

The essential indexical research program.

Daniel Morgan
University of York
d.morgan@york.ac.uk

(forthcoming in Synthese)

Abstract: The ways of thinking of things associated with a few indexical expressions – e.g. ‘I’, ‘now’, ‘that’ – have a special role in the causation of action. They have a role not had by, for example, the guise associated with the ‘Superman’, or the guise associated with any other proper name. So, at least, an orthodox view about action – often associated with the phrase ‘essential indexical’ – has it. Recently, this view has come under scrutiny. An increasing number of philosophers think it is a myth. In this paper, I do two things. First, I argue *that* the orthodox view really is an insight, not a myth. Second, I offer an explanation of *why* it is that the insight holds.

I. Introduction

An orthodoxy about intentional action (hereafter, action) holds that there is a special connection between intentional action on the one hand, and, on the other, a small subset of attitudes – the attitudes that involve guises from a fairly short list.¹ A guise associated with the first-person pronoun, the guise one deploys when one thinks ‘*I* am being pursued by bear’ is a strong candidate for being on the list. So too is a guise associated with ‘now’, the guise one deploys when one thinks ‘The meeting starts *now*’. So too is a guise associated with ‘that’, the guise one deploys when one thinks ‘*That* is the sweater for me’. There are other strong candidates besides (e.g. a guise associated with ‘here’). But clearly not every guise can be on the list. If that were so, there wouldn’t a *special* connection between action and attitudes involving just the guises on the list.

¹ Particularly influential statements include: Castañeda (1968), Perry (1977), Perry (1979), Lewis (1979), Evans (1982), Peacocke (1983), Kaplan (1989) and Perry (1990).

Ultimately, a defender of this orthodoxy will want to do several different things. They will want to give a *full account* of what relation to action distinguishes list-members from non-list-members (and whether there are important distinctions to be drawn among list-members²). They will want to say *which* guises belong on the list and whether all the guises on the list have something interesting in common beyond being on the list – e.g. is there any sense of the word ‘indexical’ on which all and only indexical guises are on the list? Finally, they will want to say, of each list member, what explains the fact that it is on the list. These questions arise because of the form of the claim that orthodoxy is. For any claim of the form ‘A few X’s stand in a special relation to A’, there are analogous questions: what is the special relation; which few X’s; do those few X’s have anything *else* in common; and why do these guises stand in the relation to A that they do? If you accept the orthodox view, those are the natural next questions to ask. They define a research program.

Addressing these questions might represent wasted energy, analogous to working out wrinkles in a pre-Copernican astronomy at just the point when a better alternative became available. They would do if orthodoxy were a myth – an increasingly popular view.³ This paper has two aims. First, it argues that the orthodox view really is an insight, thus providing a rationale for the research project in the first place. Second, it makes a start on the final question mentioned above: that of how to explain, for each list member, *why* it is on the list. More than half of the paper focuses on the first aim. The second aim only has point if the first is accomplished. Someone might think the first aim can be accomplished easily, perhaps just by gesturing at a few famous thought-experiments involving words like ‘I’ and ‘now’. But I think that critics

² See Morgan (2019a) for discussion.

³ Cappelen and Dever (2013)’s book length treatment is the most sustained attempt to show this, and the one I will focus on in some detail. Other works that take a skeptical approach to orthodoxy include Ball (2015), Devitt (2013), Douven (2013), Magidor (2015), and Spencer (2007).

have by now done enough to show that whether these thought-experiments really support the orthodoxy is far from obvious, so the right to take part in the research program needs to be earned.

I focus, in the second part of the paper, on the ‘why is this guise on the list?’ question, rather than any of the other questions that define the research program, for three reasons. First, I have something new to say about it. Second, as I explain below, a doubt that the ‘why’ question can be answered can be converted into a doubt about the whole research program. Third, the strategy for answering this question I defend looks to be in tension with Millikan’s work on this topic. Anyone pursuing the strategy needs to say how they want to reply to Millikan.

The strongest intuitions we have relevant to the orthodox view concern thought experiments involving specific cases of action, or failures to act. So, the focus for an argument in favor of the orthodox view will be on a specific such case. I use an ‘T’ case, but I think it should be clear that there is potential for generalization to ‘now’, ‘that’, ‘here’ etc.

Here is the complete argument in the case of ‘T’:

(P1) *T-Needed*: there are *some* actions that an agent needs an attitude involving the guise associated with ‘T’ to perform.

(P2) *Contrast*: for any guise, *g*, associated with a proper name there are *no* actions that an agent needs an attitude involving *g* to perform.

(C) *T-Orthodoxy*: There is something *special* about the role the guise associated with ‘T’

has in the causation of action, i.e. something that doesn't hold for the guises associated with referring expressions across the board.

Section II provides further introduction to the orthodox picture. It explains what makes it natural to think of *T'-Orthodoxy* as the core minimal commitment of authors like Castañeda, Perry and Lewis in relation to 'P'. It also makes the case for *Contrast*. Section III makes the case for *T'-Needed*. That completes the defense of *T'-Orthodoxy*. In Section IV, I sketch a strategy for *explaining why* the guise associated with 'P' has the special role in action that *T'-Orthodoxy* says it does.

The remainder of this introduction offers some clarifications about key terms used in the argument I've just set out: 'action', 'associated with', and 'guise'.

By 'action' I mean action type, rather than action token. More particularly, I mean *robust* action type. An action type is *robust* if its tokens have something in common with each other, that isn't common to the tokens of any different action type, and that goes beyond their etiology involving a certain guise. *Addressing Superman* is an example of a robust action type. What its tokens have in common with each other is that they all involve addressing someone and, in fact, the same person – Superman, aka Clark. This isn't common to the tokens of any different action type and has nothing to do with any guise. Consider, by contrast, the type: *addressing Superman, thought of as such*. Its tokens have a great deal in common with each other. They all involve addressing someone and, in fact, the same person. But what do they have in common with each other that they don't also have in common with the tokens of the different type: *addressing Clark, thought of as such*? Nothing, I submit, beyond the fact that their etiology involves

the guise associated with ‘Superman’. So, *addressing Superman, thought of as such* is not a robust action type. Performing it depends on using the guise associated with ‘Superman’, but that dependence is trivial. It is an artefact of how the type has been constructed. We don’t want our centerpiece claim, *T’-Needed*, to be trivial. This is avoided by requiring that any witness for *T’-Needed* be a robust action type. If there is an action that is a witness for *T’-Needed* then performing it will indeed depend on the guise associated with ‘T’. But that dependence won’t be a trivial artefact of the guise associated with ‘T’ being constructed into the action.

Having provide some initial sign-posting relevant to ‘action’, I switch to the phrase ‘the guise that is associated with “T”’, which appears both in *T’-Needed* and *T’-Orthodoxy*.

The guise is *associated with T’* in the sense that its *direct expression in language* is, in English, ‘T’ or, in another language, a first-person pronoun in that language. The point of the ‘direct’ qualification is to screen out the following kind of case: a subject may think of something using one guise, whose direct expression is a particular term, but prefer to use a different term they know to refer to the same thing as a way of expressing the state they are in. De Gaulle utters ‘De Gaulle is in pain’. He may nevertheless be in a state involving a guise whose direct expression is ‘T’. Even if he is stably disposed to only ever use ‘De Gaulle’ to self-refer, he may nevertheless often be in states involving a guise whose direct expression is ‘T’. For he may be stably disposed to avoid the word that directly expresses the guise of him some of his states involve. The test for what word directly expresses a guise he uses in thought is what word would be used by him on an occasion when he expresses a state involving that guise and his choice of language does not exploit any identity knowledge he has concerning the thing the guise refers to. The only examples of such occasions may all be counterfactual, if his aversion

to a particular term is stable enough.

What, in general, is a *guise*? It is something that is individuated according to Frege's criterion of difference. If a rational subject can simultaneously believe '*a* is F' and disbelieve '*b* is F', then '*a*' and '*b*' are associated with different guises, even if they have the same reference. Any other term – e.g. 'mode of presentation' / 'concept' / 'way of thinking', 'mental file', 'mental word' – stipulated to obey the same criterion of difference would have served as well as 'guise'. Someone who prefers to avoid talk of such entities altogether can use a different framing of the topic – e.g. '*de se* mental states', where a *de se* mental state is one whose most direct expression in English is the word 'I'.

II. Why '*I*-Orthodoxy really is the orthodox commitment about 'I' and why *Contrast* is true.

So called 'essential indexical cases' were the most striking and memorable thing about the papers that advanced the orthodox picture. Here is the essential indexical case that will be the focus of my discussion:

The Bear Case. Walking through the woods, François has the desire that *he himself* not be eaten by a bear (i.e. a desire whose direct expression in language would involve the first-person). As a result of an accurate but less than fully helpful conversation with a fellow walker, François has learned that *François is being pursued by a bear that is to François's left*. He believes that *if François remains where he is, François will be eaten*. He does nothing. Ultimately however, he realizes that *he himself is François* and that *he himself is being pursued*

by a bear to his left and if he remains where he is he will be eaten. Now he flees. The change in belief seems to explain the change in behavior.⁴

Writers like Castañeda, Lewis and Perry treated essential indexical cases like this as establishing something *new* about the causation of intentional action.

What is the import of ‘new’ here? Quite a while before essential indexical cases entered the philosophical stage, Frege had already made opacity cases involving co-referential names highly salient. It is pretty clear, once one has seen such cases, that the difference between the guises associated with different co-referential names can be of relevance in action explanation. For example, the difference between guises expressed by different names is relevant in the following case offered by Cappelen and Dever (hereafter, C & D):

Superman/Clark Kent. Pushing my cart down the aisle I was looking for CK to tell him he was making a mess. I kept passing by Superman, but couldn’t find CK. Finally, I realized, Superman was CK. I believed at the outset that CK was making a mess. And I was right. But I didn’t believe that Superman was making a mess. That seems to be something that I came to believe. And when I came to believe that, I stopped looking around and I told Superman to clean up after himself. My change in beliefs seems to explain my change in behavior. (2013: 33)

Writers like Castañeda, Perry and Lewis treated essential indexical cases as showing something

⁴ The case is an adapted version of the famous case involving a bear introduced by Perry at (1977: 494).

new despite writing in a context in which cases involving informative identities between co-referring names were highly salient. So it is reasonable to regard them as committed to thinking that essential indexical cases establish something that name cases do *not* establish. It is reasonable to regard them as committed to the following contrastive claim:

There is something *special* about the role the guises on a short list have in the causation of action, something that doesn't hold for referring expressions across the board.

T'-Orthodoxy is the commitment these writers have in relation to the guise associated with 'T', granted that they think that it is on the short list:

T'-Orthodoxy: There is something *special* about the role the guise associated with 'T' has in the causation of action, something that doesn't hold for referring expressions across the board.

Why C&D are right that the opacity based case for T'-Orthodoxy fails:

A natural way of undermining 'T'-Orthodoxy is to say what the upshot of essential indexical cases is, and to show that that upshot isn't special to the guises on a short list. This is what C & D aim to do. According to them, the upshot of essential indexical cases is what they call 'Indexical Opacity':

'There's a set of indexicals, I-SET, that cannot be substituted *salva veritate* in action-explanation contexts by any other expressions' (2013: 33).

But, as they point out, this is just an instance of the more general thesis they call ‘Generic Opacity’:

‘Co-referential referring expressions cannot be substituted *salva veritate* in action explanation contexts’ (2013: 33)

Their intention, in bringing up **Superman/Clark Kent**, is to show that Generic Opacity is a plausible thesis.

C & D’s line of thought is convincing as far as it goes – it shows that Perry failed to formulate theses that support the orthodox picture. Perry says that, for some explanations of action, indexicals can be essential ‘in that replacement of them by other terms destroys the force of the explanation’ (1979:4). To this, a convincing response is that the force of an explanation can *also* be destroyed if one substitutes one non-indexical for another non-indexical (e.g. ‘Superman’ for ‘CK’). So, something Perry implicitly treats as particular to indexicals is in fact not.⁵

C & D’s aim, however, is more ambitious than arguing that Perry’s candidate for being a thesis that essential indexical cases support and that in turn supports the orthodox picture. It is to argue that no such thesis exist. So, a defender of the orthodox picture can always come along and offer a different candidate for what makes the guises on the short list special. This paper does that, arguing that we should accept *T’-Orthodoxy* because it follows from *T’-Needed* and *Contrast*. The next section motivates *T’-Needed*. This section motivates *Contrast*:

⁵ See Babb (2016: 440) for further discussion.

Contrast: for any guise, g , associated with a proper name there are *no* actions that an agent needs an attitude involving g to perform.

As noted above, ‘action’ in this paper is restricted to robust action-types. In **Superman/Clark Kent**, Herman *addresses Clark, thought of as such*. His doing that very thing depends on him thinking of Clark ‘as Clark’. But that very thing is not a robust action type. He also *addresses Clark*. That is a robust action type. But performing it does not require Herman to think of Clark ‘as Clark’. He could have performed that action even if he’d never learned the name ‘Clark’. So, **Superman/Clark Kent** does not provide a counterexample to *Contrast*.

Noting that is not the same as offering a positive argument for *Contrast*. But *Contrast* should be considered innocent until proven guilty. There is never any stipulative connection between an action, i.e. a robust action, and a guise. If one thinks even so that for some particular action a given guise is needed, one needs to say what evidence there is for that. We can, I’ll argue in the next section, do this in relation to the guise associated with ‘P’. Here, I invite the reader to see if there is any guise associated with a proper name in relation to which they can do it. If they can’t, they should accept *Contrast*.

III *T’-Needed*

Here again is *T’-Needed*:

There are some actions that an agent needs an attitude involving the guise associated

with 'I' to perform (*T'-Needed*).

T'-Needed is an existential generalization. To argue for it we need to show that it has a witness. I argue that *fleeing a bear* is a witness. The argument focuses on **The Bear Case**. To recap, François has the testimony of a fellow walker that 'François is being pursued by a bear that is to François's left'. But François does not know that *he* is the one being spoken about, and he does not flee.

Ignore issues about the guise associated with 'I' aside for a moment. There are several other things that could get in the way of François's fleeing. I'll add extra details to the base version of the case over a number of versions. The end point version of the case is one about which it's very natural to have the reaction that the problem keeping François from fleeing can *only* be that he is not thinking of himself 'as I'. If this reaction is correct, then fleeing a bear really is a witness for *T'-Needed*. To strengthen the appearance that the reaction is correct, I'll discuss C & D's account of action, which seeks to identify a different candidate as the source of the blockage that could apply even to the end point version of the case. I claim that this different candidate is illusory. The shape of my argument is as follows: the account C & D endorse is rather underspecified. There are a number of features that are not explicitly built into it but that it needs to remain plausible. But the explanation of why François fails to flee that results, once we add these features, is not an alternative to the intuitive explanation, according to which the problem is the absence of the guise associated with 'I'. So, on the only plausible version of the account, the explanation it offers of why François does not flee in the end point version of the case is the same as the intuitive explanation. So, fleeing a bear is a witness for *T'-Needed*, just as it initially seemed to be. *Contrast* is also true. So, *T'-Orthodoxy* is true.

I begin, however, by refining **The Bear Case** so that we get up to the end point version of the case.

The problem in **The Bear Case** could be similar to the problem in **Superman/Clark Kent** – the agent has a belief and a desire that pick out the same person using *two different guises*. As a result, the belief and the desire don't gel together in the right kind of way to yield action. The analogous thought about **The Bear Case** would be that the guise involved in François's desire (the desire that *he* not be eaten by a bear, which he would express using 'I') and the guise involved in his belief (the belief that *François* is being pursued by a bear to *François's* left, which he would express using 'François') are different, and that's why action does not occur.

If the difference in guise were the only other potential problem though, it ought not to be possible to get a version of case where the guise involved in the belief and desire is the *same*. However, the following case is also possible:

The Bear Case*. Walking through the woods, François has the desire that *François* not be eaten by a bear. François has learned that *François* is being pursued by a bear that is to *François's* left, and believes that if *François* remains where he is *François* will be eaten. Even so, François fails to flee.

In **The Bear Case***, the problem can't be that François's belief and desire involve two different guises of François. Each of them involves the same guise. Even so, the problem need not have anything to do with 'I'. Whether appropriate action occurs typically depends on

conditions that are downstream of the subject making an attempt to do something. E.g. whether fleeing occurs depends on François not being a victim of sudden paralysis.

If that were the only other potential problem though, it ought not to be possible to get a version of the case where all the relevant conditions downstream of François making an attempt are satisfied and action still does not occur. However, the following case is also possible:

The Bear Case.** Walking through the woods, François has the desire that *François* not be eaten by a bear. François has learned that *François* is being pursued by a bear that is to *François's* left and believes that if *François* remains where he is *François* will be eaten. François fails to flee – *and is not even motivated to attempt to.*

This is the end point version of the case, the crucial version for my purposes. It is set up in such a way as to indicate that the problem is neither of the two that have already been considered. What could the remaining problem be? It's intuitive to think that the only problem left is that François is not thinking about himself 'as I'. By this I mean the following. First, if one is told: François does not know 'I am François', that does intuitively seem enough to explain why he hasn't fled yet. Second, there isn't *anything else* that intuitively strikes us as an alternative equally good explanation of why he hasn't fled yet. For example, suppose one is told: François does not know 'Recanati is François'. Our intuitive reaction to learning this is not that it too seems enough to explain why he hasn't fled yet. Now this disparity between 'I' and, e.g., 'Recanati' could turn out to be an illusion. Perhaps there is a possible etiology for François's fleeing that involves him thinking about himself 'as Recanati' and not 'as I'. What

I submit at this stage is just that our intuitive reaction is the one I've set out: what explains his not fleeing can only be that he isn't thinking of himself 'as I'. If that reaction is right, then fleeing a bear is a witness for *T-Needed*.

Is there a different explanation?

C & D want to provide a different explanation of what the problem could be in a case **The Bear Case****. They offer the following 'action inventory' model:

Every agent has a very wide range of third-person beliefs and desires that give rise to third-person intentions, which in turn rationalize or motivate actions (via their recognition). Not all of these intentions are going to produce action, at least in normal cases (perhaps in a god they would). This is because a given agent has an "action inventory": a range of actions that he can perform. An agent constantly seeks to match his intentions with his action inventory, and when he finds a match, action occurs. When there's no match, the intention idles, and doesn't motivate or rationalize action. (2013: 50)

Can this model, on its own or supplemented with further premises, block the case for *T-Needed*?

Even on the action inventory model, it seems true that there is a certain guise of François that is needed if François is to flee – namely, the guise of him that has the property of figuring in an intention that is 'matched in his inventory'. The only way of resisting this is to say that

whether an intention is matched in an agent's inventory cannot be sensitive to what guises it involves. But that rules out the following evident possibility. François has an intention that does motivate some action and so, in terms of the model, is 'matched in his inventory'. Now someone comes along and introduces François to a new term that refers to something that his original intention refers to. They tell François enough about the referent of this term that François forms a second intention whose satisfaction requires exactly what his original intention required. They don't tell him enough to put him in a position to act on that intention. The first of these intentions is matched, and the second is not. But the only difference between them is the guise they involve. So, the action inventory model should acknowledge that whether an intention is matched can be guise-sensitive. So, there is one sense in which it's true, even on this model, that there is a certain guise of François that François needs if he is to flee – namely, the guise of him *that has the property of figuring in an intention that is matched in his inventory*. How much logical space is there between this claim and *T'-Needed?* I think there are two gaps. But these can be closed, as the next two subsections show.

(a) *De dicto and de re truths about guises*

Suppose it is conceded that, even on the action inventory model, François needs to be deploying the guise of himself *that figures in an intention that is matched in his action inventory* if he is to flee. Still, it doesn't follow from that that there is some particular guise of himself such that François needs *it* if he is to flee. The first claim, a claim of *de dicto* necessity, does not entail the second, a claim of *de re* necessity. It is a contingent fact about François that his action-inventory is constituted as it is. If his action inventory had been different, then an intention's matching with the inventory could have required that it involve a *different* guise. François needs *whatever*

guise has the property of being matched in his inventory. But there is no particular guise such that François needs *it*.

A defender of *T'-Needed* can accept this. *T'-Needed* should be read as a claim of *de dicto* necessity. There are reasons that have nothing to do with the action inventory model to read the claim this way. Suppose one is a functionalist and one thinks of a guise as a kind of physical particular that, in virtue of playing a certain functional role, counts as having various mental properties. Then it won't be a *de re* necessary truth about a particular guise even that *it* is a guise. *A fortiori*, it won't be a *de re* necessary truth that *it* is a guise that is needed to flee a bear. *T'-Needed* is the *de dicto* claim that there are some actions that an agent needs an attitude involving *the guise that has the property of having T' as its direct expression* to perform. This claim is strong enough to vindicate *T'-Orthodoxy*. That's because *Contrast* doesn't become any less plausible on a *de dicto* reading. It isn't true, for example, that there are some actions that an agent needs the guise *that has the property of having 'Superman' as its direct expression* to perform.

(b) *A guise whose direct expression is T'*

Why doesn't François flee? The problem, intuitively, is that he isn't thinking of himself 'as I'. The problem, in terms of the action inventory model, is that his 'François'-intention is not matched in his action inventory. These two explanations aren't rivals, if it turns out that the intention that would be matched in the action inventory, would also have 'I' as its direct expression. So, an opponent of *T'-Needed* will need to resist adding that feature to the model. So, they ought to be able to describe a possible case in which François's matched intention is *not* one whose direct expression is 'I'.

C&D make an interesting suggestion that might seem to provide a recipe for doing this. Here it is: not all intentions that are matched in an agent's action inventory need involve *self-reference* by the agent. For example, Herman closes a door because he desires that someone else, Nora, be safe, and believes that if the door is closed, she will be safe (see 2013: 37). There will have to be some specific means by which Herman brings the desired result about – e.g. perhaps *he walks over to the door, grabs the knob and turns his hand*. But C&D deny that making this selection of means requires Herman to form a self-referential attitude (e.g. an intention *to walk over to the door, grab the knob, and turn his hand*). They insist it could be achieved subpersonally, so that the only agent-directed intention Herman has is Nora-directed but not self-directed. He doesn't intend 'I walk over to close the door' or 'I save Nora' or 'I bring it about that Nora is safe'. He intends 'Nora is safe'. Then he executes.

Is this possible? It seems to me that it might well be. If possible, the case immediately rebuts an argument for *T'-Orthodoxy* that focuses on the claim that *every* action depends on the guise associated with 'T'.⁶ It doesn't immediately rebut the current case, which is focused on *fleeing a bear*. But it might appear that if we can construct a possible etiology for *bringing it about that Nora is safe* not involving the guise whose direct expression is 'T', a tweaked version of the same recipe will enable us to construct a possible etiology for *fleeing a bear* not involving the guise whose direct expression is 'T'. This tweak isn't offered by C&D, but it does seem to me the best way for an opponent of *T'-Needed* to go.

⁶ Babb (2016) and Bermudez (2016) defend the orthodox view against C&D, and in each case the defence rests on the claim that *every* action depends on thinking of oneself 'as I'. This commits them to saying that Herman saving Nora without any self-directed intention is impossible. But neither provides a defence of this verdict.

Saving another as a model for saving oneself

Herman can act so as to bring it about that Nora is safe. In doing so, let's suppose, he need have no self-referential attitude, *a fortiori* no attitude involving the guise whose direct expression is 'I'. It's just a fact about Herman's action inventory that Nora's safety is something he can bring about directly. Here is the tweak: if that is true of Herman, then it presumably could be true of Nora herself. Identity is not a magic forcefield that prevents ways of saving things that are generally applicable from being used to act on the thing that the agent is identical to. So someone might argue as follows:

François's action inventory could be constituted so that his being a safe distance from a bear that is threatening him is something that he can bring about just by forming an intention with the content 'Recanati is a safe distance from a bear that is threatening him'. 'Recanati', unlike 'François', is the direct expression of a guise of him that is matched in his inventory. That the one guise is more motivationally potent than the other is just a fact about him and his inventory. Why, in that case, does François not flee in **The Bear Case****? Because he is thinking of himself in 'the wrong way'. But we can't move from that to the claim that he needs to be thinking himself 'as I'. In the case we've imagined, it would be fine if only François started thinking of himself 'as Recanati'. Once he did that, he'd apply some generally applicable method for bringing it about that someone is safe to himself and thereby save himself from the bear.

This, it seems to me, is a possible case in which François acts without thinking about himself

‘as I’ and, in particular, acts in such a way that he brings out the very result, his own safety, that he also brings about when he successfully flees a bear. What I deny is that it is a case in which he flees. I invite the reader to think themselves into the case and consult their ‘does he flee?’ intuitions. I note, as a prompt, that googling ‘unknowingly fleeing’/ ‘unintentionally fleeing’/ ‘fleeing but not as such’ returns nothing relevant. A natural view to take about why this isn’t a case of fleeing is that fleeing, like other actions, isn’t just individuated by its end-result but also by the kind of psychological process that leads up to that end-result. But the intuition that this is not a case of fleeing is independent of any particular explanation of why it isn’t.

Suppose, for the moment, that it’s conceded that this case is not one in which François flees. Reflecting on the question of why it isn’t might suggest a new strategy for an opponent of *T’-Needed*. Perhaps they should concede that fleeing a bear depends on thinking of oneself ‘as I’, but deny that fleeing a bear is a witness for *T’-Needed* on different grounds. They could say that *fleeing a bear* turns out not to be a *robust* action. Perhaps we refuse to count the imagined case as one in which François flees purely because he isn’t thinking about himself ‘as I’.

This is a reasonable objection, but it can be answered. *Addressing Superman, thought of as Superman* is not an example of a robust action type because tokens of *addressing Superman, thought of as Superman* have nothing in common with each other, that they don’t have in common with the tokens of *addressing Superman, thought of as Clark*, beyond the fact of their etiology involving the guise associated with ‘Superman’. But tokens of *fleeing a bear* seem to have plenty of things in common with each other, that they don’t have in common with the tokens of *bringing it about that one is safe from a bear, when one conceives of oneself just using a proper name*, and that go beyond the

fact of their etiology involving the guise associated with 'T'. I'll focus on three of them.

(i) When François *flees a bear*, his nervous system is disposed to be highly aroused. The connection between flight and the nervous system is not some contentious philosophical doctrine. It is one of the first things one learns about flight (as in, 'fight or flight'). By contrast, when he *brings it about that he is safe from a bear that is threatening him, conceiving of himself just using a proper name*, his nervous system is not disposed to be aroused (or, at least, the degree to which it is depends on the degree of his empathy). (ii) When François *flees a bear* he acts on an intention whose *predicative* component involves an action verb: 'flee'. By contrast, when he *brings it about that he is safe from a bear that is threatening him, conceiving of himself just using a proper name*, the intention has as its object a non-agential fact: 'Recanati is safe' (to be safe is not, as such, to do anything). (iii) When François *flees a bear* what an onlooker ordinarily observes is something like this: an agent catches sight of a bear and runs. When François *brings it about that he is safe from a bear that is threatening him, when he conceives of himself just using a proper name* what an onlooker observes can't be anything like this. What they observe is an agent deploying some generally applicable method which, if the onlooker follows the causal chain downstream for long enough, emerges as being deployed in this instance to act on the agent themselves. These contrasts, assuming they are conceded, give substance to the claim that tokens of *fleeing a bear* have something in common with each other, that they don't have in common with the tokens of that other action type, and that goes beyond their etiology involving the guise associated with 'T'. They rebut the worry that *fleeing a bear* is not a robust action.

Perhaps the opponent of *T'-Needed* should not concede the three contrasts highlighted in the previous paragraph. Perhaps they should insist that there could be a case in which François

acts on a ‘Recanati’-intention that resembles an uncontentious case of fleeing not just with respect to *end result* but also with respect to its *psychological process*. Here it is:

François’s intention is not of the form ‘x is safe’, but of the form ‘x flees’. As François executes this intention, he exhibits a healthy degree of nervous system arousal (and if there is some part of his brain normally associated with altruistic behavior that part of his brain is inactive). Finally, what an onlooker observing François would see is an agent apparently catching sight of a bear and running, not his closing a door or engaging the services of rescuers or anything like that. Finally, however, the direct expression of the singular component of his intention is not ‘P’ but ‘Recanati’.⁷

The first thing to notice about this case is that the Herman-saves-Nora case doesn’t do anything to support its possibility. To make this case more resemble one in which François actually does some fleeing, it’s been given features that the Herman-saves-Nora case in principle could not have. For example, even if Herman’s action inventory is so constituted that ‘Nora is safe’ is an intention he can act on directly, it can’t be so constituted that ‘Nora flees’ is an intention *he* can act on directly. The opponent of *I²-Needed* who appeals to the case just set out has abandoned the ‘saving oneself on the model of saving another’ approach. They’ve gone for a more brute force approach, stipulating a case that has everything an uncontentious example of fleeing a bear has other than an etiology involving the guise whose direct expression is ‘P’. The risk with this approach is that the stipulated case may not be a

⁷ This needs to be a scenario in which ‘Recanati’ remains a proper name, the kind of word which, if Kripke’s picture is right, has its reference fixed by a causal-historical chain. It is, of course, irrelevant to *I²-Needed* that there could be a language whose first-person pronoun is pronounced like the French and Italian surname ‘Recanati’.

possible case.

The reason to think it *impossible* is the original intuitive reaction to **The Bear Case****. Our intuitive reaction about that case was that the problem can only be that François isn't thinking of himself 'as I'. It wasn't that there is some particular way of thinking of him that François needs to be using but what the direct expression of this other way of thinking just isn't something we cannot know in advance – it could be 'I', it could equally be 'Recanati'. Now, of course, the intuitive reaction is not sacrosanct. If the 'saving oneself on the model of saving another' approach to constructing a case of fleeing without the guise whose direct expression is 'I' had worked out, it might have been suggested that the case thus constructed is rather unusual and complicated and just the kind of possibility that our intuitions are apt to overlook and, in any case, there is positive reason to believe in the possibility of that case. But that recipe did not work out. The alternative brute force approach yields a case that the intuitive reaction is precisely not neutral on the possibility of. If we let the two cases blur, of course, it may seem as though there is a single case that clearly has the right kind of psychological process to be an instance of fleeing a bear and that our intuition is neutral on the possibility of. Once the cases have been distinguished, and it has been clarified that the relevant case is the second case, an opponent of *T'-Needed* must say what reason there is to think that case possible, despite its intuitively seeming not to be.

I close the defense of *T'-Needed* by considering two proposals for what that reason could be.

Proposal 1: 'Action is fundamentally about the causation of bodily movements by mental states. Metaphysically, anything can cause anything. So, of course, a "Recanati flees"

intention could in principle cause bodily movements that constitute François fleeing (and cause nervous system arousal etc). That could even happen regularly. And when it does that would be an instance of François fleeing on the basis of a “Recanati flees”-intention.’

Here’s where I suspect Proposal 1 goes wrong: some version of functionalism is true and the way mental properties supervene on physical ones makes it impossible for a particular physical state both to regularly directly cause movements that constitute François fleeing and to have a content whose direct expression is ‘Recanati flees’. But, whether or not that’s right, the ‘anything can cause anything’ line of argument does nothing to support the possibility of François fleeing on the basis of a ‘Recanati flees’-intention. A way of seeing this is to note that the line of argument makes no play at all with the fact a ‘Recanati flees’ intention at least has something in common with an ‘I flee’ intention, had by François. If ‘anything can cause anything’, then a belief whose direct expression is ‘ $2+2=4$ ’ could equally regularly cause bodily movements that constitute François fleeing. So, unless we accept that there is possible case of François fleeing where the only mental state involved in explaining his flight is that belief, we should reject Proposal 1.

Here is the more challenging objection, which brings us to the business of the next section:

Proposal 2: ‘A connection between a linguistic action, like uttering the word “bear”, and a certain word, the word “bear”, seems readily intelligible. But why should there be any connection between an action as supremely non-linguistic as fleeing a bear and one word in particular: “I”? It would be amazing if such an action-word connection were a brute

fact, admitting of no further explanation. So, to the extent that we are unable to give any further explanation of it, we have reason to doubt whether it is a fact. We can't give any further explanation of it. This is the reason for thinking that François fleeing through executing a "Recanati flees" intention must be possible, as happens in the second case described above, even if our intuitive reaction to **The Bear Case**** seems to rule that possibility out.'

I mentioned in the introduction of the paper that, in addition to defending the essential indexical research program, I wanted to pursue one of the questions that arises within it: 'why is each guise that's on the list of special guises on the list?'. I've now argued that the guise whose direct expression is 'T' is on the list because that guise is needed if one is flee a bear. So, as applied to this guise in particular, the question becomes: why is this guise needed to flee a bear? To the extent that it's implausible that such a connection could be a brute connection, answering this question isn't just an explanatory question that we would want to answer even if the healthiness of the essential indexical research program were clear. It is that. But it is also a question that needs to be answered if we are to rebut the worry Proposal 2 articulates about the argument for *T'-Needed*.

My response to Proposal 2 is to deny that we have to accept the word-action connection as brute. The next section sets out the following strategy for explaining it: the fact that the guise under discussion has the word 'T' as its direct expression in language is a reflection of its having a (not extrinsic, not linguistic) mental property, and it becomes more readily intelligible that the guise is needed for certain actions, like fleeing a bear, once we see that it also has this mental property. Ruth Millikan's work on this topic might seem to rule out this explanatory

strategy. She appears to argue that we can't read back from what words certain guises have as their direct expression to what the guises are like in themselves. The next section therefore also says how I think Millikan's worry should be answered.

IV. The nature of the guise whose linguistic clothing is 'I'

Millikan accepts that some self-directed states are immediately motivationally potent in a way that other self-directed states are not, and the difference corresponds to whether the direct expression of the state is 'I'. She conceives of the former states as involving what she calls the agent's 'active self-name'. Millikan abbreviates 'active' to '@'. So Millikan's own active self-name, for example, is @ "RM". She writes:

My mental "I", my @ "RM", is not an indexical. More reasonably it is a (Millian) name for me, your "I", which may well have quite a different mental shape, is a (Millian) name for you. There is a special connection between @ "RM" and my use of the public language indexical "I", not because @ "RM" is indexical, but because in knowing how to effect the action of the referent of @ "RM" I know, among other things, how to effect her public self-reference (1990:732).

For my purposes, the final sentence here is the one to focus on. In it, Millikan does three things. First, she acknowledges a datum – there is a special connection between @RM and Millikan's use of 'I'.⁸ Second, she resists a possible explanation of that datum – that @RM is an indexical. Third, she offers her own explanation – the datum is explained by the fact that

⁸I drop the quote-marks around @RM as there is no risk of use/mention confusion.

Millikan knows how to effect the action of the referent of @RM and, simply as an aspect of this (“among other things”), knows how to effect her public self-reference.

How are we to understand the sentence: “There is a special connection between @RM and my use of the public language indexical ‘T’”?

‘There is a special connection between Barcelona and Messi’ has two readings. It could mean that Messi is connected to Barcelona in a way that he isn’t to any other club – e.g. he stands in the *being an employee of* relation to Barcelona and to no other club (not e.g. Madrid). Or it could mean that Messi is connected to Barcelona in a way that no other player is connected to Barcelona – e.g. Messi stands in the *has scored over 400 goals for* relation to Barcelona and no other player does (not e.g. Coutinho).

‘There is a special connection between @RM and Ruth’s use of ‘T’” likewise has two readings. The first reading is that @RM stands in a relation to ‘T’ that *no other guise* of Ruth stands in to ‘T’. The second reading is that @RM stands in a relation to ‘T’ that @RM stands in to *no other word*. The sentence is true, on both readings. So we have two bits of data to explain.

The truth of the sentence, on the first reading, corresponds to a datum that Millikan has an elegant, deflationary explanation of. Here is a thing that could happen. Millikan could fail to *utter the word ‘T’* because although she desires, of herself, that she utter ‘T’, she does not, in having that desire, think of herself in the right way, i.e. using @RM. Perhaps she desires that the author of *Biosemanitics* utter ‘T’, but doesn’t know who that is. Here though is another thing that could happen. Millikan could fail to *utter the word ‘Ruth’* (or ‘cat’) because although she

desires, of herself, that she utter ‘Ruth’ (or ‘cat’) she does not think of herself in the right way, i.e. using @RM. Here is yet another thing that could happen. Millikan could fail to *flee a bear* because although she desires of herself, that she *flee a bear*, she does think of herself in the right way, i.e. using @RM. The point about @RM being connected to Ruth’s use of ‘T’ in a way that *no other guise* of Ruth is connected to Ruth’s use of ‘T’ really is just an aspect of the broader point that @RM is connected to many of Ruth’s actions, linguistic and non-linguistic, in a way that *no other guise* of Ruth is.

What about the second reading of the sentence – that @RM stands in a relation to ‘T’ that @RM stands in to *no other word*? The sentence is also true, on this reading. Ruth’s uttering ‘T’ (rather than her uttering ‘Ruth’ or ‘cat’) is the direct expression of @RM in language. Millikan seems to acknowledge as much – in the passage I’ve quoted, she also refers to @RM as ‘my mental “T”’; she never refers to it as ‘my mental “Ruth”’, or ‘my mental “cat”’. But precisely because what her explanation appeals to is a point about @RM’s involvement with Millikan’s actions *in general*, it does nothing at all to explain why the action of *uttering one word in particular* – the word ‘T’ – is the direct expression of @RM in language.

Only equivocation on the two readings of ‘There is a connection...’ could make it appear as though Millikan has already provided an explanation of this datum. Since she hasn’t, we need some different explanation of it. So there is the possibility of offering an inference to the best explanation argument that has this datum as its starting point and has as its conclusion a claim about the nature of @RM. That is, we should argue as follows:

- (1) @RM is associated with Millikan’s use of ‘T’ in a way that it isn’t with her use of *any*

other word.

(2) The best explanation of that is that @RM has something *in common* with the word 'I' that it doesn't have in common with any other word.

(3) The obvious thing they could have in common is *how they refer to the thing they refer to.*

(4) The word 'I' refers to the thing it refers to because it is governed by a rule that *any token of that word refers to the one using it to refer.*

(*Reflexivity*) So, @RM must also refer to Ruth because she is the one using it to refer. Similarly for anyone else's active self-name.

This is not a knock-down argument for (*Reflexivity*). Each of the steps could be resisted. Still, the argument has some plausibility. What happens if we accept (*Reflexivity*)?

One consequence is that we go from a situation where we know *two* things about the guise we are interested in:

(A) it plays a special role in at least some of Millikan's actions (e.g. fleeing a bear)

(B) its direct expression in language is the first-person pronoun

to knowing *three* things about it:

- (A) it plays a special role in at least some of Millikan's actions (e.g. fleeing a bear)
- (B) its direct expression in language is the first-person pronoun
- (C) it refers to Millikan because she is the one using it to refer

If *T-Needed* is true, then the guise that has (A) also has (B). The worry we considered was that it shouldn't be a brute fact that the guise that has (A) also has (B). The strategy I favor says that the reason the guise that has (A) also has (B) is that the guise that has (A) also has some mental property, and it's more readily intelligible that the guise that has this mental property has (B). The mental property is: (C). Someone who has met Millikan's work on this topic might worry how we can know that the guise that has (A) has (C). Perhaps the linguistic clothing the guise wears gives a misleading impression of its underlying nature! My argument above is not vulnerable to that worry. Millikan's passage does not supply a good reason for resisting the argument, once the two readings of the sentence of hers I've discussed are distinguished. In particular, it doesn't once we clarify that the reading of the sentence that is the first premise of the argument is the *second* reading – @RM stands in a relation to 'T' that @RM stands in to *no other word*. This is a datum of which Millikan has no explanation, and that looks to be well explained by *Reflexivity*.

Of course, even if we accept *Reflexivity*, we can still ask: 'How come the guise that refers to Millikan *because she is using it to refer* is needed if she is to flee a bear? How come it wouldn't be equally fine if she deployed a guise that referred to her because she satisfies some great deeds description, or because she sits at the beginning of a causal chain leading up to tokens of a proper name?' It would be intellectually satisfying to have some general principles connecting reference-fixing stories and psychological role to answer these questions. How best to

articulate and motivate these and apply these to the guise associated with ‘I’ (and how to do something similar for other guises that the orthodox view also applies to) is a challenging open question within the essential indexical research program.⁹

Conclusion

We have certain quite strong intuitions about essential indexical cases. Are these strong enough to support the orthodoxy they’ve generally been taken to? Yes they are! But showing that – formulating a claim that they support, and that when accompanied with a contrast premise supports the orthodoxy – is a delicate business. The argument for the orthodox view I’ve given has *I-Needed* as its center piece. *I-Needed* isn’t a trivial corollary of the fact that action explanation contexts are opaque (which would make it too weak to support the orthodoxy). *I-Needed* doesn’t rule out the possibility of action without self-reference in the way that the corresponding universal generalization does (which would make it too strong for any intuitions we have to support it). The claim that *fleeing a bear* is a witness for *I-Needed* is given support by our intuitive reaction to a particular case, **The Bear Case**** which is that the explanation of why François does not flee in that case can only be that he isn’t thinking of himself of ‘I’. An opponent of *I-Needed* must provide some alternative possible explanation of why he doesn’t flee. Looking at C&D’s work on this, I’ve argued that the account they give is rather underspecified but that there is no version or elaboration of their account that is both plausible and that offers an alternative explanation of why François does not flee. The putative case of fleeing while just thinking of himself ‘as Recanati’ constructed on the model of the Herman-saves-Nora case is possible, but not a case of fleeing. The putative case of fleeing while just

⁹ For a proposal drawing on the non-accidental truth requirement on knowledge, see Morgan 2019b.

thinking of himself ‘as Recanati’ constructed by brute force would be a case of fleeing, but is not possible (we can accept this without accepting any brute word-action connections since the guise whose direct expression is ‘I’ also seems to satisfy *Reflexivity*). The essential indexical research program has stood up to sustained scrutiny. Those working on answering the questions that arise within it, which are after all some of the most basic questions about the nature of agency and mindedness, can continue that work in good conscience. The analogy with pre-Copernican astronomy misfires.¹⁰

References:

- Ball, D. (2015) ‘Indexicality, Transparency, and Mental Files.’ *Inquiry* 58.4 pp. 353-367.
- Babb, M. (2016) ‘The Essential Indexicality of Intentional Action’, *Philosophical Quarterly* 66: 264 (439-457).
- Bermúdez, José Luis (2016) *Understanding ‘I’: Language and Thought*, Oxford: Oxford University Press.
- Campbell, J (1994). *Past, Space and Self*. Oxford: Clarendon Press.
- Cappelen, H and Dever, J (2013) *The Inessential Indexical: on the Philosophical Insignificance of Perspective and the First Person*. Oxford: Oxford University Press.
- Castaneda, H.N (1968). ‘On the Logic of Attributions of Self-Knowledge to Others’, in *Journal of Philosophy*, 65, 439-56.
- Devitt, M (2013). ‘The Myth of the Problematic De Se’. In A. Capone & N. Feit (eds.), *Attitudes De Se: Linguistics, Epistemology, Metaphysics*. CSLI Publications, Stanford.
- Douven, I (2013). ‘The Epistemology of De Se Beliefs’. In A. Capone & N. Feit (eds.), *Attitudes De Se: Linguistics, Epistemology, Metaphysics*. CSLI Publications, Stanford.

¹⁰ Too many people to list have helped improve this paper. Thanks to audiences at London, Barcelona and York and to three excellent reviewers for this journal.

- Evans, G (1982). *The Varieties of Reference*. Oxford: Clarendon Press.
- Frege, G (1966). 'The Thought: A Logical Inquiry', in P.F. Strawson (ed.) *Philosophical Logic*. Oxford: Oxford University Press.
- García-Carpintero, M (2017), 'The philosophical significance of the *De Se*', *Inquiry* 2017 60 (3): 253-276.
- Kaplan, D (1989). 'Demonstratives', in J. Almog, J. Perry and H. Wettstein (eds.) *Themes from Kaplan*. Oxford: Oxford University Press: 481-614.
- Lewis, D (1979). 'Attitudes *De Dicto* and *De Se*', in *The Philosophical Review* 88: 513-543.
- Magidor, O (2015) 'The Myth of the *De Se*', *Philosophical Perspectives* 29 (1): 249-283.
- Millikan, R (1990). 'The Myth of the Essential Indexical', in *Noûs* 24 (5): 723-734.
- Millikan, R (2001), 'The Myth of Mental Indexicals' (revised and expanded version of 'The Myth of the Essential Indexical') pp. 163-77 in Brook and DeVidi (eds.) *Self-Awareness and Self-Reference* John Benjamins Publishing Company.
- Morgan, D (2019a). 'Temporal indexicals are essential' *Analysis* 79 (3): 452-461.
- Morgan, D (2019b). 'Accidentally about me' *Mind* 128 (512): 1085-1115.
- Peacocke, C (2008) *Truly Understood*. Oxford: Oxford University Press.
- Perry, J (1977). 'Frege on Demonstratives', in *The Philosophical Review* 86 (4): 474-497.
- Perry, J (1979). 'The Problem of the Essential Indexical', in *Noûs* 13 (1): 3-21.
- Perry, J (1990). 'Self-Notions', in *Logos*: 17-31.
- Salmon, N (1986). *Frege's Puzzle*. Ridgeview.
- Spencer, C (2007). 'Is There a Problem of the Essential Indexical?'. In M. O'Rourke & C. Washington (eds.), *Situating Semantics: Essays on the Philosophy of John Perry*. MIT Press, Cambridge.