Commentary on D. Morgan, « A Demonstrative Model of First-Person Thought »

François Recanati

1

Morgan takes for granted the existence of 'first person thoughts', but he does not tell us exactly what they are. He gives us examples, though, and from the examples he gives we see that what he means by a first person thought is the sort of thought one may express by using the first person : 'I am hungry', 'I was born in Cork', 'I have crossed knees'. Such thoughts he rightly takes to be about the thinker of the thought, but he does not use that feature – reflexivity – to define the thoughts in question. Indeed one cannot ; for a thought can be about the thinker of the thought without being the sort of thought one would express by using the first person. One may have a thought about oneself even though one does not realize that the thought is about oneself (as when one sees oneself in the mirror, without realizing that one is self-seeing). So, not all thoughts that are about ourselves are first person thoughts. In the other direction, however, the entailment holds : a first person thought will be about oneself, as a matter of necessity. The question Morgan asks in his paper is : Why ? Why is a first person thought a thought about oneself, i.e., a thought about the thinker of the thought ?

In a sense, this is an easy question, i.e. a question it is easy to answer. It is easy to answer precisely because the only characterization of first person thoughts that is (implicitly) offered in the paper is the characterization in terms of the linguistic means we would use to express the thought. Let us suppose that the thinker holds a thought that is *not* about himself or herself. Could he or she express such a thought by using the first person ? Answer : No, because we use the first person to refer to ourselves. (That is a linguistic fact about the first person.) The only situation in which it seems that we might use the first person to express a thought that is not about ourselves would be a situation in which we mistakenly *believe* that the thought is about ourselves. Let us consider such a situation. If I believe that the winner will go to Tahiti, and I mistakenly believe that I am the winner, my thought 'the winner will go to Tahiti' will not be about myself but I will take it to be about myself. As a result, I will be in a position to use the first person in expressing my belief : 'I will go to Tahiti'. In such a case, however, there are *two* thoughts : the initial thought ('the winner will go to Tahiti') is *not* about myself, but the mistaken belief that I am the winner leads me to form another thought which *is* about myself : the thought that I will go to Tahiti. It is that thought — a thought about myself — which I express by using the first person. Conclusion : a first person

thought necessarily is a thought about the thinker because (i) it is a thought one would express by using the first person, and (ii) the first person is governed by the rule that we use it (only) to refer to ourselves.

Simple as it is, there is something deeply dissatisfying about this answer. We appeal to the notion of a first person utterance to characterize first person thought, but when we attempt to characterize the (linguistic) first person, we appeal to the token-reflexive rule : the fact that we use the first person to refer to ourselves. Now, to follow the token-reflexive rule, we need to think of ourselves *as ourselves* — we need to think first person thoughts. Suppose I *am* the winner and I believe, as before, that the winner will go to Tahiti. This will not lead me to say 'I will go to Tahiti' unless I also believe that *I* am the winner and (therefore) that *I* will go to Tahiti. But what is it to believe that I am the winner, or that I will go to Tahiti ? These are first person thoughts, aren't they ? What this shows is that the capacity to use the first person in language presupposes the capacity to think first person thoughts. It follows that the characterization of first person thoughts in terms of first person utterances is ok only if it's a way of 'fixing the reference' of the phrase 'first person thought', for exposition purposes ; but it cannot be used to reach any substantive conclusion regarding first person thoughts. In particular, it cannot be used to answer Morgan's question : Why is a first person thought a thought about the thinker ? To answer that question in a satisfactory manner, we need to abstract from the linguistic expression of first person thought : we need a way of characterizing first person thoughts that is independent of issues of linguistic expression.

2

If we abstract from linguistic expression, what can we say about first person thoughts ? I said earlier that we cannot characterize such thoughts as thoughts that are about the thinker of the thought, for many thoughts will be about the thinker of the thought by accident (as in the case of the mirror). Still, there is something to the idea that first person thoughts are thoughts about the thinker of the thought. There actually are two senses in which a thought can be about the thinker : a superficial sense in which a thought can be about the thinker without being a first person thought, and a deeper sense in which it cannot. In the deeper sense, a first person thought *is* a thought about the thinker. A similar contrast arises in the language case : there too we have to draw a distinction between utterances that are accidentally about the speaker and utterances that are about the speaker in a more fundamental sense (first person utterances).

If the speaker happens to be the winner and he or she says, 'The winner will go to tahiti', her utterance is about herself but only in the accidental sense. This is not enough to make it a first person utterance. First person utterances are about the speaker in a more fundamental sense : it is a conventional rule of the language that an utterance of the first person is about the speaker of that utterance. The relevant convention governs the reference of 'I' : it is a conventional property of the word type 'I' that a token of that word refers to the speaker. So we need to distinguish two levels in the linguistic case : the linguistic meaning of the type (Kaplan's 'character') and the referential 'content' carried by a token of the type. A token of 'The winner will go to Tahiti' uttered by the winner will be about the winner at the token level but not at the type level. Clearly, the type expression 'the winner' is not such that its tokens are supposed to refer to the speaker. In the case of 'I', however, reflexivity – the fact that a token of 'I' refers to the speaker — *is* encoded in the meaning of the type and is not merely an accidental property of the token.

Can we say analogous things about first person thoughts ? Evidently, the notion of conventional meaning does not apply in the mental realm. As Morgan writes, 'Thoughts do not get their objects in virtue of being governed by conventions'. What we need is a property of the type that plays the same role as conventional meaning plays in the language case. Such a property must explain why the first person type in thought (the first person concept, as several philosophers call it) refers to the thinker of the thought, just as the conventional meaning of 'I' explains why a token of that word refers to the speaker.

The relevant property exists, I take it. As I argue at length in my last book (Recanati 2012), we can think of indexical concepts in general (and the first person concept in particular), as 'mental files' whose function is to store putative knowledge gained in virtue of standing in certain types of relation to the referent. The relevant relations are *epistemically rewarding* (ER) relation. An ER relation is a relation such that, when one stands in that relation to some object, one can gain knowledge about the object through the relation. The relation establishes a channel of information between the subject and the object. The suggestion, then, is that there are mental files which are based on such relations and whose role is to store the putative knowledge gained in virtue of standing in that relation to the object. What fixes the reference of the file (what the file is about) is not the information in the file — for that can be misinformation — but the relation : the file refers to the object to which the subject stands in the relevant ER relation.

ER relations can be described as relations that are conducive to knowledge, given the (contingent) cognitive equipment of the thinker. In general, we have ways of gaining

knowledge about individuals which depend upon our standing in the right relations to these individuals. Different ways correspond to different relations. First person ways of gaining information (through proprioception, kinaesthesis, introspection and so on) correspond to the relation in which one stands to an individual when one *is* that individual. So the relevant ER relation, in the first person case, is the relation of identity. That relation makes knowledge possible, given the cognitive equipment of the thinker. The knowledge one can get about an object in virtue of standing in the identity relation to that object is 'first person knowledge', or 'knowledge from inside' – the sort of knowledge one gains through proprioception and kinaesthesis. So I agree with Morgan's defense of Evans's epistemic approach to the reference of first person thoughts. The rule of reference – that the first person concept refers to the thinker – *follows from* the fact that an indexical file refers to the object to which the thinker stands in the ER relation which it is the function of the file to exploit. In the first person case the relation is identity, and the reference of the file is the thinker.

On this view, the type/token distinction applies to mental files. Mental files are typed according to the type of ER relation they exploit. Thus the SELF file (my name for the first person concept) exploits the relation to oneself (viz. identity) in virtue of which one can gain knowledge about oneself in a special way, 'from inside' — a way in which one can gain knowledge about no one else (as Frege puts it). My SELF file is not the same as yours, and they refer to different persons, of course, but they belong to the same type : they are both SELF files, unified by the common ER relation it is their function to exploit. We see that the *function* of files — namely, informational exploitation of the relevant ER relation — plays the same role as the conventional meaning of indexicals : through their functional role, mental file types map to types of ER relations, just as, through their linguistic meaning (their character), indexical types map to types of contextual relation between token and referent.

3

There is an important asymmetry between two types of first person thoughts : those which correspond to (putative) items of first person knowledge gained through the relevant epistemically rewarding relations to the referent ('I am hungry', 'My knees are crossed', when these thoughts are based on the appropriate first person experience), and those which do not even putatively correspond to items of first person knowledge ('I was born in Cork' — a first person thought corresponding to a piece of knowledge that can only be gained third-personally). The epistemic relation wich is conducive to first person knowledge in the first

type of case is what fixes the reference of the file, on my account, but the file *also* hosts information delivered through other avenues of knowledge than the special ways we have of gaining information about ourselves. (A hallmark of information about ourselves gained third-personally is that it is vulnerable to error through misidentification, while information gained first-personallly is immune to such error.)

Why is the SELF file hospitable to information gained in other ways than the first person way ? Because what Morgan calls the Informational Glue principle is constitutive of files. Two pieces of information are stored in the same file (and refer to the same object if they refer at all) if they are taken to be about the same object. In this way, first person thoughts which are not putative items of first person knowledge go into the same file as, and inherit their reference from, first person thoughts which are putative items of first person knowledge.

This account is similar to Morgan's in many ways. Like Morgan's account, it relies on Evans's epistemic approach to the reference of first person thoughts : the reference rule for the first concept is taken to be derivative rather than primitive. Like Morgan's account, it posits a basic asymmetry between two types of first person thoughts : first person thoughts in the second class inherit their reference from first person thoughts in the first class. As in Morgan's account, what makes it possible for first person thoughts in the second class to inherit their reference from those in the first class is the Inferential Glue principle.

What I am not sure I understand is the contrast Morgan draws between the inferential-glue explanation we both seem to accept and another explanation of how first person thoughts in the second class inherit their reference from first person thoughts in the first class. According to what he calls the Simple Rule Plus, the privileged cases (those which involve first person knowledge) fix the semantic value of the first person concept qua concept type : they make it possible to assign to that concept type a certain character, corresponding to the token-reflexive rule. So the first person concept refers to the thinker. That rule then explains why non-privileged cases of first person thought refer to the thinker of the thought.

I must confess that I don't clearly see the difference between the two accounts. Morgan writes : 'a version of the demonstrative model that incorporates the inferential glue explanation does not require us to accept as an unexplained primitive that all the tokens of the first-person type are tokens of the same semantic type'. But what *is* the first person type if it is not a semantic type ? What non semantic access do we have to the relevant type ? I simply don't understand what Morgan is talking about here. Throughout the paper Morgan seems to assume that there is such a thing as first person thoughts — a certain type of thought. As he

himself writes, ‘To be a first person thought is already to have a certains semantic profile’. His own account of the reference of such thoughts involves two ingredients : (i) an epistemic ingredient, i.e. the fact that in privileged cases a first person thought carries first person knowledge (knowledge from inside), (ii) a glue ingredient, i.e. a certain relation which holds between first person thoughts of all stripes, namely the fact that the subject takes them all to be about the same object. In my framework, these two same ingredients together amount to providing a semantics for the feature which all first person thoughts are supposed to have in common : the first person concept. In my framework, that concept is a mental file (this is what ‘glues’ all the deployments and makes it the case that, from the subject’s point of view, they are treated as referrring to the same object), and the reference of the file is fixed by the epistemic relations which it is the function of the file to exploit (and which are actually exploited in the privileged cases). This looks very much like Morgan’s account in that *the same two ingredients he appeals to are involved* ; but this is clearly a version of what he calls the Simple Rule Model Plus ! The difference, again, is that Morgan wants to avoid talk of a semantic type. But I don’t see what makes it possible for him to continue to assume that there is such a thing as first person thoughts. Does he simply mean token thoughts that are about ourselves (in the superficial sense)? But again, we don’t want to treat ‘the winner will go to Tahiti’ as a first person thought even if the thinker of the thought is the winner, and the only way to avoid that is by locating the relevant reflexivity at the type level. Morgan suggests that there is an alternative to that picture, but he hasn’t managed to make it clear enough for me to grasp it.

References

F. Recanati (2012) *Mental Files*. Oxford : Oxford University Press.