

Can you think my “I”-thoughts?

Daniel Morgan

Philosophical Quarterly 59 (234) (2009): 68-85.

Introduction

Not everyone agrees that “I” has a sense. “I” has a linguistic meaning all right, one which many philosophers believe to be encapsulated in the simple rule that any token of “I” refers to whoever produces that token. However, the sense of “I” cannot be the same thing as its linguistic meaning. Sense is supposed to determine reference. But the linguistic meaning of a token of “I” does not determine its reference: my token of “I” and your token of “I” have the same linguistic meaning but refer to different things. One response to this observation is to jettison the whole idea that “I” (and other indexical expressions, for all of which linguistic meaning fails to determine reference) has a sense. This is the response of direct reference theorists, such as Perry and Kaplan.¹ The neo-Fregean thinks this response too drastic. To avoid having to give it she attempts to provide an account of what, in general, sense is and what, in particular, the sense of “I” is which sharply distinguishes sense and linguistic meaning. In making this attempt she incurs, I think, at least the following three obligations. To secure the determination of reference by sense, her account must yield the result that my tokens of “I” and your tokens of “I” express different senses, despite being tokens of the same linguistic type. To avoid the collapse of sense into reference her account must yield the result that for at least

¹ See Perry, J. ‘Frege on Demonstratives’. *Philosophical Review* 86, 1977, pp. 474-97; Kaplan, D. ‘Demonstratives’. In J. Almog, J. Perry and H. Wettstein (eds). *Themes from Kaplan* (Oxford: Oxford University Press, 1989), pp. 481-614.

one expression which co-refers with one's tokens of "I"- e.g. one's name- the sense of that expression and the sense of one's tokens of "I" differ. Finally, if we are to have a reason for actually preferring her view, the neo-Fregean must say explicitly what mental or linguistic phenomenon which would otherwise go unexplained gets explained by her claim that "I" has a sense. An obvious candidate would be failures of substitutivity of co-referring terms in attitude contexts. The intuition that I can believe that I am in a strange hospital bed without believing that NN is in a strange hospital bed, despite the fact that I am NN, is as strong as, and strikingly similar to, the intuition that I can believe that Hesperus is a planet without believing that Phosphorous is a planet, despite the fact that Hesperus is Phosphorous. Frege claimed that names had senses as well as references in order to explain the latter possibility. Neo-Fregeans think that the claim that "I" has a sense as well as reference- or, more correctly, the claim that tokens of "I" have senses as well as references- will be required to explain the former possibility.

My sympathies are almost entirely with the neo-Fregean. I shall not though do much in this paper to defend her claim that "I" has a sense – beyond saying in passing how I think each of the three obligations incurred by the neo-Fregean can be discharged. Instead, I focus on a question that arises as soon we accept that "I" has a sense, a question about what the best account of that sense is. Can anyone other than oneself grasp the sense expressed by one's tokens of "I"? Gareth Evans's account of the sense of "I", along with other accounts which give the notion of functional role a central place, entails that the answer to this question is no – that the sense of "I" is not shareable.² I shall label this claim the unshareability claim. Evans is happy to accept the claim- seeing in it a

² See Evans, G. *The Varieties of Reference*. (Oxford: Clarendon Press, 1982)

vindication of Frege's remark that "everyone is presented to himself in a special and primitive way in which he is presented to no one else".³ In a recent paper, Jose Luis Bermudez argues that much in Evans's account is correct and important but that we need to excise just those features of it that yield the unshareability claim.⁴ Bermudez has two main arguments. The first aims to show that the unshareability claim entails the undesirable conclusion that "I"-thoughts are not objective. The second aims to show that the thesis that the sense of "I" is unshareable undermines a plausible account of the relationship between thought and linguistic understanding. I shall rebut each of these arguments in turn, thereby showing that an account of the sense of "I" which has the unshareability claim as a consequence is entirely defensible. Along the way, I will provide a general characterization of what a functional role account of the sense of 'I' is, one which explains both why it is plausible that some such account is correct and also why any such account must yield the unshareability claim. This provides a powerful motivation for the defense of the unshareability claim.

Evans's account of the sense of "I" focuses on the highly distinctive functional role "I"-thoughts have. What is an "I"-thought? Evans would say that it is a thought which is most adequately expressed through a sentence containing the first person.⁵ This definition does seem to provide an intuitive handle on the kind of thought Evans has in mind. We should notice though that a direct reference theorist may well deny that there are any thoughts

³See Frege. 'The Thought: A Logical Inquiry' p. 25 in P.F. Strawson (ed.) *Philosophical Logic* (Oxford: OUP, 1966)

⁴Bermudez, J. 'Evans on the Sense of "I"', in Bermudez (ed.) *Thought, Reference and Experience: Themes from the Philosophy of Gareth Evans* (Oxford: Oxford University Press, 2005)

⁵ See Evans, G. *The Varieties of Reference* p. 206.

which satisfy the definition. Direct reference theorists accept a very coarse-grained conception of thoughts on which they are simply composed of objects and properties. On such a conception the thought expressed by my token of the sentence ‘I am F’ and the thought expressed by my token of the sentence ‘NN is F’ will be one and the same thought: the unique thought which is composed of NN and the property of being F (at a particular time). That thought is equally adequately expressed by either of my two sentences. So it will not be *most* adequately expressed through the sentence which contains the first person.

The upshot is that Evans’s assumption that there are thoughts which satisfy his definition of ‘“I”-thought’ is slightly tendentious. However, this should not distress us overly. Most direct reference theorists will agree that the *mental state* expressed by my utterance of ‘I am F’ and the mental state expressed by ‘DM is F’ are different. A neo-Fregean like Evans or Bermudez says that they are different mental states because they involve attitudes to different thoughts. A direct reference theorist who wishes to say they are different mental states must find some other way to label the difference. Perry, for example, talks of thoughts about oneself which are apprehended ‘in a first-personal way’. Kaplan talks of thoughts about oneself which are entertained ‘under the character of “I”’. We could, if we wanted, translate Evans’s talk of “I”-thoughts into Perry or Kaplan’s terminology. Equally, we could offer Evans’s positive account of what it is to think an “I”-thought as an account of what it is to apprehend a coarse-grained thought ‘in a first personal way’ or ‘under the character of “I”’.

Admittedly, there are other more extreme direct reference theorists, such as Nathan Salmon, who deny not only that my utterance of ‘I am F’ and my utterance of ‘NN is F’ express different thoughts but also deny that they express different mental states.⁶ On Salmon’s view, attitudes are two place relations which hold between a subject and a coarse-grained thought. I stand in the belief relationship, for example, to a particular thought so long as there is at least one mode of presentation under which I believe that thought. So long as I believe that I am in strange hospital bed, I will also believe that NN is in a strange hospital bed- even if I am suffering from amnesia and have forgotten that my name is ‘NN’. The oddness of attributing to me the belief that NN is in a strange hospital bed in this kind of situation is given a pragmatic explanation. The attribution sounds odd because, despite being true, it is infelicitous. In particular, it is infelicitous because it suggests that I would assent to the sentence ‘NN is in a strange hospital bed’. Neo-Fregeans must ultimately say something to counter Salmon’s strikingly simple view of attitude ascriptions. In particular, they must say why the phenomenon for which Salmon gives an explanation at the level of pragmatics deserves to be explained at the level of semantics. I shall ignore Salmon’s view here. It is worth pointing out though that even on Salmon’s view the question arises of how my mind must be if I can felicitously, and not just truly, say ‘I believe that I am in a strange hospital bed’. It seems plausible that the positive account Evans’s puts forward as an account of what it is to think an “I”-thought could also figure as a large part of the answer to Salmon’s question. The more general point is that although Evans himself presents his functional role account within the context of a particular neo-Fregean semantic framework there is no reason to assume that his account is wedded to that framework. This is why it should not come as a surprise

⁶ See Salmon, N. *Frege’s Puzzle*. (Cambridge, MA: MIT Press, 1986).

that John Perry, who works within a neo-Russellian framework, is the philosopher other than Evans who has done most to explore the possibility of providing a functional role account of the sense of “I”.

The sense of “I” is defined as the sense expressed by tokens of the word “I”. An “I”-thought is defined as any thought expressed by an utterance of a sentence containing the word “I”. So the sense of “I” is the component which is common to all and only “I”-thoughts. As I understand him, Evans aims to give an account of what all and only “I”-thoughts have in common. Since the sense of “I” just is what all and only “I”-thoughts have in common he can put this account forward as an account of the sense of “I”. It is because his account of “I”-thoughts is a functional role account that his account of the sense of “I” is a functional role account. And it is because his account of “I”-thoughts entails that they are not shareable that his account of the sense of “I” entails that it is not shareable. This should become clearer once a bit more of the general picture has been roughed in.

The central idea behind a functional role account of “I”-thoughts is a claim about how one can informatively characterize such thoughts. It is the claim that one can informatively characterize “I”-thoughts by saying something about the causal and/or normative relationships in which they stand to some or all of: other thoughts, ways of gaining information and ways of acting. Among positions which accept this core claim there is obviously scope for considerable variation – e.g. in how much attention is paid to each of the three relata just mentioned, and in how much stress is put on the causal

aspects, and how much on the normative aspects, of the relevant relationships. What makes it plausible that *some* such position will prove correct is that “I”-thoughts clearly do stand in a relationship to ways of gaining information and to ways of acting which is highly distinctive. This basic insight, which became explicit in the work of Casteneda, Anscombe, Perry and Evans, was responsible for a shift in philosophical orthodoxy. It convinced most philosophers to abandon the view (held, e.g., by Carnap and most of his contemporaries) that indexicals are a mere practical convenience and to accept instead that they are, to use Perry’s word, ‘essential’— i.e., that they can be used to express things which cannot be expressed unless one uses them. The insight behind a functional role account of “I”-thoughts is that we can appeal to the highly distinctive relationships in which “I”-thoughts stand in order to say what “I”-thoughts are. I will now set out in a bit of detail some of the distinctive relationships that obtain between “I”-thoughts and ways of gaining information and between “I”-thoughts and ways of acting. I will also, in passing, look at one example of a functional account of “I”-thoughts- my own.

Many “I”-thoughts are directly sensitive to special ways of gaining information which provide information about oneself but about no one else. For example, my thought that I am cold is directly sensitive to a way of gaining information about temperature which only provides information about my own temperature. By contrast, even if I am NN, my thought that NN is cold will not be *directly* sensitive to that way of gaining information. (This shows up in the fact that I am suffering from amnesia, and do not know that I am NN, the fact that my special way of gaining of providing information about my own temperature is operating as normal will not help me to decide whether NN is cold. It will

not since the ordinary connection between my thought that NN is cold and my special way of gaining information about my own temperature is *mediated* by my knowledge of the identity premise that I am NN, and, in this special case, that knowledge is missing.) The functional role of an “I”-thought like the thought that I am cold thus has a distinctive input, or informational, component. “I”-thoughts’ output, or action, components are also distinctive. Given the appropriate desires on my part, the thought that I am being approached by a hungry bear is an immediate spur to action. By contrast, the thought that NN is being approached by a hungry bear depends for its impact on the additional knowledge that I am NN. This is knowledge which, in certain circumstances, I may not have- e.g. if I am suffering serious amnesia.

It is easy to point out that many “I”-thoughts stand in a distinctive relationship to action and information. It is somewhat harder to provide a correct statement of necessary and sufficient conditions for a thought to be an “I”-thought. One attempt to provide such a statement is encapsulated in the following quote from Evans:

We clearly do have special ways of gaining knowledge of ourselves, and “I”-thoughts are thoughts which are controlled, or disposed to be controlled, by information gained in these ways.⁷

There are many counterexamples to Evans’s claim. My thought that I shall die in Havana is an “I”-thought but it is not controlled, or disposed to be controlled, by information gained in any special way I have of gaining information about myself. Neither is the

⁷ See Evans, G. (1982). *The Varieties of Reference* p. 253.

thought that I was breastfed (an example Evans himself mentions). Some “I”-thoughts – e.g., the thought that I am hot, the thought that my knees are crossed – are controlled by special ways of gaining information of ourselves. Some though are not. Evans was aware of the existence of thoughts which cannot be established on the basis of special ways of gaining information about oneself, and considered it important to be able to account for our grasp of their truth-conditions. But he failed to notice that they are counterexamples to the general claim he makes about “I”-thoughts in the quotation we have just looked at.

My own preferred functional-role characterization of “I”-thoughts, like Evans’s, gives special ways of gaining information about oneself a key role in saying what “I”-thoughts are. However, it is designed so as to be compatible with the simple point that there are plenty of “I”-thoughts which are *not* sensitive to any special way of gaining information about oneself. My proposal is that a thought is an “I”-thought just in case either:

- (i) it is sensitive to a special way of gaining information about oneself.

or

- (ii) it involves the same singular sense as any thought which meets condition (i).

A thought like the thought that I am now hot qualifies as an “I”-thought in virtue of meeting condition (i). A thought like the thought that I shall one day die in Havana qualifies as an “I”-thought in virtue of meeting condition (ii). Evidence that the thought that I shall one day die in Havana really does meet condition (ii) is provided by the fact that the following inference:

I am now hot

I shall one day die in Havana

Something which is now hot will one day die in Havana.

is a perfectly good one to perform. The inference would not though be a perfectly good one to perform if the singular senses involved in the two thoughts were different. In that case, the inference would involve a fallacy of equivocation.

Certainly, the particular functional role based account of “I”-thoughts I have just offered needs more development. For example, some justification for connecting up the notion of sense with that of a good inference would need to be supplied. Perhaps the account will even turn out to be incorrect. Its correctness however is not what concerns me in this paper. I have argued that the claim that *some* functional role based account of “I”-thoughts is correct has plausibility which is independent of the plausibility of any particular version of such an account. Having argued that, my main concern will be to show that the fact that functional role accounts entail the unshareability claim should not be held against them. First though I need to explain why it is that functional role accounts of “I”-thoughts do entail the unshareability claim.

I stand in an input relationship to myself which no one else stands in- for example, only I receive proprioceptive information about myself (of course, other people stand in a corresponding privileged relationship to themselves). I also stand in an output

relationship to myself which no one else stands in – only I am capable of directly determining what my actions are to be (of course, other people stand in a corresponding privileged relationship to their actions). Any functional role account of “I”-thoughts will write one or other or both of these privileged relationships into the identity of an “I”-thought. For example, the idea may be that a thought is not identical with my thought that I am being pursued by a bear unless it feeds directly into my actions. Or the idea may be that a thought is not identical with my thought that I am cold unless it is directly sensitive to my way of gaining information about my own temperature. But it is clear that no thought which anyone else can entertain is going to be directly sensitive to *my* way of gaining information about my temperature, and that no thought which any one else can entertains is going to feed directly into *my* actions. Hence, a functional role account of “I”-thoughts will entail that “I”-thoughts are unshareable.

The argument of the last paragraph presupposes that it can be an aspect of the functional role of my “I”-thought that it is sensitive to *my* special way of gaining information about myself. On a different, ‘narrow’, way of individuating functional roles, the functional role of an “I”-thought will include only the fact that it is sensitive to *someone’s* special way of gaining information about herself. In that case, my “I”-thoughts and your “I”-thoughts may well have the same functional roles. So, on that narrow way of individuating functional roles, it isn’t clear why a functional role account of “I”-thoughts should be thought to entail the unshareability claim.

Neither Evans nor Bermudez considers this objection. There is however a ready reply to it which either could give. The “I”-thoughts Evans and Bermudez are trying to give an account of are broad thoughts. On their view, my “I”-thought and your “I”-thought are different thoughts because they are about different objects. If functional roles are narrow, and “I”-thoughts are broad, then we cannot individuate “I”-thoughts just in terms of their functional roles. Rather, we will have to think of “I”-thoughts as composite entities each of which has a functional role and a person as its components. But, in that case, to entertain my “I”-thought it will not be enough to entertain a thought with a particular narrow functional role. Entertaining my “I”-thought will require that one entertain a thought which (i) has a particular narrow functional role and (ii) is about me. And this is something only I can do. Your “I”-thoughts, though they may have the right sort of functional role, are about the wrong object to be identical to my “I”-thoughts. Your “you”-thoughts, though they are about the right object, have the wrong sort of functional role to be identical to my “I”-thoughts. (Bermudez thinks that some of your “you”-thoughts *must* be identical to my “I”-thoughts and it is precisely for this reason that he rejects a functional role account of “I”-thoughts). So any functional role account of “I”-thoughts will entail that no one other than myself can entertain my “I”-thoughts. From the point of view of this entailment, it makes no difference whether we individuate functional roles in a broad or a narrow way.

If we accept a functional role account of “I”-thoughts we will have to accept that no one else can entertain my “I”-thoughts. Since entertaining, or grasping, a sense just is entertaining some thought which has that sense as a component, we will also have to

accept that no one other than myself can grasp the sense expressed by my tokens of “I”- i.e. that the sense expressed by my tokens of “I” is unshareable. Notice though that one cannot just gloss this by saying that a functional role account of “I”-thoughts forces us to accept that no one else can *understand* what I say when I use the word ‘I’. Grasping is stipulated to be a relationship in which a subject stands to a sense. Understanding is a relationship in which a subject stands to a linguistic entity such as a word or a sentence. The inference from the claim that no one else can grasp the sense I express when I use the word “I” to the claim that no one else can understand what I say when I use the word “I” presupposes a particular view of the relationship between thought and linguistic understanding. Precisely because it would be disastrous to have to deny that others can understand my uses of the word “I”, anyone defending the unshareability claim ought to reject that view of the relationship between thought and linguistic understanding. A defender of the unshareability claim can allow that there are *some* connections between someone’s having the capacity to entertain thoughts of a particular kind and their having the capacity to understand tokens of certain words.⁸ For example, she can accept the plausible claim that only someone who is capable of entertaining “I”-thoughts herself will be capable of understanding my use of the word “I”. What she must deny is only that understanding one of my tokens of “I” requires another person to entertain a thought which is identical to my “I”-thought as part of that very act of understanding. Later, I will argue for a view of the relationship between thought and linguistic understanding which underpins this denial.

⁸ Thanks to an anonymous referee for encouraging me to make this point explicit.

It is worth making explicit something which I have been glossing over – namely, that there are two strands to Evans’s account of the sense of ‘I’, of which the functional role strand I am defending is just one. The other strand largely consists in Evans’s attempt to show how a subject thinking an “I”-thought *knows which* thing her “I”-thought is about. Evans thinks the functional role account of the sense of “I” does not explain this, and that it needs to be explained in terms of a subject’s capacity to know where she is. This knowledge is itself explained in terms of a subject’s capacity to identify her egocentric location within a cognitive map of the world. Since the knowing which requirement is very important to Evans, so too is this strand of his account of the sense of “I”. One reason I have not discussed it is that I find the knowing which requirement deeply obscure and doubt that there is any cleaned up version of it which is both plausible and non-trivial.⁹ Others will disagree about this. The more important point is that we can be convinced that an account of the sense of “I” must mention aspects of functional role without first having decided whether it must include other things as well- in particular, whether it must include a substantial account of how subjects know which things their “I”-thoughts are about. I have argued that the account of the sense of ‘I’ must mention the functional role of “I”-thoughts and that, if it does, it will entail the unshareability claim. So the account of the sense of ‘I’, whether or not it includes an account of how the knowing which requirement is satisfied, will entail the unshareability claim. That is enough to justify an attempt to defend that claim, which is the substance of what follows. My focus will be on two arguments against the unshareability claim given by Bermudez.

⁹ For a clear sense of the difficulties involved in explaining and defending the requirement see Sainsbury, Mark ‘Evans on Names’ in Bermudez (ed.) *Thought, Reference and Experience: Themes from the Philosophy of Gareth Evans* (Oxford: Oxford University Press, 2005)

Perhaps there are other arguments against the unshareability claim, though I am not aware of any.

Objectivity and Shareability

Evans insists that “I”-thoughts are not shareable. But he, like Frege, also insists that all thoughts are objective. He makes no exception for “I”-thoughts. Given that, one way of putting pressure on Evans to modify his account of the sense of “I” so that it does not entail the unshareability claim is to establish that if an “I”-thought is not shareable, then it cannot be objective either. Bermudez tries to establish this:

Evans proposes two respects in which thoughts involving the sense of “I” (as he is proposing to understand it) should count as objective. First, they are not reducible to the contents of an individual consciousness. Second, they exist and have a truth-value independently of anyone entertaining them.

Why might thoughts of a given class not be reducible to the contents of an individual consciousness, even though they are not shareable? There would be such irreducibility if there were thoughts of that class that have never been thought by anyone. This certainly makes perfectly good sense for e.g., mathematical thoughts. It cannot however be applied to non-shareable “I”-thoughts, simply because the identity of any given token-reflexive thought is a function of the episode of thinking. Recall that the reference of “I” in a sentence involving the first-person pronoun is fixed by the token-reflexive rule that (in

standard cases) picks out the speaker of the relevant token sentence as the person to whom the pronoun refers. Since an “I”-thought is a thought expressed by a sentence involving the first-person pronoun, one would expect the reference to the first-person component of the thought to be fixed in an analogously token-reflexive way. *The whole point of token-reflexivity, and indexicality in general, is that the identity of the thought is determined by the context in which it is thought – , there is no thought without an episode of thinking.* If this episode has to be an episode featuring the subject of the thought (as it must be if the thought is non-shareable), then I see no sense in which the thought can exist and have a truth-value independently of being thought by the particular person who is the subject of the thought (my italics).¹⁰

Bermudez and Evans agree on three things. They agree that if a thought is irreducible to the contents of an individual consciousness (if, for short, it is irreducible), then it is objective. They agree that if a thought has a truth-value independently of being thought (if, for short, it is truth-value independent), then it is objective. They also agree that if a thought is neither truth-value independent nor irreducible, then it is not objective. They disagree about whether objectivity and unshareability are compatible. Evans thinks that unshareability and objectivity are compatible because he thinks that unshareability is compatible with irreducibility and also with truth-value independence. Bermudez thinks that they are not compatible because he thinks that unshareability is incompatible with irreducibility and also with truth-value independence. I think that Bermudez’s argument for his position is supposed to run as follows:

¹⁰ Bermudez. “Evans on the Sense of ‘I’” p. 183.

((P1)) The existence of an indexical thought, such as an “I”-thought, depends on there being some episode of thinking that thought (“There is no thought without an episode of thinking.”).

(P2) If “I”-thoughts are unshareable, then there can be no episode of thinking an “I”-thought which is not performed by the subject of that “I”-thought.

(C) If “I”-thoughts are unshareable, then the existence of any “I”-thought depends on there being an episode of thinking that thought which is performed by the subject of that thought.

(C) fairly obviously secures the result that unshareability is incompatible with truth-value independence – if my “I”-thought depends for its very existence on being thought by me then it will certainly depend for its having a truth-value on being thought by me (non-existent thoughts don’t have truth-values). And if it depends for its having a truth-value on being thought by me then it isn’t truth-value independent. The connection between (C) and irreducibility isn’t quite as obvious. But I think that Bermudez means us to assume that unless “I”-thoughts exist and have a truth-value independently of being thought there is no good reason not to think them reducible to the contents of an individual consciousness. Granted that assumption, Bermudez, so long as he able to establish (C), will also be able to establish both of the conditionals which he accepts and Evans rejects: the conditional that if “I”-thoughts are unshareable, then they are not truth-value

independent and the conditional that if “I”-thoughts are unshareable, then they are not irreducible. Once he has established these conditionals, he will be in a position to conclude that, pace Evans, if “I”-thoughts are unshareable they are not objective. After all, Evans thinks that unshareability and objectivity are compatible precisely because he thinks that unshareability is compatible with irreducibility and with truth-value independence.

Without (P1), Bermudez’s argument for (C) collapses. What should we make of (P1)? In deciding, the sentence in the passage we need to look at is the one I have italicized. Its final clause, the one that follows the dash, is (P1).

The whole point of token-reflexivity, and indexicality in general, is that the identity of the thought is determined by the context in which it is thought—, there is no thought without an episode of thinking.

The sentence is surprising, given that Bermudez generally follows Evans in individuating thoughts in a broad way. On a broad way of individuating thoughts, the thought I express with ‘I am F’ and the thought you express with ‘I am F’ are different thoughts. There is no single “I”-thought which is sometimes entertained in a context in which I am its producer and sometimes entertained in a context in which you are its producer. So it is at best highly misleading to say that ‘the identity of the thought is determined by the context in which it is thought’. (That misleadingly suggests that there is more than one thought in which that very thought might be entertained).

I think we should understand Bermudez as making a claim about narrow thoughts, which we can harmlessly think of as the *vehicles* of broad indexical thoughts. The claim made by the part of Bermudez's italicized sentence running up to the dash is the claim that which broad indexical thought an indexical vehicle expresses depends on context. Suppose, for the sake of vividness, that there are mental utterances of indexical sentences and that these are the vehicles of indexical thoughts. Then the claim will be that which thought a mental utterance of an indexical sentence expresses depends on context. For example, whether a mental utterance of the sentence 'I am F' expresses a thought which is about me will depend on whether I am its producer, and that I am its producer is a fact about context. This claim seems innocuous enough. The clause which follows the dash – (P1) – is presumably supposed either to be a gloss on, or an implication of, this claim. However it is nothing of the sort. From the fact that context determines which indexical thought a mental utterance of an indexical sentence expresses it does not follow that the existence of an indexical thought depends on someone or other's actually having produced a mental utterance which expresses that thought. Compare: which target, if any, at the fair-ground shooting stand one hits depends on which direction one aims in. But it would be a non-sequitur to conclude that the existence of a particular target depends on someone or other's actually have fired a shot which hits that target. If there is a slack day at the fair and no one fires a shot which is aimed in the right direction to hit a particular target, that target can still exist. The whole point of indexicality may be that the identity of an indexical thought expressed by a particular indexical vehicle is determined by the context in which that indexical vehicle is produced. This does not show though that there

can be no indexical thought without an indexical vehicle – i.e. without an episode of thinking. So we have no good reason to accept (P1). Without (P1), there is no sound argument for (C) and without (C) the argument that unshareability and objectivity are incompatible collapses. In that case, the attractiveness of the view that “I”-thoughts are objective cannot be used in an argument against the unshareability claim. In the next section, I focus on Bermudez’s back up argument against the unshareability claim. It raises some much more general issues about the relationship between thought and linguistic understanding.

The Relationship between Thought and Linguistic Understanding

What, in general, does linguistic understanding require you to think or know? Understanding a sentence, presumably, requires only that one have knowledge of the linguistic meaning of that sentence which one has derived from one’s knowledge of the linguistic meaning of the words which make it up. Bermudez assumes though that can we also apply the notion of understanding to *utterances* of sentences. Understanding an utterance of a sentence is tied to the success of the particular communicative act the utterance is involved in. It seems plausible that understanding an utterance will at least require that one think a thought about the particular object the utterance is about (Of course, this does not apply to utterances of existentially general sentences. Here, my focus is entirely on utterances which are about particular objects.). Bermudez thinks, however, that understanding an utterance requires something much more demanding than this. He thinks we should accept the following claim:

Understanding a speaker's utterance requires thinking the very thought the speaker expresses in that utterance (U).

U is incompatible with the unshareability claim. For, as Bermudez observes, you can understand what I say when I utter the sentence 'I am F'. Moreover you can do so by thinking a thought which you would express by uttering the sentence 'You are F'. With U in place it immediately follows that the thought my utterance of 'I am F' expresses is the very same thought as the thought your utterance of 'You are F' expresses. That is, your "you"-thought is the same thought as my "I"-thought. But the unshareability claim implies that no one can entertain anyone else's "I"-thought. Something has to be given up- either U or the unshareability claim. The thrust of Bermudez's second argument against the unshareability claim is just that maintaining the unshareability claim requires giving up U. Bermudez acknowledges that Evans explicitly rejects U¹¹ but he regards this as an unacceptable manoeuvre Evans is forced into by his acceptance of the unshareability claim.

One unfriendly response to Bermudez's argument would be to deny that there really is any robust notion of utterance understanding which we can be expected to have definite intuitions about. I wish to avoid relying on this response though since I accept that we do have an intuitive grasp on the notion of utterance understanding (Evans also accepts this). Instead, I shall argue that the requirement U imposes on utterance understanding is too

¹¹ See Evans, *The Varieties of Reference* p. 316 for a clear statement of the rejection. Evans rejection of the claim constitutes one of his main departures from Frege.

demanding. There are cases in which, intuitively, understanding occurs but in which U entails that understanding does not occur. I describe one such case. Having done so, I will consider Bermudez's motivation for accepting U. If the motivation for U were strong enough, then we might be prepared to accept it, despite the unintuitiveness of some its verdicts about particular cases. But the motivation is not strong enough. In arguing for U Bermudez does point to a genuine insight. However, to make sense of that insight we do not need to accept U. We can make sense of it by accepting a somewhat similar, but weaker, claim. Crucially, this weaker claim is compatible with the unshareability claim.

Before describing the counterexample to U, I need to say a bit about the conception of thoughts Bermudez and Evans share. On their fine-grained conception of thoughts, two thoughts can count as different, even though they ascribe the same property to the same object, because they involve two different ways of thinking of the same object (or two different ways of thinking of the same property). One way of cashing out this basic idea is to appeal to Frege's intuitive criterion of difference for thoughts. The intuitive criterion says that it is not possible for a subject coherently to take different attitudes to the same thought at a given time.¹² If this criterion is acceptable, then we can show that two thoughts are distinct, and involve different ways of thinking of their object, by showing that it is possible for a single subject rationally to take different attitudes to them at a given time. To refute U it suffices to provide an example of a case in which a hearer understands a speaker's utterance even though the thought the hearer thinks and the thought the speaker expresses count as different by the intuitive criterion of difference. I will now do this, using a perceptual demonstrative example. That I shift from "I"-

¹² See Evans, *The Varieties of Reference* p. 193.

thoughts to perceptual demonstrative thoughts is not unreasonable. U is an entirely general claim about the relationship between thought and linguistic understanding. It applies to perceptual demonstrative thoughts as much as it does to “I”-thoughts. One reason it makes sense to shift from “I”-thoughts to perceptual demonstrative thoughts at this point is that the latter are somewhat easier to individuate than the former. Our best guide to individuating fine-grained thoughts is the intuitive criterion of difference. But, as Evans points out, this criterion cannot be appealed to in order to decide whether the thought expressed by my utterance of ‘I am F’ and the thought expressed by your utterance of ‘You are F’ are the same or different. The criterion only applies to thoughts which can be entertained by a single subject. But, as it happens, to work out whether the thought expressed by my utterance of ‘I am F’ and the thought expressed by your utterance of ‘You are F’ can be entertained by the same subject we already need to know whether they are the same or different.¹³

The example I offer as a counter-example to U is one used by Perry to make a somewhat similar point.¹⁴ Two people standing at different ends of a room are looking out at the same ship through two different windows, w1 and w2. One of them, looking through w1, says ‘That ship was made in Japan’. It seems entirely possible that the other person, looking through w2, will understand this utterance by thinking a demonstrative thought he would express with his utterance of ‘That ship was made in Japan. That’s what he’s saying’. If we deny this, if we say that the mere fact that the two people have different views onto the same object entails that they cannot understand each others’ demonstrative

¹³ See Evans, *The Varieties of Reference* p. 193.

¹⁴ See Perry, J. (1977). “Frege on Demonstratives” p. 483.

utterances, then we are forced to accept that one hardly ever understands an utterance of someone else's in which a demonstrative is used to refer to some perceptually salient object. For one hardly ever has exactly the same view onto an object as one's conversational partners. Nevertheless, the thoughts the two people think do qualify as different thoughts by the intuitive criterion of difference. One way to see this is to consider a case which is exactly like this one except that instead of having two subjects we just have a single subject. The single subject first stands in the position of the speaker in the first example and looks at the ship through w_1 and then stands in the position of the hearer in the first example and looks at the ship through w_2 . Such a subject could coherently wonder whether the ship he sees through w_1 and the ship he sees through w_2 are the same ship. We can even imagine him being persuaded that they are not and accepting the thought that that ship (seen through w_1) was made in Japan while refusing to accept the thought that that ship (seen through w_2) was made in Japan. By the intuitive criterion of difference, then, his two thoughts are different. It seems right to say that the single subject's first thought is the same thought as the thought entertained by the speaker in the original case- after all, both are perceptual demonstrative thoughts made on the basis of the very same kind of perception of the ship through w_1 . It also seems right to say that the single subject's second thought is the same thought as the thought entertained by the hearer in the original case- after all, both are perceptual demonstrative thoughts made on the basis of the very same kind of perception of the ship through w_2 . Since his two thoughts are different, their two thoughts are different. So the original ship case is a case in which speaker and hearer understand the same utterance despite thinking different thoughts. It is a counterexample to U. The fact that the unshareability claim is

incompatible with U is not a reason for rejecting the unshareability claim. Examples which have nothing to do with “I”-thoughts force us to reject U anyway.

Should we conclude on the basis of the ship case that all that is required for utterance understanding is that both hearer and subject each think thoughts which are about the same object (and the same property)? Such a view would certainly be compatible with the unshareability claim. However, it does not seem to be correct. In seeing why not, we will see what motivated Bermudez to accept U. There are cases in which, intuitively, understanding does not occur even though a hearer and speaker do both think about the same object and the same property. Bermudez makes a plausible suggestion about what is goes wrong in such cases. His suggestion is that, in general, if a communicative act successful then it should permit the transmission of knowledge from speaker to hearer.¹⁵

That the speaker and hearer both think about the same object and property is not enough to guarantee understanding precisely because it is not enough to guarantee the transmission of knowledge from speaker to hearer. There is considerable work to be done in making this kind of knowledge-condition on understanding defensible and precise (we need, for example, to decide how it applies to cases in which the speaker’s utterance does not express knowledge; and to cases in which the hearer, despite understanding, fails to form a belief, or does form a belief but has his knowledge defeated by conflicting evidence). Nevertheless, some such condition may well be correct. The objectionable inference Bermudez makes is the inference from the claim that there is some such knowledge-condition on understanding to the claim that there is always just *one* way of thinking of the object an utterance is about – the way of thinking about that object

¹⁵ Bermudez, *Evans on the Sense of “I”* p. 186.

employed by the speaker – such that understanding the utterance requires one to think of that object in that way. That there is such a requirement, and that it is violated in the cases where knowledge is not transmitted and understanding does not occur despite the fact that both subjects are thinking about the same object, is one explanation of what is going wrong in such cases. It is the explanation Bermudez favors and it constitutes his reason for accepting U. However, there is an alternative explanation.

The alternative explanation is that there is, in general, a *range* of ways of thinking of an object such that thinking about the object using any member of that range is sufficient for understanding an utterance about that object. What is going wrong, on this view, in the kind of case in which understanding does not occur even though the hearer is thinking about the right object, is that the hearer employs a way of thinking of the object which is outside the range of acceptable ways of thinking. Ultimately, of course, a defender of this view will want to give some account of what determines membership of the acceptable range.¹⁶ Even before this has been done though it is clear that there is an advantage which the second explanation has over the first: it does not entail that understanding does not occur in cases, such as the case involving the ship viewed through two different windows, in which two subjects do not think of the same object in the same way. The converse claim – that the first explanation enjoys some advantage over the second – does not seem plausible. The main reason it does not is that someone who defends the second explanation can happily accept that there may be special cases in which all but one way

¹⁶ For an attempt to do this see Heck, Richard, "Do Demonstratives have Senses?", *Philosophers' Imprint* vol. 2, no. 2 (2002), pp. 1-33.

of thinking of an object is excluded from the range of ways of thinking about that object which are sufficient for understanding. The only point she need insist on is that there is no reason to expect such cases to be rule. Given that, there would need to be some special consideration to persuade us to think that understanding an utterance of ‘I am F’ is only compatible with one way of thinking about the object of that utterance, the way of thinking employed in the utterer’s “I”-thought. Bermudez provides no such special consideration. His strategy is rest everything on an appeal to an entirely general claim – U – about the relationship between thought and understanding.

It is plausible that you can understand my utterance of “I am F” by thinking a “you”-thought about me even though thinking about someone using an “I”-thought and thinking about someone using a “you”-thought are two quite different ways of thinking about that person. It is also plausible that not just any way of thinking of an object will do if one is to understand an utterance about that object. We can accept both of these plausible claims because it is consistent to hold both that, for understanding, ‘not just any way of thinking of an object will do’ and also that it isn’t the case that ‘there is only one way of thinking which will do’.

The price of accepting U, as Bermudez does, is that we have to concede that understanding occurs much less frequently than we ordinarily take it to occur. I have tried to show that there is no good reason to pay that price. However, since it appears that Bermudez might well be willing to pay it, it is worth saying why his somewhat persuasive attempt to make it seem acceptable ultimately fails. Bermudez appeals to a

distinction between cases in which understanding actually occurs and cases in which understanding comes close enough to occurring for ascriptions of understanding to pass muster, even though they are not strictly speaking true. Appealing to such a distinction certainly requires giving some account of what is special about the ascriptions of understanding which are strictly speaking true. But it is relatively easy to see what Bermudez's suggestion on this score is going to be, at least when it comes to utterances involving "I" and "you". According to Bermudez, what is distinctive about "you", what makes it the case that thinking a "you"-thought about John is sufficient to understand sentences of his containing "I" although thinking a "John"-thought about John would not be, is that thinking about someone using "you" involves having the capacity to locate him. On Bermudez's view, thinking an "I"-thought involves having the capacity to locate oneself.¹⁷ This ensures that there is a sense in which the way in which you think of me and the way in which I think of me is the same when I think 'I am F' and when you think 'You are F': each of our ways of thinking of me involves having the capacity to locate me. So Bermudez's suggestion will be that only in cases in which one has the capacity to locate a particular utterer of the sentence 'I am F' is it strictly speaking correct for one to be ascribed understanding of what he says with that utterance.

This way of distinguishing cases of understanding from cases of near-understanding delivers some unintuitive results. Consider one such result which Bermudez himself mentions: it seems that I can understand your utterance of "I" perfectly well when I talk to you on the phone. However, since I may well not have the capacity to locate someone I am speaking to on a phone, Bermudez has to deny this (Of course, in one sense, I may

¹⁷Bermudez, *Evans on the Sense of "I"*, p. 192-4 passim.

have the capacity to locate him since I could persuade him to tell me where he is. But the notion of a capacity to locate is being understood by Bermudez in a more demanding sense than this, and needs to be if it is to the theoretical work he wishes it to do). To make this denial more palatable Bermudez says the following:

There are of course degrees of success in communication and degrees of understanding. A conversation can proceed perfectly well in circumstances in which one might hesitate to attribute full understanding to the parties concerned. The notion of understanding is an ideal to which we may approximate more frequently than we attain.¹⁸

On its own this reply is unconvincing. The main problem with it isn't just that the claim that the notion of understanding is an ideal to which we do not often attain (like Unger's claim that the notion of knowledge is an ideal to which we do not often attain) requires an error-theory of our practice of ascribing understanding (or, in Unger's case, knowledge) to each other. The most obvious problem with it is just that *we have no idea what the source of the ideal in question is supposed to be* (by contrast, one can just about see why someone might think that infallibility is an ideal at which beliefs in some sense aim). Why should there be such a tight connection between having the capacity to locate someone and being able to understand utterances they produce containing "I"? From an intuitive point of view, chatting away fluently to someone on the phone can be a paradigm case of understanding. Perhaps intuition leads us astray here. But, in that case, it is surely correct to ask why it should do so. The only convincing way of answering this

¹⁸ Bermudez, *Evans on the Sense of "I"*, p. 190.

question is to appeal to theoretical considerations to which untutored intuition is likely to be insensitive. One might buy into one important aspect of Evans's theoretical framework: accept that there is a 'knowing which' requirement on understanding and accept as well that in the case of a use of a singular term to refer to a material object such as a human-being, satisfying that requirement involves having the capacity to locate the object one is thinking about. In that case, we might have a good theoretical reason for saying that what happens when one hears tokens of "I" produced by someone whom one lacks the capacity to locate is not really understanding 'in the strict sense'. However, Bermudez evinces no willingness to accept Evans's 'knowing which' requirement. This is, in a way, a good thing. The knowing which requirement isn't an obviously attractive feature of Evans's theoretical framework and it is an important virtue of the functional role account of the sense of "I" I have argued for that it does not presuppose that requirement. The problem is that, without appealing to the knowing which requirement, it is very hard to see why we should believe that there is a tight connection between having knowledge of someone's location and understanding utterances of theirs which contain the word "I". On the face of it, the telephone-conversation case we have just discussed reveals that there is no such tight connection.

The way forward, for those trying to give an account of the sense of "I", is (i) to concentrate on fleshing out the suggestion that there is a connection between the sense of "I" and the distinctive functional role of "I"-thoughts (and on saying exactly what the relevant distinctive functional role is). The fact that such an approach entails the

unshareability claim should not be held against it, since no good arguments have yet been provided for denying the unshareability claim.