup> Doesn't this depend on what I think of myself? The argument could have a plausible premise for Anslem, and not for me.

<sup>71</sup> In the original version of this talk, I was worried that Lewis might face the same difficulty as Anscombe about logical inferences containing the word 'I'. In the sense of classical logical inference, I still think this might be true, since in this classical sense such an inference is valid simply in virtue of its form. (The Quinean notion of a logical truth has a similar motivation.) It is indeed true that an inference involving 'I' in the premise does not, for Lewis, even express a genuine premise (proposition), let alone one involving 'I' as a singular term. But Gilbert Harman and Robert Stalnaker have pointed out to me that Lewis can easily explain in his apparatus why we can validly infer from a self-ascription *p* to a self-ascription entailed by it: it is simply a matter that one who self-ascribes the property of being in a set of possible worlds is thereby committed to self-ascribing to himself the property of being in any larger set of possible worlds. I haven't really thought about whether such problems arise on Chisholm's version. Note that in the crudest sense the Lewis-Chisholm theory could be regarded as one that solves the supposed self-ascription problem simply by taking self-ascription as a special primitive; then, questions as to the logic of this predicate would obviously arise.

<sup>72</sup> From this point of view, 'I am Saul Kripke' is little different in principle from 'Hesperus is Phosphorus', or 'Cicero is Tully', and the like, though 'I' determines its referent in a very special way. Similarly, 'I am Jesus Christ', said falsely, misdescribes an actual or possible situation, as would 'Hesperus is Mars'. A better analogy might be 'that man is Jesus Christ', which could be uttered by a believer, or denied by a disbeliever (and might even be expressed by a proper name for 'that man'), but it is no longer first person. On my view, the statement would also misdescribe every possible world.

Well, what I have been arguing? Not anything really so special, on a topic that has a considerable literature. But each of us does have a special acquaintanceship with himself or herself, as philosophers from Descartes to Frege have held. This self-acquaintance is more fundamental than anything purely linguistic, and is the basis of our use of first person locutions. And each of us can use them to make genuine claims, to express genuine propositions.<sup>73</sup>

## REFERENCES

- Almog, J., J. Perry, and H. Wettstein, eds. (1989). *Themes from Kaplan*. New York: Oxford University Press.
- Anscombe, G. E. M. (1975). 'The First Person'. In *Mind and Language*, ed. S. Guttenplan. Oxford: Clarendon, 45–65.
- Beaney, M. (1997). *The Frege Reader*. Malden, MA: Blackwell.
- Brouwer, L. E. J. (1948). 'Consciousness, Philosophy and Mathematics'. In Brouwer (1975), 480–94.
- . (1975). *L. E. J. Brouwer, Collected Works. Volume 1*. Ed. A. Heyting. Amsterdam: North-Holland.
- Buber, M. (1923). *Ich und Du*. Leipzig, Germany: Insel-Verlag.
- Castañeda, H. N. (1966). "'He": A Study in the Logic of Self-Consciousness'. *Ratio* 8:130–57.
- . (1967). 'The Logic of Self-Knowledge'. *Noûs* 1:9–22.
- . (1968). 'On the Logic of Attributions of Self-Knowledge to Others'. *Journal of Philosophy* 65:439–56.
- Chisholm, R. M. (1976). *Person and Object*. London: George Allen & Unwin.
- . (1979). 'The Indirect Reflexive'. In Diamond and Teichman (1979).
- . (1981). *The First Person: An Essay on Reference and Intentionality*. Minneapolis: University of Minnesota Press.
- Descartes, R. (1971). *Descartes: Philosophical Writings*. 2nd ed.. Ed. and trans. G. E. M. Anscombe and P. T. Geach. Indianapolis: Bobbs-Merrill.
- . (1996). *Meditations on First Philosophy, with Selections from the Objections and Replies*. Ed. and trans. J. Cottingham. Cambridge: Cambridge University Press.
- Diamond, C., and J. Teichman, eds. (1979). *Intention and Intentionality*. Sussex, UK: Harvester.

One could also use tense in describing counterfactual as well as actual situations. However, there could be special problems in using 'now' in the description of a hypothetical total world history. For one thing, the description would necessarily be very fleeting. For another, there could be relativistic histories for which its referent is problematic. Also, one might be accused of smuggling in views about 'A-series/B-series' problems, if these are problems. A tenseless version of such a total world description could be thought of as not taking a stand on these issues.

<sup>73</sup> I would like to thank Romina Padró for transcribing the original lecture and for helpful suggestions and discussions. I also want to thank Jeff Buechner, Gary Ostertag, and Harold Teichman for editorial assistance. This paper has been completed with support from the Saul A. Kripke Center at the City University of New York, Graduate Center.

- Evans, G. (1985). 'Pronouns, Quantifiers, and Relative Clauses (I)'. In Evans (1985), pp. 76–152.
- . (1981). 'Understanding Demonstratives'. In *Meaning and Understanding*, ed. H. Parret and J. Bouveresse. Berlin: W. de Gruyter. Reprinted in Evans (1985), pp. 291–321.
- . (1985). *Collected Papers*. Oxford: Oxford University Press.
- Frege, G. (1892). 'Über Sinn und Bedeutung'. *Zeitschrift für Philosophie und philosophische Kritik*. 100:25–50. Translated by Max Black, in Beaney (1997), pp. 151–71.
- . (1918–19). 'Der Gedanke'. *Beiträge zur Philosophie des deutschen Idealismus* 1:58–77. Translated as 'Thoughts' by Peter Geach and R. H. Stoothoff, in Gottlob Frege, *Logical Investigations* (New Haven, CT: Yale University Press, 1977), and included in Beaney (1997), pp. 325–45, under the title 'Thought'. Also translated by A. M. and M. Quinton, as 'The Thought: A Logical Inquiry,' in *Mind* 65 (1956):289–311.
- Geach, P. (1957a). *Mental Acts: Their Content and Their Objects*. London: Routledge & Kegan Paul.
- . (1957b). 'On Beliefs about Oneself'. *Analysis* 18:23–24.
- Hempel, C. G. (1942). 'The Function of General Laws in History'. *Journal of Philosophy* 39:35–48.
- Hume, D. (2000). *A Treatise of Human Nature*. Eds. David Fate Norton and Mary Norton. Oxford: Oxford University Press.
- Irving, D. J. C. (1977). *Hitler's War*. New York: Viking.
- Kaplan, D. (1968). 'Quantifying In'. *Synthese* 19:178–214.
- . (1989). 'Demonstratives'. In *Themes from Kaplan*, ed. Joseph Almog, John Perry, and Howard Wettstein. New York: Oxford University Press.
- . (n.d.). 'What Is Meaning? Explorations in the Theory of *Meaning as Use*'. Unpublished manuscript.
- Kripke, S. (1961). 'History and Idealism: The Theory of R. G. Collingwood'. Unpublished manuscript.
- . (1977). 'Speaker's Reference and Semantic Reference'. *Midwest Studies in Philosophy* 2:255–76. Reprinted in this volume as Chapter 5.
- . (1979). 'A Puzzle about Belief'. In *Meaning and Use*, ed. A. Margalit. Dordrecht, Netherlands: D. Reidel. Reprinted in this volume as Chapter 6.
- . (1980). *Naming and Necessity*. Cambridge, MA: Harvard University Press. First published in *Semantics of Natural Language*, ed. D. Davidson and G. Harman. Dordrecht, Netherlands: D. Reidel, 1972, 253–355, 763–69.
- . (1982). *Wittgenstein on Rules and Private Language*. Cambridge, MA: Harvard University Press.
- . (1986). 'Rigid Designation and the Contingent *A Priori*: The Meter Stick Revisited.' Unpublished manuscript.
- . (1992). 'Logicism, Wittgenstein, and *De Re* Beliefs about Natural Numbers'. Unpublished manuscript.
- . (2008). 'Frege's Theory of Sense and Reference: Some Exegetical Notes'. *Theoria* 74:181–218. Reprinted in this volume as Chapter 9.
- . (2010). 'Unrestricted Exportation and Some Morals for the Philosophy of Language'. Reprinted in this volume as Chapter 11.

- Lewis, D. (1970). 'Anselm and Actuality'. *Nous* 4:175–88. Reprinted in Lewis (1983).
- . (1973). *Counterfactuals*. Cambridge, MA: Harvard University Press.
- . (1979). 'Attitudes *De Dicto* and *De Se*'. *Philosophical Review* 88:513–43. Reprinted in Lewis (1983); page references are to the reprint.
- . (1983). *Philosophical Papers, Volume I*. Oxford: Oxford University Press.
- Loar, B. (1976). 'The Semantics of Singular Terms'. *Philosophical Studies* 30:353–77.
- Ludlow, P., and S. Neale (1991). 'Indefinite Descriptions: In Defense of Russell', *Linguistics and Philosophy* 14:171–202.
- Ludlow, P., and G. Segal (2004). 'On a Unitary Semantical Analysis for Definite and Indefinite Descriptions'. In *Descriptions and Beyond*, ed. M. Reimer and A. Bezuidenhout. Oxford: Oxford University Press.
- Moore, G. E. (1954/1955). 'Wittgenstein's Lectures in 1930–33'. *Mind* 63 (1954), and 64 (1955). Reprinted in Moore (1959), 252–324.
- . (1959). *Philosophical Papers*. London: George Allen & Unwin.
- Peacocke, C. (2008). *Truly Understood*. Oxford: Oxford University Press.
- Perry, J. (1977). 'Frege on Demonstratives', *Philosophical Review* 86:474–97. Reprinted with a postscript in Perry (1993).
- . (1993). *The Problem of the Essential Indexical and Other Essays*. Oxford: Oxford University Press.
- . (1997). 'Indexicals and Demonstratives'. In *A Companion to the Philosophy of Language*, ed. Bob Hale and Crispin Wright. Oxford: Blackwell.
- Plantinga, A. (1974). *The Nature of Necessity*, Oxford: Oxford University Press.
- Prior, A. N. (1967). 'On Spurious Egocentricity'. *Philosophy* 42:326–35.
- Quine, W. V. O. (1968). 'Propositional Objects'. *Critica: Revista Latinoamericana de Filosofia* 2:3–29. Reprinted in Quine (1969).
- . (1969). *Ontological Relativity and Other Essays*. New York: Columbia University Press.
- . (1976). 'Worlds Away'. *Journal of Philosophy* 73:859–63. Reprinted in Quine (1981).
- . (1981). *Theories and Things*. Cambridge, MA: Harvard University Press.
- Sider, T. (2001). *Four Dimensionalism*. Oxford: Oxford University Press.
- Stalnaker, R. (2008). *Our Knowledge of the Internal World*. Oxford: Oxford University Press.
- Wittgenstein, L. (1953). *Philosophical Investigations*. Ed. and trans. G. E. M. Anscombe. Oxford: Blackwell.
- . (1961). *Tractatus Logico-Philosophicus*. Trans. David Pears and Brian McGuinness. London: Routledge & Kegan Paul.