

## *De se* thoughts and immunity to error through misidentification

Manuel García-Carpintero<sup>1</sup>

Received: 6 March 2015 / Accepted: 25 June 2015  
© Springer Science+Business Media Dordrecht 2015

**Abstract** I discuss an aspect of the relation between accounts of *de se* thought and the phenomenon of *immunity to error through misidentification*. I will argue that a deflationary account of the latter—the *Simple Account*, due to Evans (The varieties of reference, 1982)—will not do; a more robust one based on an account of *de se* thoughts is required. I will then sketch such an alternative account, based on a more general view on singular thoughts, and show how it can deal with the problems I raise for the Simple Account.

**Keywords** First-person · Self-reference · Self-knowledge · Immunity to error through misidentification

Inspired by Castañeda (1966, 1968), Perry (1979) and Lewis (1979) showed that, among singular thoughts in general, thoughts about oneself “as oneself”—which the latter called *de se* thoughts—call for special treatment, and advanced rival accounts. While Lewis aims to account for *de se* thoughts by taking the subject away from their contents, thus making them properties instead of full traditional propositions, Perry offers an account compatible with traditional Fregean views on contents.<sup>1</sup> In this paper I discuss an aspect of the relation between accounts of *de se* thought and the phenomenon of *immunity to error through misidentification*. I will argue that a deflationary account of the latter—the *Simple Account*, due to Evans (1982)—will not do; a more robust one based on an account of *de se* thoughts is required. I will

---

<sup>1</sup> García-Carpintero (forthcoming) develops further the debate between the Lewisian and Fregean options, and defends a version of the latter.

✉ Manuel García-Carpintero  
m.garciacarpintero@ub.edu

<sup>1</sup> Departament de Lògica, Història i Filosofia de la Ciència, Universitat de Barcelona, Barcelona, Spain

then sketch such an alternative account, based on a more general view on singular thoughts, and show how it can deal with the problems I raise for the Simple Account. The first section introduces the issue; the second objects to the Simple Account, and the third outlines my alternative proposal, based on the account of *de se* thoughts I have presented elsewhere (García-Carpintero, forthcoming).

## 1 Immunity to error through misidentification: Recanati's Lewisian account

Wittgenstein (1958, pp. 66–67) pointed out a distinction between two uses of 'I', one "as object" and another "as subject", which Shoemaker (1968) elucidated in terms of a notion of *immunity to error through misidentification* ('IEM' henceforth). This is an epistemological phenomenon that it is *prima facie* related to the special character of the *de se* thoughts expressed with uses of the first-person. One should hence expect that accounts of *de se* thoughts explain or otherwise illuminate IEM (Peacocke 2014, pp. 109–110). On that assumption, Recanati (2007, 2009 and 2012) argues for a Lewisian account of some "basic" or "implicit" *de se* thoughts, contending that this provides a good explanation of IEM. Several writers (Stanley 2011, pp. 91–93; Morgan 2012; Wright 2012) have rejected Recanati's claim. Here I want to agree with Recanati that the IEM of *de se* thoughts should be explained on the basis of an account of those thoughts, even if I ultimately disagree with him that the required substantive account is a Lewisian take on the *de se*.

Shoemaker (1968, p. 557) specifies the sort of mistake that IEM excludes as follows: "to say that a statement '*a* is  $\varphi$ ' is subject to error through misidentification relative to the term '*a*' means that the following is possible: the speaker knows some particular thing to be  $\varphi$ , but makes the mistake of asserting '*a* is  $\varphi$ ' because, and only because, he mistakenly thinks that the thing he knows to be  $\varphi$  is what *a* refers to". In order to illustrate the idea, let me use Perry's (1979, p. 33) famous shopper example: "I once followed a trail of sugar on a supermarket floor, pushing my cart down the aisle on one side of a tall counter and back the aisle on the other, seeking the shopper with the torn sack to tell him he was making a mess. With each trip around the counter, the trail became thicker. But I seemed unable to catch up. Finally it dawned on me. I was the shopper I was trying to catch."

Now, consider the moment Perry introduces later when the shopper sees what in fact is his image in a mirror with a torn sack (ibid., p. 42): "Suppose there were mirrors at either end of the counter so that as I pushed my cart down the aisle in pursuit I saw myself in the mirror. I take what I see to be the reflection of the messy shopper going up the aisle on the other side, not realizing that what I am really seeing is a reflection of a reflection of myself." Let us imagine a variation on the story, in which this is instead the ground for the shopper's epiphany, because he recognizes himself in the mirror; assume then that he judges on this basis what he would express by 'I am making a mess'. Made on such epistemic grounds, the claim is vulnerable to the error that Shoemaker identifies. For the shopper might have been wrong (although in the story he is not) in identifying himself with the person whose back is reflected in

the mirror; he would then have been right that someone is pushing a cart with a torn sack and is thus making a mess, but wrong to think that it is he who is responsible.

Pryor (1999) offers a definition of IEM alternative to Shoemaker's linguistic characterization, free from concerns that this might raise.<sup>2</sup> Pryor's definition is in terms of justified belief instead of knowledge, and it highlights the relativity to ways of reaching the relevant judgment that, as this example already makes clear, any proper characterization should capture: a judgment can be IEM if made on a given epistemic basis in normal circumstances, and not if made on different grounds. Pryor defines first *de re misidentification*, which occurs when the following three conditions obtain—here they are illustrated relative to the imagined variation on Perry's story:

- (i) There is some singular proposition about  $x$ , to the effect that it is F, that a subject believes or attempts to express. (The shopper believes a singular proposition about himself, to the effect that he is making a mess.)
- (ii) The subject's justification for believing this singular proposition rests on his justification for believing, of some  $y$ , and of  $x$ , that  $y$  is F and that  $y$  is identical to  $x$ . (In the imagined circumstances, the shopper's justification derives from his justification for believing of the person whose back he sees that he is making a mess, and for identifying himself with that person.)
- (iii) However, unbeknownst to the subject,  $y \neq x$ .

Pryor (1999, p. 279) then stipulates that a judgment justified by grounds G is IEM just in the case that it is impossible for the judgment to be a case of *de re misidentification* when made on grounds G. As explained, the shopper's judgment in the example is vulnerable to this sort of mistake, even if it is not in fact a case of it. Now, in the original version of the story the epiphany is complete when the shopper visually confirms what has dawned on him after seeing the trail getting thicker, that the torn sack is in his own cart. Consider now his physical self-ascription, 'I am pushing a cart with a torn sack', or his psychological self-ascription, 'I see that I am pushing a cart with a torn sack', both based on his visual experience. Neither of these claims *prima facie* appears to be open to the sort of error that Shoemaker and Pryor describe; nor is the thought he expresses by 'I am making a mess', when based on the same grounds.

Our examples show that, if there is a connection between *de se* thoughts and IEM, it must be indirect; for the thought that the shopper expresses by accepting 'I am making a mess' is *de se* not just in the original IEM case, but also in the previously imagined alternative non-IEM situation. The subtler connection might be this: 'I'-thoughts that are IEM are *fundamentally de se*; those that are not are only *de se derivatively*, in that, in making them by using the first-person concept, the speaker identifies himself under that concept as the object of other, *fundamentally de se* thoughts. In the terms of this speculation, Recanati's defense of the Lewisian account of *de se* thought concerns only the *fundamentally de se*. He himself (2007, p. 177) distinguishes *explicit* from *implicit de se* thoughts. The former are attitudes to traditional propositions, the latter

<sup>2</sup> Here is one, suggested to me by Ernest Sosa. On Shoemaker's explanation, any statement appears to be vulnerable to error through misidentification. For it is always possible that the term ' $a$ ' might have meant something other than what it means and that the speaker might then have mistakenly thought that the thing he knows to be F is what ' $a$ ' refers to.

to properties. This explains the IEM of thoughts such as ‘my legs are crossed’, based on proprioception (2007, p. 147):

The content is a relativized proposition, true at a person, and the internal mode determines the person relative to which that relativized content is evaluated: myself. In Perry’s terms, the self is an unarticulated constituent of the judgment. The content of the judgment, therefore, is not the proposition that *my* legs are crossed, strictly speaking. What is articulated is only the person-relative proposition that one’s legs are crossed. Since I am not explicitly ‘identified’ as the person whose legs are crossed, I cannot be misidentified. The judgment, Evans (1982, pp. 180–181) says, is ‘identification-free’.

Now, Recanati’s appeal to Evans here might be misleading. Various writers including John Campbell, Christopher Peacocke and Crispin Wright have developed an account of IEM suggested by Evans’s notion of identification-freedom that Wright (2012, p. 253) calls “the Simple Account” (‘SA’ henceforth). As I will show now, SA provides an account of IEM that is not only compatible with the Lewisian account of the *de se*, but also with one along the lines of Perry’s. On SA, non-IEM thoughts are thoughts the structure of whose epistemic justification is such that they depend on an identity claim; thus, for instance, the shopper’s judgment ‘I am making a mess’ in the previously imagined version of the story, which was not IEM, depends on the shopper’s identity judgment, *I am that person reflected in the mirror*. This is why he himself might coherently conceive that, although the existential “part” of his claim—that someone is making a mess—is correct, he is mistaken in the identification, and it is not in fact he himself but someone else who is making a mess. SA characterizes IEM thoughts negatively as those that are not thus dependent on an identity claim.

The account that SA affords appeals to the absence of an identity claim *in the justificational structure* of IEM thoughts; Recanati’s explanation appeals instead to the absence of a conception of the self *in their content*. SA leads us to distinguish between two senses for a thought to have an identification component. In the first sense (‘identification<sub>p</sub>’), the epistemic grounds for the thought include an identity-premise. In the second sense (‘identification<sub>C</sub>’), the thought includes a concept that identifies what it is about. Given SA, having an identification<sub>C</sub> component is compatible with a thought being IEM; for being IEM is *lacking* an identification<sub>p</sub> component, and thoughts having identification<sub>C</sub> components may well be as much identification<sub>p</sub> free as Lewisian relativized thoughts surely are—they might even be epistemically basic. In contrast, Recanati’s account explains core cases of IEM by the absence of an identification<sub>C</sub> element.

SA thus opens up the possibility of thoughts including identification<sub>C</sub> components, which are nonetheless IEM with respect to them. Hence, supporters of a Fregean account of the *de se* can happily appeal to SA to explain IEM.<sup>3</sup> And it seems to be the case that there are such examples of IEM thoughts with identification<sub>C</sub> constituents when we consider other indexical thoughts than *de se* thoughts. Wright (2012, p. 254)

<sup>3</sup> One can thus understand Evans’s reliance on SA, given his neo-Fregean conception of the contents of thoughts exhibiting IEM, and his lack of sympathy for any form of truth-relativism. Cf. Grush (2007) for an accurate exegesis of Evans’s notion of identification-freedom.

offers as examples ‘you are standing very close’ and ‘he is a long way off’, both based on observation; Peacocke (2008), ‘this keyboard is black’, again based on observation. Hence, without further elaboration, Recanati’s account of IEM will not do. He (2009, p. 259; 2012, pp. 190–196) accepts this, even for some first-personal IEM thoughts. A thought expressed by ‘my legs are crossed’ based on proprioceptive evidence is IEM on his original account because the content is just the property of having crossed legs, which the subject self-ascribes. The thought expressed by ‘it is my legs, not my neighbor’s, that are crossed’ cannot plausibly be thought not to include the concept of the thinker’s leg; but it is still IEM.

In response, Recanati argues (ibid.) that “subject-explicit” thoughts are only IEM because they are derived through a process he calls ‘reflection’ from a “subject-implicit” property-ascription, and hence have the same grounds as the latter, which are identification<sub>C</sub>-free and hence also identification<sub>P</sub>-free: “a judgment is immune to error through misidentification if it is implicitly *de se*, that is, if the subject is not represented in the content of the judgment but his or her involvement is secured by the mode of the grounding experience; yet an explicit *de se* thought may also be IEM if it has the same grounds as an implicit *de se* thought.” But how could this work for demonstrative thoughts? *Prima facie*, the idea seems hopeless. While the subject, together with the world and the time of evaluation, can with some plausibility be taken to be given by the attitude or representational act, as opposed to being constituents of the represented content, it is difficult to make sense of the same idea when it comes to demonstrated objects. Recanati (2012, pp. 199–200) attempts to motivate such an extension of his proposal, but Wright (2012, pp. 273–278) raises serious concerns.

## 2 Immunity to error through misidentification: the simple account is not enough

We have seen in the previous section that Evans’s notion of identification-freedom does not particularly support the explanation of IEM offered by the Lewisian view, contrary to Recanati’s initial suggestion: the Fregean line that assumes a full-proposition view of *de se* thoughts can also rely on SA to explain IEM. However, supporters of the Fregean line would be well advised to look for an alternative, more robust account of IEM than the one afforded by SA; to that extent, Recanati is right. This is the point I want to make in this section. I will provide three reasons for it, in order of increasing importance.<sup>4</sup> These are: (i) the need to explain *in virtue of what* some beliefs can dispense with an identity-judgment in their background; (ii) the need to make room within the account for a distinction between *de facto* and logical/*de jure* IEM; and (iii) the need to accommodate the phenomenon of wh-IEM.

- (i) In the first place, Recanati can argue that the Lewisian view, which is as compatible with SA as Fregean accounts, in fact offers the best explanation for there being

<sup>4</sup> Wright (2012, p. 255) is well aware that SA needs at least “clarification and qualification”, essentially for the reasons I will suggest. I’ll discuss his proposals below.

no identity premise involved in the justificational structure of IEM thoughts.<sup>5</sup> Now, I do not think this response is ultimately satisfactory, but the reason why it is not manifests the need to go beyond SA to properly explain IEM. The Lewisian view is indeed compatible with SA, but Recanati purports to offer a more specific explanation than the one afforded by it. To that effect, he invokes a feature (absence of an identification<sub>C</sub>) that entails the absence of an identity premise, without being mandated by it. If this more specific proposal cannot account for some cases (e.g., IEM thoughts that do involve an identification<sub>C</sub>), perhaps another explanation, equally compatible with SA but alternative to the Lewisian one, might be the truly operative one in all cases, accounting for them in a unified way. I think this is in fact the case. To defend it we should obviously go beyond SA, indicating specifically how accounts allowing for identification<sub>C</sub> without identification<sub>p</sub> explain IEM. The final section offers one such more substantive account.

- (ii) The second reason for going beyond SA will need more elaboration. It relates to the fact that the IEM of the cases we have offered so far appears to be merely *de facto* (cf. Shoemaker 1970) obtaining only if circumstances are *normal*, and the need to explain why they are still cases of IEM; examples will be given presently. Let us first note that SA relies on the notion of *doxastic* justification, hence on the *grounding* relation. *Propositional* justification is justification for a proposition, whether or not one believes it; *doxastic* justification is justification that supports a belief a subject has, in terms of a structure of mental states that explain the holding of the belief by the subject in an epistemically valuable way.<sup>6</sup> For a judgment A to be epistemically grounded on B does not require that the subject consciously experience his coming to judge A as a result of an inference in part from B. In our example above, the shopper might well lack awareness of inferentially relying on his identification with the person whose back he sees in the mirror, but his judgment in fact relies on it—as shown in the fact that on appropriate prompting he would be prepared to admit to the dependence.

There are propositions that affect the epistemic status of a belief, but there is an ongoing debate about how exactly they do so, and, relatedly, to what extent the belief needs to be based on them for it to be justified. Consider Moore's (in-)famous awkward inference, (i) here are two hands, (ii) if there are hands, there is an external world, hence (iii) there is an external world. Given its validity, someone who judges (i) is thus rationally committed to (iii); but there are different ways of understanding such commitments and thereby explaining the oddness of the inference. Pryor (2004) distinguishes two epistemological attitudes we may have with respect to them, a *liberal* and a *conservative* one. On the conservative attitude, having *prima facie* justification to believe (i) requires antecedent justification to believe (iii); the liberal denies this, even though he agrees that evidence against (iii) would defeat our justification to believe (i). We should further distinguish between two versions of the conservative attitude;

<sup>5</sup> Wright (2012, pp. 263–264, 273) grants this, which is also the line that Recanati (2013, pp. 236–238) takes in response to the criticism so far, as presented in García-Carpintero (2013).

<sup>6</sup> Chalmers (2012, pp. 94–99) outlines an interesting analysis; Audi (1986) provides a more detailed account, which I like.

on the most straightforward *strong* version, justification for (i) would require a priori or empirical evidence for (iii); on a *weak* one along lines explored by Wright (2004), it suffices for justification that (iii) is a presupposition that one is entitled to make by default.<sup>7</sup>

These views can be applied to the identity claims that might be discerned in the justificatory structure of our singular thoughts. The conservative strong attitude is adequate concerning the shopper's identification with the person whose back he sees in the mirror. The reason is that, as we indicated above, the shopper's judgment that he is making a mess in such a case is grounded on the premise that he is the person whose back he sees in the mirror, and the former would not be justified if the latter were not. Consider, however, the following example inspired by a similar one by Wright (2012, p. 267): you see someone you take to be your wife, and judge: my wife is across the street. This is vulnerable to misidentification. Perhaps unbeknownst to you, you just met a perfect look-alike of your wife; you are right that someone is across the street, but she is just a double of your wife, whom you have wrongly identified with her. So, the justification of your belief depends on the justification of this identity claim, but is it reasonable to suppose that your belief was grounded on it? It appears to be *immediately* justified by your perceptual recognition. We could thus instead adopt a liberal view about the case: your judgment was not based in any way on the identity judgment, but you nonetheless were justified. Or we can perhaps adopt a conservative view of the weak form: the identity judgment can be said to be a background presupposition of your judgment, to which you were entitled by default.

Similar issues arise about judgments mentioned earlier as examples of IEM, as the ongoing debate between Evans and Shoemaker reveals.<sup>8</sup> Consider the shopper's physical self-ascription that he is pushing a cart with a torn sack, made on the basis of visual perception, or his psychological self-ascription that he *sees* (not just *appears* to see) that he is pushing a cart with a torn sack. Shoemaker (1970, p. 283) takes them as cases of *de facto* IEM; they are IEM, but they are vulnerable to misidentification error under abnormal conditions. Suppose that the science fiction technologies that the film *Avatar* imagines are available, allowing us to temporarily break off our own visual impressions and receive instead those coming from another body.<sup>9</sup> Under those circumstances, the judgments might be wrong because the shopper mistakenly identifies himself with the body that is the source of the relevant visual impressions.

We can interpret *de facto* IEM along the conservative lines suggested above for the wife case. In normal circumstances, the weak form rules: the subject reasonably and correctly takes for granted the identity premise, so that the judgment is appropriate even though not based on it. Abnormal conditions—a circumstance in which the otherwise merely presupposed identification might fail (your wife has an identical twin) becomes

<sup>7</sup> Cf. Neta (2011) for an introduction and a sophisticated discussion of these matters. Cp. my distinction between liberalism and weak conservatism with Neta's distinction (op. cit., p. 691, 695) between *simple* and *strong* liberalism.

<sup>8</sup> McGlynn (2015) and Morgan (ms) provide recent useful appraisals of the debate. In contrast to what I will suggest, both side with Evans against Shoemaker and Pryor.

<sup>9</sup> Dennett (1978) presents such a scenario; cf. also Higginbotham's (2010, p. 262) "Davidson's eyes" example from H. G. Wells.

a relevant alternative—override this, mandating the strong stance. In such contexts the judgment must be based on the identity premise for it to be warranted, and the identity premise substantively justified. On this view, the commitment to identifications that ordinary self-ascriptions based on visual perception in fact carry is understood along the lines of weak conservatism, as opposed to the strong form that suggests itself as more appropriate for the mirror example. In normal contexts these identities are presuppositions to which we are entitled by default, without the need to have ordinary *a priori* or empirical evidence for them.<sup>10</sup>

Examples of thought-insertion in schizophrenic patients discussed in the literature (cf. Recanati 2007) might show that some mental self-ascriptions (*I am judging that I am good and omnipotent*) are liable to the same account.<sup>11</sup> I might have conscious awareness of seeming to make the judgment that I would express by ‘I am good and omnipotent’, while worrying whether it has in fact been planted in me by an evil scientist in control of my brain. (I assume I cannot have the impression of *actively making the judgment*, in contrast to the ordinary case; it must be rather like being aware of an unbidden utterance in inner speech, which would in normal circumstances express the judgment.)<sup>12</sup> If the worries are sensible, the strong form is required and the judgment, when it is nonetheless justifiably made, must be based on the identification of myself with the producer of the judgment; otherwise, the weak attitude prevails and the judgment is identification<sub>p</sub>-free, but nonetheless justified.

Now, how could an account that relied merely on SA justify the contention that the examples above are in fact cases of IEM, even if only *de facto*, perhaps along such lines? Consider Cappelen and Dever’s (2013, pp. 130–133) skepticism about IEM. They want to establish that *de se* thoughts are philosophically unimportant. On their way to this goal, they argue that no significance can accrue to them from their role in explaining IEM, simply because there is nothing there to explain: there is no IEM. They adduce in support the scenarios just contemplated to show that previous

<sup>10</sup> Cf. Coliva (2006, 2012) and Wright (2012, pp. 267–272); cf. also Peacocke (1983, pp. 139–151), and Peacocke (2008, pp. 92–103).

<sup>11</sup> Fernández (2010, pp. 67–70) provides a nice description of the data to be accounted for in cases of thought insertion.

<sup>12</sup> Cf. the debate between Coliva and Campbell (Coliva 2002a, b; Campbell 2002) about two different relevant senses of ownership, one—related to the “self-intimating” views on introspection to be mentioned below—on which the subject of an introspected state is guaranteed to be the introspecting subject, and another specific instead to the subclass of mental states including those that can be legitimately counted as mental actions. In the analysis of (rationally intelligible) cases of thought insertion suggested in the main text I am refusing to use a single category (such as the indistinct “thought”) to describe them. (Fernández 2010 talks of *beliefs*, but this is even less of a good candidate, because the relevant cases are episodic or occurrent states. Note, in response to some of Fernández’s criticisms of a suggestion like mine, that even judgments that are responses to perceptual experiences can be sensibly counted as actions, in that one could have refused to take the experience at face value and hence refused to form them, as we all do with illusions we recognize as such.) The subject introspects a passive state, such as an episode of inner speech, and she acknowledges it as her own; this is what I describe as *the impression of being making a judgment*. She disowns the corresponding judgment, and hence does not introspect it. Whether or not this is an accurate description of the relevant cases in schizophrenic patients, I cannot tell. However, it fits well with a tendency in psychiatry to make sense of delusions as rational responses to weird experiences; as far as I can tell, it fits the data; and it is consistent with both “subpersonal” and “personal” explanations of the disowning of the judgment in the literature (cf. Fernández, op. cit., pp. 72–75).



examples are only *de facto* IEM; they appear to take them to show that there is no *de jure* (as we may call it) IEM, and as a result no genuine IEM. Now, the examples do not in fact support anything of this sort. In the thought-insertion case, for instance, one is still conscious of *being under the impression* of making the judgments or inner utterances; and I cannot see how an analogous worry extends to the IEM of that state, against what Cappelen and Dever contend.<sup>13</sup> Similar points can be made about other cases.<sup>14</sup>

Let us put this aside, though, for I think it is important to establish that ordinary cases are also IEM, even though only *de facto*.<sup>15</sup> It should be clear that by relying on SA we would be deprived of the resources to do this in any illuminating way in response to skeptical challenges such as this one. Firstly, if we just apply the criterion of the absence of identification<sub>p</sub> without further embellishment, we would end up by misclassifying the judgment in the wife example above as IEM. Secondly, no straightforward appeal to the distinctions above would help to distinguish that case from cases of *de facto* immunity.<sup>16</sup> Defending that *de facto* IEM is a form of IEM, and explaining the distinction between it and *de jure* IEM, thus requires us to go beyond SA.

Someone elaborating on SA along Recanati's lines by appealing to the Lewisian account of the *de se* has more resources. On this view, in a normal case the subject is not part of the content, whereas in abnormal ones he is. The proponent of this view can now explore whether this theoretical distinction allows further justification for a claim

<sup>13</sup> Cappelen and Dever consider this possibility (op. cit., p. 131): “maybe what is IEM for Gareth is his belief that his legs *seem* to be crossed”, and go on to contend: “But now imagine that Gareth is deviantly wired, so that he receives “introspective” awareness from John’s seemings”. Now, in the first place intuition does not support the suggestion that one could be introspectively “wired” to someone else’s seemings. To make sense of this requires a view of introspection as a pure causal process that most contemporary philosophers reject. The possibility would not exist, for instance, on Gertler’s (2012), Peacocke’s (1999, Chap. 6), or Shoemaker’s (1994) otherwise different accounts; more on this below. A proper defense of their claim would thus require an outline of an account of introspection that makes sense of the case (say, a straightforward perceptual model). Cappelen and Dever do not provide anything of the sort.

<sup>14</sup> Lane and Liang (2011) provide another skeptical argument against IEM, based on interesting abnormal cases. Thus, a patient with somatoparaphrenia refers an introspected tactile sensation in her arm to the arm of her niece; and a subject of a “body swap illusion” experiences as her own an action (shaking her own hands) which she knows someone else, the experimenter, is in fact carrying out. In my view, these cases should be treated in the way I have suggested for thought-insertion. They only show that cases of what we would have *prima facie* considered *de jure* IEM are in fact cases of merely *de facto* IEM—and call for an account of the distinction. On the basis of similar considerations, Howell (2007) defends a more nuanced skepticism about IEM than these writers, which does not extend to introspection-based judgments. He is only skeptical that IEM affords a kind of non-Cartesian privilege that can be acknowledged without indulging Cartesian metaphysical consequences. His point is that (*de facto*) IEM ultimately depends on the (*de jure*) IEM of less committal mental states such as impressions or seemings, which are precisely those possessing the potentially problematic form of Cartesian privilege.

<sup>15</sup> This is so because I agree with Evans (1982, pp. 242–243) and Peacocke (2014, pp. 86–99) that ordinary cases have a foundational explanatory character; more on this below.

<sup>16</sup> McGlynn (2015) raises this point against Coliva (2006). On account of it, Wright (2012, pp. 271–272) ends up acknowledging a huge restriction of the extent of IEM; in spite of this, unlike Howell (2007), who, as mentioned above, argues for a similar restriction, he contends that the qualified version of SA that he accepts suffices to reject the “Psychological Hypothesis” that IEM is explained by features of psychological self-knowledge. I question this below.

of *de facto* IEM in the former case.<sup>17</sup> Similarly, the proponent of the Fregean line can also appeal to what she takes the *Self* concept or mode of presentation to be; cf., for example, the account that Peacocke (2014, pp. 108–113) offers, assuming such a Fregean account of *de se* thoughts.<sup>18</sup> I myself prefer the second option, having offered an account of this kind in a companion piece (García-Carpintero forthcoming). I will outline the account it affords in the final section.

(iii) The third and final reason lies in that, as Pryor (1999, pp. 281–286) argues, immunity to *de re* misidentification is not the truly general epistemological phenomenon in need of explanation; it is rather immunity to what he calls “wh-misidentification”.<sup>19</sup> I peel a few oranges, separate the segments, put them on a dish and go out to buy some sugar and port to make a dessert; when I come back, all the segments have vanished. I judge *someone ate my oranges*, without assuming a unique culprit; several family members are around, and for all I know a few of them might have each eaten some segments. Then I see Mary with some orange peel in her hand, and judge *Mary ate my oranges*. Here no identification (no identity premise) is involved in my reasoning; I just infer that Mary is a witness to an existential generalization that I accept, open to the possibility that there are others. The judgment is IEM, given SA; but my singular thought is vulnerable to a mistake very similar to the one Shoemaker had in mind: the properly singular, identifying “part” of my judgment can be wrong, without the existential “part” being wrong.

Pryor analyzes this talk of *parts of contents* in terms of *undercutting defeaters*: the identification can be defeated (I may come to have evidence that Mary did not touch the oranges) while the existential claim stays undefeated.<sup>20</sup> A judgment to the effect that

<sup>17</sup> Cf. Recanati’s (2007, pp. 150–160) distinction between *grade-1* and *grade-2* immunity.

<sup>18</sup> Even though Peacocke does not take SA to provide the full story about IEM, but relies instead on his own account of the *Self* concept to offer a positive and helpful account, he does wrongly rely on SA to characterize the phenomenon to be explained. As we are about to see, this is something that the case of wh-immunity challenges. Peacocke (2014, p. 109) rejects modal characterizations of IEM like the one I rely upon following Pryor and Shoemaker. His reason is that modal definitions would force us to intuitively misclassify some cases as instances of IEM. “Suppose ... that in normal circumstances mirrors were always entirely flat, and that they only reflected faces very close to the mirror, so that in those normal circumstances, for any person *x*, only *x* can see the face and head of *x* in the mirror. In those circumstances, it would not be possible for someone, basing his judgment on what he sees in the mirror, to be right that someone’s hair is tidy, but wrong that his hair is tidy.” Why not? Even in such a world, we can easily conceive of being in a *Duck Soup* situation, not facing a true mirror but a transparent piece of glass behind which someone looking very much like us devilishly apes our doings.

<sup>19</sup> Coliva (2006, pp. 7–11) questions the need for differentiating this form of immunity from the one previously introduced. I think she relies on some idiosyncrasies of Pryor’s example; the one I offer is intended to circumvent the problem. Cf. Wright (2012, pp. 267–269).

<sup>20</sup> An undercutting defeater for *p* just takes away one’s justification for believing it; and *additive* defeater gives, in addition, positive evidence for *not-p* (Pryor 1999, p. 283). Defeaters should not be additive, because otherwise there would not be cases of wh-immunity, not even *I am in pain*. This is precisely what Smith (2006, p. 279) argues for, unfortunately for his argument by using only additive defeaters: “suppose you experience what you take to be a pain. Now suppose someone reliably tells you that you are not really experiencing a pain but an itch, and that this has been caused by someone else’s suffering a real pain. Whenever they experience a pain, they press a button which causes you to feel a pain-like itch.” Smith confronts the obvious problem that this does not affect Pryor’s notion by providing (*op. cit.* 280) an argument that ruling out the example on the grounds that the defeater is additive—packing the justification for the existential generalization—makes wh-misidentification collapse with *de re* misidentification; but the

$a$  is  $F$  based on grounds  $g$  is immune to this kind of mistake just in the case  $g$  justifies the singular judgment—and to that extent its existential generalization too—without providing independent justification for the latter: the justification for the existential generalization comes indissolubly together with a justification for the corresponding ascription to a witness.<sup>21</sup> In the previous example, the justification I had warrants the existential generalization that someone had eaten the orange slices independently of the justification for the belief that it was Mary who ate them.<sup>22</sup>

It seems that the core examples of IEM (*I am in pain*) are also immune to this new kind of mistake. Now, wh-immunity on grounds  $G$  entails immunity to *de re* misidentification relative to  $G$ , because when wh-immunity obtains the justification that  $G$  offers to believe *something is F* must suffice by itself to justify the belief *it is a who is F*. The converse does not obtain under SA, because orange-thief cases exhibit immunity to *de re* misidentification (the judgment does not rely on any identity premise); but intuitively they are not examples of the sort of immunity we are after. Now, the interesting epistemic phenomenon appears to be the more encompassing one: what we would like to understand better is why in some cases the justification for the existential “part” of a singular claim comes indissolubly together with a justification for the corresponding ascription to a witness. But SA cannot explain this because, as we have illustrated, the absence of an identification premise is compatible with wh-vulnerability.<sup>23</sup>

---

Footnote 20 continued

orange example shows the argument to be spurious. He also expresses doubts that the distinction “can bear the required weight” (op. cit., p. 281, fn. 14), but he does not motivate them. Morgan (ms) provides a subtler but related argument against Pryor’s characterization of the notion of wh-immunity, without questioning that it is a distinctive kind.

<sup>21</sup> Wright (1998, p. 19) and Campbell (1999, p. 89) suggest similar views. The path Pryor follows to reach his own final definition of vulnerability to wh-misidentification, and hence immunity to it, is tortuous. In trying to avoid the complications, in a previous version of this material I just relied on this characterization, which I took to be at the heart of the notion, and which glosses what Pryor says at some points (“if your belief of  $a$  that it is  $F$  is immune to wh-misidentification when justified by certain grounds, that means that the justification those grounds give you for believing that something is  $F$  suffices by itself to justify you in believing that it’s  $a$  which is  $F$ ”, op. cit., p. 285). Morgan (ms.) convincingly argues that it is better to replace Pryor’s definition by this core. I am very much indebted to him for discussion of these issues.

<sup>22</sup> Morgan (ms.) suggests the following test for the independence of the justification that  $g$  provides for the existential generalization from the justification it provides for the singular claim: the former could be known on the basis of  $g$ , without the latter being true.

<sup>23</sup> The qualification to SA that Wright adds in order to deal with this is that the justificational architecture of IEM thoughts “is non-inferential”, in addition to not involving any identity background presupposition of the kind needed to deal with the previous issue (op. cit., p. 271). Judgments liable to either kind of misidentification “involve an inferential justificational architecture involving the synthesis of multiple items of presumed information: ... an identification; ... a unique existential” (op. cit., p. 260). In addition to the big restriction in the scope of IEM already pointed out in fn. 16 above, this proposal questions Wright’s main claim—his rejection of the “Psychological Hypothesis” that IEM is explained by features of psychological self-knowledge. For accounts assuming forms of that hypothesis such as Howell’s (2006), capable of explaining on that basis the absence of the justificational architecture that Wright simply takes as explanatorily primitive, are more illuminating. I’ll make a similar point below in favor of my own proposal, which has the additional virtue of preserving *de facto* IEM.

In the next section I will show how a Fregean account of *de se* thought improves on the explanation of IEM offered by SA on each of the three counts raised in this one. Given the point just made, I'll be assuming that immunity to wh-misidentification, as just characterized, is the form of IEM that we are interested in accounting for.

### 3 The presuppositional view of the *de se* and IEM

The account I have provided of *de se* thought applies a more general account of singular thoughts/claims that I have been suggesting over the years (cf. García-Carpintero (2000, 2006)), on which they involve *reference-fixing presuppositions*. Let me briefly introduce the view, focusing first on the linguistic expression of such thoughts.

Although the state of information we end up in by accepting 'John stole the camera' and 'it was John who stole the camera' is the same, these two sentences pack the information they convey in different ways. The second, cleft sentence *presupposes* that someone stole the camera, but the former, plainer sentence does not. For present purposes, we can think of presuppositions along the well-known lines that Stalnaker (1978, 2014) has suggested.<sup>24</sup> Speech acts like assertions take place relative to a *common ground*, a set of already accepted propositions. Linguistic presuppositions are requirements on the common ground, whose satisfaction should be checked at an ideal time just after the utterance is made (because it might well be that it is the utterance itself that generates the common ground information that satisfies them) but before the resulting assertion is accepted—in which case it then goes to conform the common ground, licensing further presuppositions in the ensuing discourse. The difference between our two sentences lies in that an utterance of the cleft sentence will feel inappropriate (at presupposition evaluation time) if it is not common ground that someone stole the camera. But the state of information that we get into by accepting either of our sentences will be the same.

Consider then an utterance of 'he is hungry'. The proposal agrees with the *direct reference* theorist that the asserted content is a singular proposition, *x is hungry*, for some contextual assignment to *x*. It is expressed, however, in a context in which another singular proposition is presupposed—in this case, one semantically triggered by something akin to a Kaplanian character for 'he'—which we could express thus: *x is the male picked out by the demonstration associated with he*,<sup>25</sup> where the bold-faced 'he' refers to the relevant token. These semantically triggered presuppositions are about linguistic expressions—in fact, reflexively about tokens of the very linguistic expressions conveying them. They will typically be supplemented by further pragmatically triggered presuppositions, specifying additional features of the demonstrated referent,

<sup>24</sup> My own views, although strongly influenced by him, differ at some points (cf. García-Carpintero 2015). I think of all speech (and mental) acts, including ancillary acts such as *presupposing* and *referring*, as constitutively normative; in particular, I think of presuppositions as constituted by normative requirements that their contents are already known. I also think that some linguistic presuppositions are semantically triggered, as will transpire below.

<sup>25</sup> This statement of the presupposition contains another presupposition associated with the definite description, which I do not unpack further for the sake of perspicuity.

perceptually accessible or accessible from previous discourse.<sup>26</sup> (These presuppositions, by the way, illustrate the point made in the previous paragraph about the time at which presuppositions are to be appraised.) The descriptive identification embodied in such presuppositions is “reference-fixing” and not “meaning-giving”, in Kripke’s (1980) sense.

This is not, it should be clear, a reductive view: far from aiming to reduce singular representational states/acts to descriptive general ones, it assumes primitively singular attitudes. The suggestion is only that general descriptive information helping to fix the individuals the utterances are about is a constitutive feature of them. This information figures in associated presuppositions. But the presuppositions are themselves singular, and not just because they may mention singular token representational states; the intended referents, if there are any, contribute to characterizing their contents. Singularity, like presuppositionality, is here understood as a constitutive feature of the representational devices, a “semantic requirement” (Fine 2007) on them—a fact to be embedded in a theory of such representations, which must be grasped for them to be fully comprehended.<sup>27</sup> It is thus that singular terms are *de jure* rigid, in Kripke’s (1980) sense. The source of those requirements is diverse. Extending his views to the case of thoughts, Fine speaks of “representational requirements”, whose sources are whatever grounds a “psychosemantics”—for instance, (teleo-) functional features if functionalism is the proper account, as in fact I think. Similarly, in the linguistic case the source cannot be semantics narrowly understood (the compositional grounds of linguistic meaning), but must also include Gricean maxims and other pragmatic factors. Subjects making the presuppositions envisaged here need only be potentially aware of it, on condition of full understanding and possession of all the required concepts (cf. García-Carpintero forthcoming for details).

The outlined proposal applies to utterances paradigmatically expressing *de se* thoughts. Thus, an utterance of ‘I am hungry’ asserts that *x* is hungry, presupposing that *x* is the utterer of *I*. In the companion piece I have referred to above (García-Carpintero forthcoming), I show how parallel claims apply to the realm of thoughts.

<sup>26</sup> Proposals along these lines are quite standard nowadays in the linguistic semantics literature; cf. Heim (2008), Maier (2010) and Hunter (2013). Maier (forthcoming) extends this sort of account to mental states in a way compatible with my own proposal outlined below.

<sup>27</sup> Singular representations thus understood may fail to have an object; there are, e.g., singular presuppositions associated with singular terms in fictions that are merely pretend. Our theoretical claims are to be understood as made in the framework of a free logic. Also, the proposal raises the question whether Frege puzzles can arise at the level of the presuppositions themselves, which after all are ultimately nothing more than background attitudes (I thank Gregory Bohner, Josh Dever and Dilip Ninan for presenting the worry): if this man is Joseph, how come I do not make the connection when I presuppose in one context, of him, that he is the demonstrated male when a given token is produced, and, in another, that he is picked out by the naming practice to which a given token of ‘Joseph’ belongs? In reply, I am assuming something like the apparatus of Fine’s (2007) Semantic Relationalism (cf. also Heck 2012). In some cases we assume anaphoric relations among the singular terms we deploy in thought, in some others we don’t; in the latter cases, but not in the former, the variables we use in the metalanguage to characterize the presuppositions can be assigned different values. Unlike Fine, I think these anaphoric dependences are token-reflexive monadic properties, invoked in additional presuppositions. This raises the further worry whether the initial presuppositions are needed: am I putting forward a redundant, overdetermined set of explanations? But they are independently needed, to capture contextually valid inferences (*he is Bolivian, therefore a male is Bolivian*) which we are disposed to make as part of our full linguistic competence.

On this view, a token *de se* thought of  $x$  would deploy a token self-concept [ $\text{self}_x$ ] (Peacocke 1981) whose content would be  $x$  and whose associated presupposition is something like  $x$  is the subject of [ $\text{self}_x$ ]. The value of  $x$  could equally be the content of other concepts, such as [that man reflected in a mirror]. But it is constitutively assumed in the thought-state in which [ $\text{self}_x$ ] occurs that  $x$  is what [ $\text{self}_x$ ] picks out. Unlike the related proposal by Higginbotham (2003, 2010), neither tokens of ‘I’, nor their mental counterparts, contribute to thoughts the token-reflexive description *the producer/subject of  $t$* ; such a description merely “fixes their referent”, on a specific way of cashing this out.<sup>28</sup>

Let me now outline the account of IEM that this view affords. Campbell (1999, p. 90) advanced the main idea relative to descriptive names such as Evans’s ‘Julius’, whose reference is fixed by means of the description ‘the inventor of the zip’. On the application of the proposal above to names, the referent  $x$  of a proper name  $n$  is fixed by the semantically triggered presupposed condition that  $x$  is picked out by the naming practice associated with  $n$ ; in the assumed context, it will be further pragmatically presupposed that the relevant naming practice is one on which  $x$  is whoever invented the zip, if anybody uniquely did. As Campbell points out, on these assumptions the thought that *Julius invented the zip* is IEM, and the explanation is obvious: it cannot be (counterfactually, i.e., considering alternative worlds as actual, in the jargon of two-dimensional semantics; cf. García-Carpintero 2006) that somebody else invented the zip, because that is precisely what identifies Julius. The warrant for the existential generalization that someone invented the zip comes together indissolubly with a warrant for the corresponding ascription to a witness, because the identity of the witness is established precisely by assuming the truth of the generalization. The grounds we have do not offer a justification for the existential claim independent of the justification they offer for the singular claim; it cannot be that such grounds could secure knowledge of the former, even when the latter is false. We thus have an account of wh-immunity and thereby a way of dealing with the third complaint I made above about SA.

Campbell, however, argues that a proposal along these lines cannot be extended to other cases, but this is precisely what I purport to do; I’ll confront his argument against it below. Let us see first how the proposal deals with the first and second points above against SA. Consider for illustration a demonstrative case of IEM, ‘that keyboard is black’. In many cases (in core cases), fully understanding a demonstrative requires grasping from the context more reference-fixing information than just the one associated with the expression character. I have proposed to understand this by assuming that the character contains a determinable, *being the demonstrated entity*, to be determined in context. In the context assumed for the keyboard example, the further determination is given by perception; the demonstrated entity is the salient, perceptually available black keyboard. On the view outlined above, the asserted content is  $x$  is black, with the keyboard assigned to the variable, and it is presupposed that  $x$

<sup>28</sup> In addition to Higginbotham, Peacocke (1983, Chap. 5; 2008 Chap. 3; 2014, Chap. 2, 4), and Howell (2006) have also defended related views. I think my proposal improves on theirs in elaborating on the role the reference rule is supposed to play in characterizing the person level, conscious first-person awareness that the subjects of those states are supposed to have.

is the perceptually salient *black* keyboard when **that** is produced—a presupposition pragmatically, contextually triggered.<sup>29</sup> This is why the claim is a case of IEM. Perhaps we are having a hallucination and there is nothing that we are seeing; but it cannot be that something is in fact black, but it is not *that*. As before, the justification for the existential entailment of the judgment comes together with a justification for the ascription to a witness. Assuming that a structurally similar story works for the first-personal cases (and cases involving temporal or spatial demonstratives, ‘now’ and ‘here’), as I will argue below, we thus have a way to handle the first objection to SA. We have an explanation for the absence of an identification<sub>P</sub> component as substantive as the one afforded by Recanati’s Lewisian view of contents: the relevant cases are those where, in judging *x is F*, we think of *x* by way of a representation triggering the presupposition that what is being referred to as *x* is a thing that is F. Unlike Recanati’s, this account applies neatly to demonstrative cases.

In fact, the account just offered could be questioned, in a way that I now must examine; this will allow me to explain how the proposal deals with the *de facto/de jure* distinction and thus the second objection above to SA. Campbell (1997, p. 70) provides the following example: “if you judge, ‘that chair is yellow’, it may be that you thereby know of something that it is yellow, but that thing is not the chair, if, for instance, the chair is transparent and set against a yellow background. If your judgment is mistaken, you can rectify it by retreating to the more cautious, ‘At any rate, something is yellow’.” *Mutatis mutandis*, the point could be made about the ‘that keyboard is black’ example.

I suggest dealing with these cases on the model of presupposition *accommodation*. In some contexts, the fact that a speaker uses an expression that triggers a presupposition that is not yet common ground suffices (if the proposition is sufficiently uncontroversial) to make it so, so that the presuppositional requirement is fulfilled in this roundabout way after all: uttering ‘I cannot come to the meeting, I have to pick up my sister at the airport’ suffices for the proposition that the speaker has a sister to become common knowledge, although it would not have been otherwise.<sup>30</sup> Consider now a variation on a famous example by Donnellan (1966, p. 287). You say, ‘the man drinking martini is Bulgarian, I think’; your audience replies, ‘you are right, he is, but in fact he is a teetotaler drinking water in a martini glass’. Or someone asks, gesturing at your pet, ‘what do you call her?’, and you pointedly reply, ‘*he* is Toby’. Now, the pronouns in these examples are anaphoric on, respectively, ‘the man drinking martini’ and ‘she’. On the present account, this means that the determinable property *male being demonstrated* which figures in their respective reference-fixing presupposition is intended to be further determined by means of information associated with the expression on which they anaphorically depend. But how can that be? If what the speaker manifestly wants to communicate by using these tokens of ‘he’ is correct, they must fail to refer. In fact, the expressions on which they anaphorically depend must already fail to refer, assuming the reference-fixing presupposition account.

<sup>29</sup> Actually, this can at most be a case of *de facto* IEM; I am ignoring here possibilities that will be examined below, for the sake of keeping the presentation sufficiently simple.

<sup>30</sup> Cf. García-Carpintero (2015) for references and discussion.

Some writers have used cases like these to defend radical versions of *direct reference*—the view that reference does not depend on descriptive identification in any way (Martí 2008). However, as many others have pointed out (cf. Sosa 1995), the examples do not establish anything of the sort. The second speaker would not be able to get his point across if there was not an available fallback successful description which the first was also assuming to fix the referent; say, something such as *the person apparently drinking martini* in the first case, and *the female-looking pet* in the second. If, for instance, the first speaker in Donnellan’s example was truly confused, and there were no salient candidates for a perceived object, the communicative strategy of the second speaker could not be successfully implemented.

Given this, some writers would talk at this point of “presupposition cancellation”; the use of ‘he’ by the second speaker somehow revokes or annuls the first speaker’s presupposition that the person he was referring to was drinking a martini. Now, given that the relevant presuppositions are lexically triggered, I do not think they can be cancelled like conversational implicatures. I agree instead with Chierchia and McConnell-Ginet’s (1990, pp. 314–315) diagnosis of these cases: the presuppositions are (semantically speaking) still there; the speaker rhetorically contradicts them with the goal of challenging, and eventually changing, the contextual assumptions. In Donnellan’s case, the second speaker is conveying something like this: there is no man drinking martini to refer to, so let us pretend that you were not really asserting of  $x$  that  $x$  is Bulgarian, relative to the presupposition that  $x$  is the salient person drinking martini, but relative instead to the presupposition that  $x$  is the salient person appearing to drink martini. The second speaker brings to the fore the tacit presupposition that  $x$  is the salient person appearing to drink martini, ignoring the presupposition that  $x$  is the salient person drinking martini, and asserts that the first speaker is right that  $x$  is Bulgarian relative to that now highlighted presupposition.

I suggest using this model to understand our intuitions about examples such as Campbell’s. In an ordinary situation, the judgment expressed by ‘that chair is yellow’ (‘that keyboard is black’) is IEM. This is of course so relative to SA, given that the judgment does not seem to be based on any identification<sub>p</sub>; but it is also so given the previous account, for the demonstrative is a perceptual one whose reference is fixed relative to a perceptual experience identifying it as a salient yellow chair. Nonetheless, the situation that Campbell conjectures is an epistemic possibility. When considering it, we are assuming that “it” is not really yellow, but merely yellow-looking, illuminated by an independent yellow source. In so doing, we are on the above suggestion reconfiguring the assumptions that were being made before. We are dropping as a merely apparently correct assumption that it is the perceived yellow chair, and pretending instead that the correct reference-fixing assumption was that it is the yellow-looking chair.<sup>31</sup>

<sup>31</sup> I have mentioned above demonstrative cases as an advantage of the present proposal over Recanati’s. As a referee pointed out, however, Recanati’s view might deal better with some intuitions about the present case. If, in an abnormal situation, a speaker judges *that chair is yellow*, her presuppositions would be the standard ones, because she is unaware that anything is unusual. The present proposal declares the judgment a presupposition failure—and hence, on my view, truth-valueless. Still, one might have the intuition that the judgment is simply false, which could be easily explained on Recanati’s account. In my view, the intuition is explained by our tacitly moving to the alternative context I have described.



The epistemic possibility of the envisaged situation establishes that the original judgment is not *de jure* immune to wh-misidentification. It shows that the grounds we have might justify the existential judgment, that something is yellow, independently of the justification they provide for the singular judgment, that that chair is yellow: they might sustain knowledge of the existential judgment, even when the singular judgment is false. Only something like *that source of this chair-appearance looks yellow* is a candidate for being *de jure* immune to wh-misidentification. Parallel points could be made about a similar case that Pryor (1999, p. 297) discusses:

Suppose I have a visual experience that represents that a little girl is standing ten feet in front of me. Call this proposition L. Proposition L entails that there is a little girl somewhere or other in my spatial vicinity. Call this latter proposition L\*. My visual experience represents L as being the case; it does not merely represent L\* as being the case. The experience justifies me in believing both L and L\*. But it seems plausible that the justification the experience gives me for believing L can be undercut while still leaving me justified in believing L\*. For instance, someone might tell me that I'm in a hall of mirrors, and that there is a full-length mirror two feet in front of me. If I were to acquire that evidence, I would no longer be justified in believing that there is a little girl standing ten feet in front of me, but my visual experience would still justify me in believing that there is a little girl somewhere or other in my spatial vicinity.

In ordinary circumstances, the reference of the perceptual demonstrative 'there' in the judgment *a little girl is standing there* is fixed relative to the location of the little girl in the content of the perceptual experience. But we could easily consider the possibility that Pryor imagines. In doing so, we are assuming that the reference of 'there' is fixed instead relative to the assumption that it is where the little girl *appears to be*. The epistemic possibility shows, as before, that the judgment was not *de jure* immune to wh-misidentification.

Now, do cases like these point in the direction of the wiser form of skepticism about IEM that Howell (2007) suggests—namely, that it is only judgments concerning what we introspect (hence candidates for more traditional, Cartesian forms of epistemic privilege) that are IEM, i.e., judgments about how things (phenomenally) look, about sensations, thoughts or intentions, etc? I do not think they do. The issue depends on difficult epistemological and metaphysical matters about the relation between the judgments we are considering and the evidence on which they depend for reference-fixing purposes. The above discussion assumes that perceptual judgments depend on awareness of the phenomenal appearances of their objects. But such awareness might also depend, in its turn, on the appearances in question making external objects available for judgment, in normal conditions. It might be constitutive of the appearances we have in the yellow chair case that in normal circumstances they correctly integrate color and shape features into an accurate experience of the presence of a material object with such features.<sup>32</sup> There is no contradiction in assuming such reciprocal

<sup>32</sup> Cf. Campbell's (1997) discussion of the epistemic role of such integration of observable features. Burge (2003) and Schellenberg (2014) provide accounts of the epistemology of perceptual experience that would be helpful to develop a more detailed justification of the view suggested in the text. Both of them,

dependencies, even granting that ontological dependence is asymmetrical, for they are distinct: perceptual access to objects *specifically* depends on awareness of their phenomenal appearances; awareness of phenomenal appearances *generically* depends on (some of) them making objects perceptually available.<sup>33</sup>

If an account can be elaborated along such lines, it would vindicate *de facto* IEM as a genuine form of IEM. *De facto* IEM is, on the present proposal, IEM explained by reference-fixing presuppositions to which subjects are entitled, but which might fail under epistemically weird circumstances. Under such circumstances, they are not the true reference-fixing presuppositions, that role being taken on by more guarded assumptions involving the way things phenomenally look. *De jure* IEM is IEM that cannot fail under any epistemic circumstances. Under the assumptions sketched in the previous paragraph, *de facto* IEM stands in perfectly good stead; the demonstrative judgments that Campbell and Pryor consider, if made in normal circumstances as opposed to the abnormal ones they consider, would be IEM. I am therefore adopting a version of weak conservatism. In epistemically normal circumstances, the reference-fixing presupposition for ‘that chair’ is that it is the yellow chair currently perceived, and the identity of the yellow chair currently perceived with the source of the shape and color appearances is a presupposition to which the subject is entitled and on which the judgment need not be based for it to be justified. Under abnormal circumstances, there is no such entitlement and the reference-fixing presupposition need therefore be the more guarded one, securing *de jure* immunity.<sup>34</sup>

After this discussion of IEM in cases involving demonstratives, let us move to the first-person. On the present proposal, judgments such as ‘I am thinking this very thought’, and related *cogito* thoughts, are obviously *de jure* IEM, for the reference of ‘I’ is fixed relative to the judgments. When it comes to judgments like ‘I have toothache’, the reference of ‘I’ is also fixed relative to the phenomenal judgment, not to the experience. However, given *acquaintance* views on the nature of introspection such as the one carefully articulated by Gertler (2012), the fact that the subject of such judgment relative to whom it is fixed is introspecting the pain, given the nature of introspection, guarantees the inseparability of the justification for the existential judgment, that someone has toothache, and the singular judgment, that I have it.<sup>35</sup> Gertler (op. cit., p. 99) characterizes acquaintance views thus: some introspective knowledge consists in judgments that (1) are directly tied to their truthmakers; (2)

---

Footnote 32 continued

however, ignore in my view the constitutive role played by awareness of phenomenal features in a full account. I hope to elaborate on these matters in future work; an earlier version of these ideas can be found in García-Carpintero (2002).

<sup>33</sup> Fine (1995) articulates the distinction between those two forms of grounding or ontological dependence.

<sup>34</sup> What determines that circumstances are abnormal? I do not think it suffices that the subject considers the possibility that they are—she might be overreacting to skeptical considerations. Nor is it necessary that she does—she might be carelessly ignoring close possibilities. This is a difficult epistemological issue for which I cannot offer a helpful discussion here.

<sup>35</sup> The account of the *de se* I assume should thus be understood as “two-tiered” in the sense of O’Brien (2007, p. 65). Deploying representations whose significance is captured by the self-reference token-reflexive rule is a necessary *but insufficient* part of what thinking (and talking) about oneself as such is. The full account involves in addition (in the present case of the self-ascription of phenomenal states) a direct awareness of one’s own states, a self-knowledge that is not itself explained by the reliance on the self-reference rule.

depend, for their justification, only on the subject's conscious states at the time of the judgment; and (3) are more strongly justified than any empirical judgments that do not meet conditions (1) and (2). The *directness* mentioned in (1) consists in that "metaphysical reality and the epistemic intersect" in such judgments, a metaphor that Gertler (op. cit., pp. 104–109) elaborates thus. Firstly, the relevant mental features (a phenomenal type, an instance thereof) constitutively (and not merely causally) contribute to the content of the judgment: the judgment is about them in virtue of a direct, constitutive relation between the mental reality and the phenomenal judgment. Secondly, the way things *epistemically* seem to the subject is constituted by the way they *phenomenally* seem to her.

Now, let us assume that the case we are considering is a witness for the existential generalization constituting the acquaintance claim: the pain is not borderline, it occurs concurrently with the judgment, it fully occupies the subject's attention. In such a case, a subject with the *self<sub>x</sub>* concept would also self-ascribe it; the judgment is not just *pain*, but something like *pain in me now*. This manifests a presupposition of *de jure* coreference (more on this below) between the subject to which the judgment is ascribed, and the subject of the pain. In other words, the application of the *self<sub>x</sub>* concept in the judgment is not just opportunistic, one-off, but the information is added to a "mental file" (which I take to be constituted by such relations of *de jure* coreference)<sup>36</sup> to which the information that the subject instantiates the thereby introspected phenomenal features is also added. The presuppositions thereby manifested are also reference-fixing, like corresponding ones in the case of perceptual demonstratives, for they essentially contribute to filling up who is the subject to which the introspective judgment is ascribed: it is also the one instantiating the thereby introspected phenomenal features. These presuppositions are ones to which the subject is entitled, as the acquaintance account of introspection justifies; for in such judgments the subject is directly presented with his own mental features.<sup>37</sup>

Given plausible views about the relation between a subject and her body, which (following work by Shoemaker) Peacocke (2014, pp. 86–99, 106–113) has spelled out, judgments such as *I see a canary in front of me, there is a canary in front of me, I am raising my arm*, when judged "from the inside" on the basis of introspective awareness of our bodily or visual experiences, or actions, are *de facto* IEM. The main assumption here is that one's body is the body from which one normally perceives, receives bodily information, and executes one's basic actions. This again grounds a defeasible reference-fixing presupposition, manifested in assumptions of *de jure* coreference, which justifies along the previous lines the IEM of examples such as those.<sup>38</sup> It is

<sup>36</sup> Cf. García-Carpintero (forthcoming) for some discussion and references.

<sup>37</sup> Cf. Peacocke (2014), pp. 110–111 for a similar explanation, assuming a different framework on introspective judgments. It should be clear why, on such accounts, the possibility that Cappelen and Dever (see above, fn. 13) casually toss about does not exist.

<sup>38</sup> As indicated above, the present proposal—like Peacocke's (1983), Coliva's (2006) and Wright's (2012)—adopts *weak conservatism*: a non-inferential judgment might be vulnerable to EM if it depends on a "background identification presupposition" to which in the circumstances one is not entitled. Morgan (ms) and McGlynn (2015) forcefully argue that, unembellished, such accounts would make the extension of IEM too narrow. For, even if skeptics such as Cappelen and Dever (2013) are wrong in their objections to IEM, the issues they raise might be enough to lead rational subjects to withhold judgments such as 'I

merely *de facto* in this case, because the connection between the mental states and the judgment lacks the directness of introspective judgments, as the conceivability of Davidson's eyes or *Avatar* situations witnesses. They establish the independence of the justification we have in those circumstances for the existential judgments (say, that someone has his legs crossed, given by proprioception) from the one we have for the singular ones (that my legs are crossed).<sup>39</sup>

As I said, my proposal is similar to Howell's (2006) and Peacocke's (2008, 2014) in assigning a mere reference-fixing role to the token-reflexive descriptions which it associates with the relevant referential expressions and which it invokes to account for IEM. In this respect, it differs from Higginbotham's (2003, 2010). In the companion piece (García-Carpintero forthcoming) I invoke this fact to show how my proposal is not subject to a compelling objection by Recanati (2007), based on cases of thought insertion. Morgan (2012, pp. 108–111) has another interesting objection to Higginbotham's account of IEM that can be dealt with also in part on the basis of the indicated difference between the two proposals. The objection is that, as we have seen above, IEM is basis-relative: the same judgment may be and may not be IEM, depending on its epistemic grounds. On Higginbotham's account the relativity is lost, because IEM is explained by the content's token-reflexive character—a semantic feature, independent of the nature of the judgment's epistemic grounds. In the reference-fixing version of the token-reflexive account the basis-relativity lies in whether the relevant evidence plays a reference-fixing role in the elaboration of the determinable aspects of the more abstract reference-fixing rule. If I judge 'I have toothpaste on my forehead' on the basis of proprioceptive and haptic information (the sensation of toothpaste on one's skin), the evidence plays a reference-fixing role, relative to the considerations about the role of one's bodily sensations in determining which body is ours gestured at above. If I judge that on the basis of seeing what I take to be myself in a mirror, the evidence does not play such a reference-fixing role.

I will conclude by discussing Campbell's (1999, p. 94) argument against extending the reference-fixing model beyond the 'Julius' case. The problem, he claims, is that the model

... generates a quite spurious problem. If "I" is glossed as "whoever is having this experience of hearing trumpets" to explain the immunity to error through

---

Footnote 38 continued

am in pain' (at least in fallible, borderline cases) based on introspection, or even 'I am judging this very thought' (perhaps by developing "thought-insertion" misgivings). I think these arguments do not affect the present proposal, because it does not appeal to any old assumption that might counterfactually be doubted; it appeals to assumptions that are requirements given by the very nature of the representational system about who we are and which body is our own. The fact that a thinker who under ordinary circumstances judges *I am in pain* might be (mis-)led to worry whether the pain is hers does not make it any less IEM.

<sup>39</sup> I thus side with Shoemaker and Pryor against Evans. Morgan (ms) and McGlynn (2015) opt for the opposite side. I cannot discuss their interesting considerations here. Let me just say that I cannot see any relevant difference between, on the one hand, the possibilities that Campbell and Pryor imagine for perceptual demonstratives, and on the other those contemplated in the main text for perceptual access to one's body, or by Shoemaker for memory judgments—other than the currently easier factual availability of the former situations.

misidentification of “I hear trumpets,” then “I” will have to be glossed as “whoever is having this headache” in order to explain the immunity to error through misidentification of “I have a headache.” At this point we have generated a problem: how do I know that it is the same person that is having this experience of hearing trumpets as is having this headache? Until the problem is resolved, I have no right to use “I” in both cases as referring to the same object. But once the problem is allowed to arise, there is no way of resolving it. And the problem is entirely spurious: the reason it does not arise ordinarily is that I know perfectly well that I am hearing the trumpets and I have the headache.

I agree that the account generates the problem that Campbell mentions. Unlike him, I think the problem is not spurious, but inescapable. I also disagree with him that once it is posed it cannot be solved. Peacocke (2012, p. 79) explicitly considers it and provides the same answer I would. Campbell can be so sure that one knows perfectly well that one is hearing trumpets and one is having a headache, because the subjects of those judgments are in relations of *de jure* coreference, like subjects inferring from the two premises *Pavarotti recorded Othello*, *Pavarotti sang yesterday*, that *someone who sang yesterday recorded Othello*. Writers pose mental files to account for such relations; Peacocke (2014, pp. 14–15, 25–29, 86–99) elaborates in more recent work on how such a model applies to the first person. Any account that works for the general case should also handle the problem that, Campbell thinks, the application of the reference-fixing model spuriously creates.

In this paper I have argued that Evans’s, Peacocke’s and Wright’s “Simple Account” of IEM is not adequate. After presenting the phenomenon and the account in the first section, I raised three worries in the second, suggesting that an adequate account of IEM should be more substantive—should be one relying on a substantive account of *de se* thoughts. In the final section I have indicated how a Fregean account of such thoughts, based on a more general account of singular reference, is in a position to dispel these concerns. Here I have merely outlined a proposal, which certainly needs much more defense and elaboration. I hope it contributes to sharpening the worries about the Simple Account raised in the previous section, by showing how a specific substantive account might deal with them.

**Acknowledgments** Financial support for my work was provided by the DGI, Spanish Government, research project FFI2013-47948-P and Consolider-Ingenio project CSD2009-00056; and through the award ICREA *Academia* for excellence in research, 2013, funded by the Generalitat de Catalunya. Previous versions of this paper were given at talks at: ESPP 2013, Granada; Semantic Content, Barcelona 2013; PLM 2, Budapest 2013; Reference, Indexicality and Singular Thought, Campinas; Journée ‘autour du travail de François Récanati’, Paris 2014; Self-knowledge, Expression and Transparency, Murcia 2014; ESPP 2014, Oslo; CSMN, Oslo; Workshop on the First Person Perspective, Athens 2015. Thanks to the audience for useful questions and remarks, and in particular to Gregory Bochner, Herman Cappelen, Annalisa Coliva, Josh Dever, Jose Díez, Bjørn Jespersen, Alisa Mandrigin, Giovanni Merlo, Daniel Morgan, Dilip Ninan, Ernest Sosa, Stephan Torre, and Clas Weber for very helpful comments. Special thanks are due to Marie Guillot, who provided very detailed terrific comments on the penultimate version. Thanks also to Michael Maudsley for the grammatical revision.

### Compliance with Ethical Standards

**Conflicts of interest** The author declares that he has no conflict of interest.

## References

- Audi, R. (1986). Belief, reason and inference. *Philosophical Topics*, 14, 27–65.
- Burge, T. (2003). Perceptual entitlement. *Philosophy and Phenomenological Research*, 67, 503–548.
- Campbell, John. (1997). Sense, Reference and Selective Attention. *Proceedings of the Aristotelian Society*, *sup.*, 71, 55–98.
- Campbell, J. (1999). Immunity to error through misidentification and the meaning of a referring term. *Philosophical Topics*, 26, 89–104.
- Campbell, J. (2002). The ownership of thoughts. *Philosophy, Psychiatry and Philosophy*, 9, 35–39.
- Cappelen, H., & Dever, J. (2013). *The inessential indexical*. Oxford: Oxford University Press.
- Castañeda, H. (1966). ‘He’: A study in the logic of self-consciousness. *Ratio*, 8, 130–157.
- Castañeda, Hector. (1968). On the Logic of Attributions of Self-Knowledge to Others. *Journal of Philosophy*, 65, 439–456.
- Chalmers, D. (2012). *Constructing the world*. Oxford: Oxford University Press.
- Chierchia, G., & McConnell-Ginet, S. (1990). *Meaning and grammar*. Cambridge, MA: MIT Press.
- Coliva, A. (2002a). Thought insertion and immunity to error through misidentification. *Philosophy, Psychiatry and Philosophy*, 9, 27–34.
- Coliva, A. (2002b). On what there really is to our notion of ownership of a thought. *Philosophy, Psychiatry and Philosophy*, 9, 41–46.
- Coliva, A. (2006). Error through misidentification: Some varieties. *Journal of Philosophy*, 103, 403–425.
- Coliva, A. (2012). Which ‘key to all mythologies’ about the self? A note on where the illusions of transcendence come from and how to resist them. In S. Prosser & F. Recanati (Eds.), *Immunity to error through misidentification: New essays* (pp. 22–45). Cambridge: Cambridge University Press.
- Dennett, D. (1978). “Where am I?”, in *his brainstorms*. Cambridge, MA: MIT Press.
- Donnellan, K. (1966). Reference and definite descriptions. *Philosophical Review*, 75, 281–304.
- Evans, G. (1982). *The varieties of reference*. Oxford: Clarendon Press.
- Fernández, J. (2010). Thought insertion and self-knowledge. *Mind and language*, 25, 66–88.
- Fine, K. (1995). Ontological dependence. *Proceedings of the Aristotelian Society*, 95, 269–290.
- Fine, K. (2007). *Semantic Relationalism*. Oxford: Blackwell.
- García-Carpintero, M. (2000). A presuppositional account of reference-fixing. *Journal of Philosophy* *XCVII*, 3, 109–147.
- García-Carpintero, M. (2002). Sense-data: The sensible approach. *Grazer Philosophische Studien*, 62, 17–63.
- García-Carpintero, M. (2006). Two-dimensionalism: A Neo-Fregean interpretation. In M. García-Carpintero & J. Macià (Eds.), *Two-dimensional semantics* (pp. 181–204). Oxford: Oxford University Press.
- García-Carpintero, M. (2013). The SELF file and immunity to error through misidentification. *Disputatio (Portugal)*, 36(2013), 191–206.
- García-Carpintero, M. (2015). Accommodating presupposition. *Topoi.* doi:10.1007/s11245-014-9264-5.
- García-Carpintero, M. (forthcoming): Token-reflexive Presuppositions and the De Se. in García-Carpintero, M. and Torre, S. (eds.), *About Oneself*, Oxford: Oxford University Press.
- Gertler, B. (2012). Renewed acquaintance. In D. Smithies & D. Stoljar (Eds.), *Introspection and consciousness* (pp. 93–127). Oxford: Oxford University Press.
- Grush, R. (2007). Evans on identification-freedom. *Canadian Journal of Philosophy*, 37, 605–618.
- Heck, R. G., Jr. (2012). Solving’s Frege’s puzzle. *Journal of Philosophy*, 109, 132–174.
- Heim, I. (2008). Features on bound pronouns. In D. Harbour, D. Adger, & S. Béjar (Eds.), *Phi Theory*. Oxford: OUP.
- Higginbotham, J. (2003). Remembering, imagining, and the first person. In A. Barber (Ed.), *Epistemology of language* (pp. 496–533). Oxford: Oxford University Press.
- Higginbotham, J. (2010). On words and thoughts about oneself. In I. Stojanovich, F. Recanati, & N. Villanueva (Eds.), *Context-dependence, perspective, and relativity* (pp. 253–282). Berlin: Mouton de Gruyter.
- Howell, R. (2006). Self-knowledge and self-reference. *Philosophy and Phenomenological Research*, 72, 44–69.
- Howell, R. (2007). Immunity to error and subjectivity. *Canadian Journal of Philosophy*, 37, 581–604.
- Hunter, J. (2013). Presuppositional indexicals. *Journal of Semantics*, 30(3), 381–421. doi:10.1093/jos/ffs013.
- Kripke, S. (1980). *Naming and necessity*. Cambridge, MA: Harvard University Press.

- Lane, T., & Liang, C. (2011). Self-consciousness and immunity. *Journal of Philosophy*, 2011, 78–99.
- Lewis, D. (1979). Attitudes de dicto and de se *Philosophical Review*, 88, 513–43. Also in D. Lewis, *Philosophical Papers*, vol. 1, Oxford: Oxford University Press.
- Maier, E. (2010). Presupposing acquaintance: A unified semantics for de dicto, de re and de se belief reports. *Linguistics and Philosophy*, 32(5), 429–474.
- Maier, E. (forthcoming). Why my I is your you: On the communication of de se attitudes. In M. García-Carpintero & S. Torre (Eds.), *About oneself*. Oxford: Oxford University Press.
- Martí, G. (2008). Direct reference and definite descriptions. *Dialectica*, 62, 43–57.
- McGlynn, A. (2015). Immunity to error through misidentification and the epistemology of de se thought. In M. García-Carpintero & S. Torre (Eds.), *About oneself*. Oxford: Oxford University Press.
- Morgan, D. (2012). Immunity to error through misidentification: What does it tell us about the de se. In S. Prosser & F. Recanati (Eds.), *Immunity to error through misidentification: New essays* (pp. 104–123). Cambridge: Cambridge University Press.
- Morgan, D. (ms). Thinking about the body as subject.
- Neta, R. (2011). Liberalism and conservatism in the epistemology of perceptual belief. *Australasian Journal of Philosophy*, 88, 685–705.
- O’Brien, L. (2007). *Self-knowing agents*. Oxford: OUP.
- Peacocke, C. (1981). Demonstrative thought and psychological explanation. *Philosophical Studies*, 49, 187–217.
- Peacocke, C. (1983). *Sense and content. Experience, thought, and their relations*. Oxford: Clarendon Press.
- Peacocke, C. (1999). *Being known*. Oxford: Oxford University Press.
- Peacocke, C. (2008). *Truly understood*. Oxford: Oxford University Press.
- Peacocke, C. (2012). Subjects and consciousness. In A. Coliva (Ed.), *The self and self-knowledge* (pp. 74–101). Oxford: Oxford University Press.
- Peacocke, C. (2014). *The mirror of the world*. Oxford: Oxford University Press.
- Perry, J. (1979). The problem of the essential indexical. *Noûs* 13, 3–21. Also in his the problem of the essential indexical and other essays (1993, pp. 33–50, from which I quote). Oxford: Oxford University Press.
- Pryor, James. (1999). Immunity to error through misidentification. *Philosophical Topics*, 26, 271–304.
- Pryor, J. (2004). What is wrong with Moore’s argument? *Philosophical Issues*, 14, 349–378.
- Recanati, F. (2007). *Perspectival thought*. Oxford: Oxford University Press.
- Recanati, F. (2009). De re and de se. *Dialectica*, 63, 249–269.
- Recanati, F. (2012). Immunity to error through misidentification: What it is and where it comes from. In S. Prosser & F. Recanati (Eds.), *Immunity to error through misidentification: New essays* (pp. 180–201). Cambridge: Cambridge University Press.
- Recanati, F. (2013). Mental files: Replies to my critics. *Disputatio (Portugal)*, 36(2013), 207–242.
- Schellenberg, S. (2014). The epistemic force of perceptual experience. *Philosophical Studies*, 170, 87–100.
- Shoemaker, S. (1968). Self-knowledge and self-awareness. *Journal of Philosophy*, 65, 555–567.
- Shoemaker, S. (1970). Persons and their pasts. *American Philosophical Quarterly*, 7, 269–285.
- Shoemaker, S. (1994). Self-knowledge and ‘inner sense’. *Philosophy and Phenomenological Research*, 54, 249–315.
- Smith, J. (2006). Which immunity to error? *Philosophical Studies*, 130, 273–283.
- Sosa, E. (1995). Fregean reference defended. In E. Villanueva (Ed.), *Content. Philosophical issues* (pp. 91–99). Atascadero, CA: Ridgeview Publishing Company.
- Stalnaker, R. (1978). Assertion. In P. Cole (Ed.), *Syntax and semantics* (Vol. 9, pp. 315–332). New York: Academic Press.
- Stalnaker, R. (2014). *Context*. Oxford: Oxford University Press.
- Stanley, J. (2011). *Know how*. Oxford: Oxford University Press.
- Wittgenstein, L. (1958). The blue and brown books, ed. R. Rush, New York: Harper and Row.
- Wright, C. (1998). Self-knowledge: The wittgensteinian legacy. In C. Wright, B. C. Smith, & C. McDonald (Eds.), *Knowing our own minds* (pp. 13–45). Oxford: Clarendon Press.
- Wright, C. (2004). Warrant for nothing (and foundations for free)? *Proceedings of the Aristotelian Society, Supplementary*, 78, 167–212.
- Wright, C. (2012). Reflections on François Recanati’s ‘immunity to error through misidentification: What it is and whereit comes from’. In S. Prosser & F. Recanati (Eds.), *Immunity to error through misidentification: New essays* (pp. 247–280). Cambridge: Cambridge University Press.