

Hazel Pearson  
University of Tübingen/ZAS, Berlin  
[pearson@zas.gwz-berlin.de](mailto:pearson@zas.gwz-berlin.de)

December 2015

# Counterfactual De Se

## Abstract

This paper addresses a long-standing debate concerning the derivation of de se construals. One camp holds that there are ‘dedicated de se LFs’, characterized by a lambda abstractor in the embedded clause that binds the de se pronoun (Chierchia 1990). Another treats de se as a special case of de re under a SELF acquaintance relation (Lewis 1979; Reinhart 1990; Maier 2006, 2010, 2011). This debate is premised on the assumption that the two different routes to de se result in identical truth conditions. I argue that this assumption is incorrect for a class of cases that can be delineated in a principled fashion - counterfactual attitude reports involving counter-identity, such as ‘Sophia imagined that she was Michelle Obama and she lived in the White House’. The argument builds on Ninan (2008), who noticed that de re construal works differently with counterfactual attitudes, and that this has consequences for de se interpretation in this type of sentence. I develop a strong version of this insight, and show that the de se-as-de re route cannot in principle yield a construal that is equivalent to that obtained via binding in counterfactual reports with counter-identity; dedicated de se LFs are indeed needed. This view casts doubt on the potential of de se-as-de re analyses of control to capture the unambiguously de se nature of PRO, as recently proposed in (Maier 2011; Landau 2015a,b).

## 1. Introduction

Here is a story that was told to me by a friend who wishes to remain anonymous.

1. One morning I was sitting with friends in the local McDonald's, as was our custom, prior to heading up the hill to school. The corner in which we frequently perched ourselves -- a spot, enclosed on two sides by wall-high glass, that allowed us to gaze at the antics of pedestrians outside -- this corner was today already taken, and so we grumpily found another table to congregate around nearer the middle of the establishment. While my friends were talking amongst themselves, I found myself momentarily distracted by a man sitting a short distance away, staring roughly back at me. For reasons I can't define, I took an immediate dislike to the appearance of this man, his face and countenance, and came to the swift judgment that he was some form of loathsome idiot. It was startling, therefore, a short moment later, to realize that I was gazing into a wall-high mirror -- and had thus cast such an aspersion upon myself.

In this story, there is a certain individual whom our friend – call him Jonathan – believes to

be a loathsome idiot (at least up until the moment when he realizes that he is looking in a mirror). Who is that individual? We know, even though Jonathan has not yet figured it out, that it is Jonathan himself. So can we then describe what happened with the following sentence?<sup>1</sup>

2. Jonathan<sub>i</sub> believed that he<sub>i</sub> was a loathsome idiot.

Yes and no. On the one hand, it seems that we can coherently utter (3).

3. Jonathan<sub>i</sub> believed that he<sub>i</sub> was a loathsome idiot, although he didn't realize that he was the person that he had in mind.

But we could also say:

4. Jonathan<sub>i</sub> didn't believe that he<sub>i</sub> was a loathsome idiot, because he didn't realize that he was the person that he had in mind.

In the second discourse, the fact that Jonathan didn't realize that the person he was ascribing loathsome idiocy to was himself is grounds to reject (2); in (3), (2) is taken to be true *despite* this fact. It seems then that (2) has two interpretations: one on which Jonathan's failure to recognize himself is relevant to the truth value of the sentence, and one on which it is not. On the former reading, the sentence is judged false, and on the latter one it is judged true.

The difference seems to come down to the interpretation of the pronoun: on the former reading, we shall say that the pronoun is read *de se*, while on the latter one it is read *de re*. A *de se* report of an attitude (a belief, desire, speech act, etc.) is a report of an attitude that in some intuitive sense is 'about' the attitude holder, and which furthermore the attitude holder is aware is about himself. Jonathan's belief that the person he is looking at is a loathsome idiot has the first property, since the person Jonathan is looking at is none other than Jonathan himself. However it lacks the second property, since Jonathan does not know that he is looking in a mirror. Attitudes that have both properties – that is, that satisfy both the 'aboutness' condition and the 'awareness' condition – are attitudes *de se*. If a pronoun is construed *de se*, then the attitude report in which it occurs is only true if both conditions are satisfied, and the reported attitude is therefore an attitude *de se*. Hence (2) is judged false on the *de se* construal: Jonathan's belief of himself that he is a loathsome idiot does not satisfy the awareness condition. On the other hand, *de re* beliefs that are about the attitude holder are not subject to the awareness condition; hence if the pronoun is construed *de re*, then we judge the sentence true.

The topic of this paper is the proper analysis of *de se* reports. Beginning with Chierchia (1990), some researchers have argued that there are 'dedicated *de se* LFs': syntactic structures that are assigned *de se* truth conditions (eg Anand 2006; Percus and Sauerland 2003a,b; von Stechow 2002, 2003). On this view, (2) is structurally ambiguous: the string is compatible both with a *de se* LF, and with one that does not assign *de se* truth conditions. On the other hand, Lewis (1979) noticed that *de se* construal could be derived via *de re* construal of the pronoun under a special acquaintance relation which we might call 'SELF'. Reinhart (1990)

---

<sup>1</sup> I use coindexing merely to indicate intended coreference; the indices are not meant to have theoretical status.

and subsequently (Maier 2006, 2010, 2011) argued that this ‘de se via de re’ route is sufficient: de se construal can be derived without appeal to dedicated de se LFs, which therefore should not be postulated.

Throughout this debate, it has been taken for granted that the two ‘routes to de se’ result in equivalent truth conditions: the interpretation that the semantics would assign to a de se LF is identical to that which would result from de re construal under the SELF acquaintance relation. I shall argue that this assumption is incorrect for a particular class of environment – namely reports of an attitude holder counterfactual supposing that she is someone other than who she actually is. Such cases arise with counterfactual attitude verbs such as *imagine*, *wish* and *pretend*, as in the following examples:<sup>2</sup>

- 5a. Sophia is imagining that she is Michelle Obama and **she** is married to Barack Obama.
- 5b. Sophia is wishing that she is Michelle Obama and **she** is married to Barack Obama.
- 5c. Sophia is pretending that she is Michelle Obama and **she** is married to Barack Obama.

On the most salient construal of the bolded pronoun in (5a), Sophia is not imagining that *Sophia* is married to Barack Obama. This would require her to imagine that Barack Obama has different properties than he actually has with respect whom he is married to. Rather, the counterfactual worlds that Sophia is entertaining resemble the actual world (and resemble the worlds that Sophia’s belief state designates as candidates for the actual world) in as much as in those worlds too, Barack Obama is married to Michelle Obama. What is different is that according to what Sophia imagines, she *is* Michelle Obama. What it could mean for Sophia to be Michelle Obama is a vexed topic – surely there are no worlds in which this is the case? – and we shall be returning to it shortly. For now, I shall just assume that (5a) reports an event involving Sophia putting herself into Michelle Obama’s shoes, say because she is wondering what it would be like to be her.<sup>3</sup> This captures our intuitions about the meaning of the sentence well enough for now.

I shall claim that this construal of the bolded pronoun in (5a) can only be captured by a dedicated de se LF; de se as a special case of de re cannot yield the intended interpretation for principled reasons having to do with the semantics of counterfactual de re (Ninan 2008, 2012). Therefore, there are dedicated de se LFs.

The paper is structured as follows. Section 2 describes the two theories of de se construal that are the focus of this paper. In Section 3, I discuss in greater detail the environments in which de se via binding and de se as a special case of de re are argued not to be equivalent – namely, reports of counterfactual attitudes with counter-identity. In Section 4, I consider three possible replies that might be available to proponents of the de se-as-de re view who wish to eliminate de se binding, and demonstrate that none of these succeeds in generating an interpretation that is equivalent to that obtained by appealing to a dedicated de se LF. This section includes new data regarding the interpretation of PRO in obligatory control –

---

<sup>2</sup> By ‘counterfactual attitude verb’ I mean a verb denoting an attitude that one can hold coherently towards a content while simultaneously believing that that content is false.

<sup>3</sup> Notice incidentally that Sophia can only put *herself* in someone else’s shoes – she cannot put her friend Billy into Michelle Obama’s shoes. Is this merely a linguistic quirk of the locution *put x in x’s shoes*? I don’t think so, for reasons having to do with data discussed in section 5.2.

an element that is well known to be unambiguously interpreted de se (Morgan 1970). Since Chierchia (1990) this fact has usually been handled by appeal to de se binding of PRO; more recently, Maier (2011) and Landau (2015a,b) have analyzed the de se construal of PRO as arising as a special case of de re. The new data presented in Section 4 are difficult to accommodate within this type of view, and instead argue for a return to a de se binding analysis of PRO. Section 5 shows that the resulting picture provides a solution to a long-standing puzzle about so-called ‘de re blocking effects’, and considers a potential challenge from data involving counter-identity about individuals other than the attitude holder. Section 6 is the conclusion.

## 2. Theories of de se construal

In this section, I describe in greater detail the two different approaches to de se.

### 2.1 De se via binding

The ‘dedicated de se LFs’ postulated by Chierchia and subsequent authors involve binding of the de se pronoun by an abstractor in the left periphery of the complement clause. We shall assume an extensional framework where worlds are represented in the syntax as unpronounced indexed elements (pronouns) that are bound by co-indexed abstraction operators higher in the structure. Thus the LF for (2) (on its de se reading) is as in (6).

6. Jonathan believes  $[_{CP} \lambda x_1 \lambda w_2 [w_2 \text{ he}_1 \text{ is a loathsome idiot}]]$ .

This provides an implementation of Lewis’s (1979) idea that to believe de se that you are P is to self-ascribe the property P. If I believe that I am looking at a beautiful view right now, then I self-ascribe the property of looking at a beautiful view. In our story, Jonathan did not self-ascribe the property of being a loathsome idiot; rather he ascribed it to the individual that he was looking at, not realizing that that individual was him.

In (6), binding of the pronoun results in the CP being interpreted as a property rather than a proposition:

7.  $[[CP]]^{c,g} = \lambda x \lambda w. x \text{ is a loathsome idiot in } w$ .

The standard truth conditions of attitude reports require that the embedded clause express a function that when applied to each of the attitudinal alternatives of the attitude holder returns the truth value 1. If the embedded clause in a de se report is a function from individuals to functions from worlds to truth values (type  $\langle s, \langle e, t \rangle \rangle$ ), then attitudinal alternatives cannot be worlds, as on a Hintikka semantics (Hintikka 1969), but rather world-individual pairs – so-called ‘centred worlds’. The attitudinal alternatives for de se belief, for example (*doxastic* alternatives), can be defined as follows:

8.  $\text{Dox}_{x,w} = \{ \langle w', y \rangle : \text{it is compatible with what } x \text{ believes in } w \text{ for } x \text{ to be } y \text{ in } w' \}$ .

We correspondingly have the semantics for *believe* given in (9), and the truth conditions for the de se reading of (2) given in (10):

9.  $[[\text{believe}]]^{c,g} = \lambda P: P \in D_{\langle e, \langle s, t \rangle \rangle} . \lambda x: x \in D_e . \lambda w: w \in D_s . \forall \langle w', y \rangle \in \text{Dox}_{x,w} , P(y)(w')$

10.  $[[\text{(2)}]]^{c,g} = \lambda w. \forall \langle w', y \rangle \in \text{Dox}_{\text{Jonathan},w} , y \text{ is a loathsome idiot in } w'$ .

According to (10), for every world-individual pair  $\langle w', y \rangle$  such that for all Jonathan believes, he could be  $y$  in  $w'$ ,  $y$  is a loathsome idiot in  $w'$ . The individual coordinate thus ranges over those inhabitants of Jonathan's belief worlds that he considers to be candidates for himself. Since in our story, Jonathan does not ascribe the property of being a loathsome idiot to any  $y$  that is a candidate of Jonathan's for himself, the sentence is correctly predicted to be false on the de se reading.

## 2.2 De se as a special case of de re

An alternative strategy to the one described above is to ask whether de se construal can be derived via some other mechanism that is independently required.<sup>4</sup> The apparatus involved in the derivation of truth conditions for de re belief reports looks like an excellent candidate. Remember Quine's old example:

- 11a. Ralph believes that Ortcutt is a spy.  
11b. Ralph believes that Ortcutt is not a spy.

If we heard these sentences out of the blue and took them to be true, then we would conclude that Ralph is irrational: his belief state seems to ascribe incompatible properties to the same individual. But Quine noticed that in the following scenario, the two sentences are true even though Ralph is not irrational:

12. 'There is a certain man in a brown hat whom Ralph has glimpsed several times under questionable circumstances on which we need not enter here; suffice it to say that Ralph suspects he is a spy. Also there is a grey-haired man, vaguely known to Ralph as rather a pillar of the community, whom Ralph is not aware of having seen except once at the beach. Now Ralph does not know it but the men are one and the same [Bernard Ortcutt].'

[(Quine, 1956): 56]

Intuitively, (11a) is true in this situation because (i) Ralph believes that the person that he saw in the brown hat is a spy and (ii) that person is Ortcutt. (11b) is true because (i) Ralph believes that the person that he saw at the beach is not a spy and (ii) that person is Ortcutt. These facts can hold without Ralph being irrational because he does not believe that the person that he saw in the brown hat is identical to the person that he saw at the beach. Roughly speaking, 'the person Ortcutt saw in the brown hat' and 'the person Ortcutt saw at the beach' pick out distinct individuals at each of Ralph's belief worlds, even though they

---

<sup>4</sup> In this paper, I focus on the alternative strategy that I will refer to as 'de se as a special case of de re', as it has been pursued in the linguistic literature. However, there is a close connection between this tradition and a philosophical tradition of 'de se skepticism' that denies that putative puzzles about de se are anything other than instances of a broader phenomenon of substitution failure in attitudinal environments (Boër and Lycan 1980; Cappellen and Dever 2013; Magidor 2014).

pick out the same individual in the actual world. Thus Ralph's belief state ascribes the incompatible properties of being a spy and not being a spy to distinct individuals; perfectly coherent.

According to this account, the truth conditions of the reports of Ralph's de re belief about Ortcutt are mediated by the kinds of objects that serve as the semantic value of definite descriptions – namely, individual concepts. (For now, we can just think of individual concepts as functions from worlds to individuals, although that assumption will be amended in a moment.) The concepts  $C$  that play this mediating role must be 'acquaintance-based' in that there must be an acquaintance relation  $R$  such that for every world  $w'$  in the domain of the concept,  $R$  holds in  $w'$  between the attitude holder and  $C(w')$ .<sup>5</sup> Thus the concept expressed by 'the person Ralph saw at the beach' is based on the acquaintance relation 'see at the beach'; that expressed by 'the person Ralph saw in the brown hat' is based on the acquaintance relation 'see in the brown hat' and so on.

Reinhart's idea, building on Lewis, was that de se reports are de re reports about the attitude holder that are mediated by that concept that is based on the acquaintance relation SELF. In order to see how this works, we firstly need to flesh out the truth conditions for de re belief reports.<sup>6</sup>

The first task is to determine what the attitudinal alternatives for de re belief are. One might have hoped (as Reinhart seems to have done) that one could revert to treating these as worlds, rather than appealing to the centred worlds discussed in the previous sub-section. But this will not do, for reasons discussed in (Anand 2006). In Quine's scenario, the belief of Ralph's in virtue of which (11a) is true is not 'the man Ralph saw at the beach is a spy', but rather 'the man *I* saw at the beach is a spy'. If Ralph did indeed see Ortcutt at the beach, and in addition he believes (de se) that he is Ronald and that some other guy is Ralph, and he says sincerely 'the man Ralph saw at the beach is a spy', then we are clearly not entitled to conclude that Ralph believes that Ortcutt is a spy. This shows that in Ralph's belief worlds  $w'$ , it is not Ralph himself who bears the acquaintance relation to the individual returned by the concept in  $w'$ , but rather those individuals that Ralph is prepared to designate with the first person pronoun. Such individuals are just Ralph's candidates for himself – precisely those individuals designated by the individual coordinate of a centred world.

What is needed is to let the domain of acquaintance-based concepts be world-individual pairs, not worlds. We can then say that if  $C$  is an acquaintance-based concept for some attitude holder  $x$ , then there is an acquaintance relation  $R$  such that for every world-individual pair  $\langle w', y \rangle$  that is in the domain of  $C$ ,  $R$  holds between  $y$  and  $C(w', y)$ . If we are

---

<sup>5</sup> I won't have much to say about what exactly acquaintance relations are. Let's just assume that they are perceptual relations holding between attitude holders and individuals, which however may be relatively indirect (not only 'saw at the beach', but also 'heard about from a friend', 'read a paper by', and what have you). Underlying the approach is the assumption that you can only have a de re belief about an individual that you stand in some acquaintance relation to, and that this is reflected in the truth conditions of de re belief reports. Kaplan (1968) also assumed that the acquaintance relations must be sufficiently 'vivid', however I will set that aside.

<sup>6</sup> It will suffice for the purposes of this paper just to give the truth conditions, without committing ourselves to a particular view on the compositional semantics that derives them. For discussion, see (Charlow and Sharvit 2014; Cresswell and von Stechow 1982; Percus and Sauerland 2003b).

to capture the semantics of de re belief reports, then included in the domain of any acquaintance-based concept should be the attitude holder's doxastic alternatives – we want a concept that gives us access to the individual that Ralph's *belief* state designates as the guy he saw in the brown hat, for example. So doxastic alternatives – the elements over which *believe* quantifies – must be world-individual pairs, and not merely worlds. We can then state the truth conditions of *Ralph believes that Ortcutt is a spy* as follows.

13. 'Ralph believes that Ortcutt is a spy' is true in  $w$  with respect to a contextually-supplied concept  $C$  iff  $C$  is a suitable concept of Ortcutt for Ralph and for every  $\langle w', y \rangle$ :  $\langle w', y \rangle \in \text{Dox}_{\text{Ralph}, w}$ ,  $C(w', y)$  is a spy in  $w'$ .

#### 14. *Suitability*

A concept  $C$  is a suitable concept of a res  $u$  for an attitude holder  $x$  in  $w$  iff:

- (i)  $C$  is *acquaintance-based* for  $x$  in  $w$ ; and
- (ii)  $C$  is a *reliable* concept of  $u$  for  $x$  in  $w$ .

#### 15. *Acquaintance-based*

A concept  $C$  is an *acquaintance-based* concept for  $x$  in  $w$  iff there is an acquaintance relation  $R$  such that for every world-individual pair  $\langle w', y \rangle$  that is in the domain of  $C$ ,  $R$  holds in  $w'$  between  $y$  and  $C(w', y)$ .

#### 16. *Reliability*

A concept  $C$  is a *reliable* concept of  $u$  for  $x$  in  $w$  iff  $C(w, x) = u$

Intuitively, the res in a de re report is just the individual that the reported attitude is about – Ortcutt in this case. A concept  $C$  witnesses (13) only if in addition to being acquaintance-based for Ralph, it is a *reliable* concept of the res (14i). This captures the fact that the truth of 'Ralph believes that Ortcutt is a spy' is dependent upon the fact that the man that Ralph saw in the brown hat is Ortcutt: the concept that witnesses the sentence returns Ortcutt when applied to the pair consisting of the actual world and Ralph. In a world in which that man is in fact Guy (and Ralph still believes that the man he saw in a brown hat is a spy), there is no *reliable* concept that witnesses (13), and the sentence turns out to be false.

So much for the truth conditions for one particular de re belief report. In order to show how de se can in general be treated as a special case of de re, we need to state the truth conditions for an arbitrary de re belief report  $\phi$ . Let us do that now:

#### 17. *Truth conditions for de re belief reports*

Let  $\phi$  be a report of a belief held by an attitude holder  $x$  that ascribes a property  $P$  to a res  $u$ .  $\phi$  is true in  $w$  with respect to a contextually-supplied concept  $C$  iff  $C$  is a concept of  $u$  that is suitable for  $x$  in  $w$  and for all  $\langle w', y \rangle$ :  $\langle w', y \rangle \in \text{Dox}_{x, w}$ ,  $C(w', y)$  is  $P$  in  $w'$ .

According to the view that we are considering, a de se belief report is just a de re belief report that can only be witnessed by a concept  $C$  if  $C$  is based on the acquaintance relation SELF, where SELF is defined as follows.<sup>7</sup>

<sup>7</sup> This is slightly different from the usual definition of SELF, in as much as I have added the condition that SELF( $x, y$ ) holds only if  $x$  is sentient. Typically, SELF is defined simply as identity, but notice that identity is

### 18. *SELF*

$SELF(x, y, w)$  iff (i)  $x$  is sentient in  $w$  and (ii)  $x = y$ .

We can then define a notion of SELF-based concept as follows.

### 19. *SELF-based concept*

$C$  is a SELF-based concept of  $x$  for  $x$  iff for every world-individual pair  $\langle w', y \rangle$  that is in the domain of  $C$ , SELF holds between  $y$  and  $C(w', y)$ .

Notice that (19), together with the definition of *SELF* entails that if  $C$  is a SELF-based concept of  $x$  for  $x$ , then for every  $\langle w', y \rangle$  in the domain of  $C$ ,  $C(w', y) = y$ . If  $C$  takes as an argument a centred world  $\langle w', y \rangle$  that is a doxastic alternative of the attitude holder, it will return the attitude holder's 'doxastic centre'  $y$ . To see how this derives the de se 'construal' of (2), let's assume that the sentence is compatible with only one LF, which is assigned the following interpretation:

20. 'Jonathan<sub>i</sub> believes that he<sub>i</sub> is a loathsome idiot' is true in  $w$  with respect to a contextually-supplied concept  $C$  iff  $C$  is a concept of Jonathan that is suitable for Jonathan in  $w$  and for all  $\langle w', y \rangle$ :  $\langle w', y \rangle \in \text{Dox}_{\text{Jonathan}, w}$   $C(w', y)$  is a loathsome idiot in  $w'$ .

Given the definition in (19), if the concept with respect to which the sentence is evaluated is SELF-based, then the sentence is true just in case (i) Jonathan is sentient and he is Jonathan and (ii) at each of Jonathan's doxastic alternatives  $\langle w', y \rangle$ ,  $y$  is a loathsome idiot in  $w'$ . Setting aside the trivially satisfied condition in (i), this is equivalent to the truth conditions assigned to the dedicated de se LF postulated in section 2.1.

## 2.3 Comparison of the two approaches

At least as far as belief reports go, it seems that one can conclude that the de se via binding and de se-as-de re routes derive equivalent truth conditions. On the version of the de se-as-de re route presented here, the SELF acquaintance relation that is responsible for de se interpretation is not imposed by the grammar. Rather, the interpretations that the grammar assigns to de re reports are underdetermined with respect to the value of the concept variable  $C$ , which is instead contextually supplied. If the context picks a SELF-based concept, then the result will be a sentence that is judged false in a situation where the reported attitude does not satisfy the awareness condition. This approach has the advantage that the only grammatical machinery that it depends on is independently needed for the analysis of de re belief reports; the SELF relation has generally been assumed to come for free: it is simply that acquaintance relation that all sentient individuals bear to themselves.

Until recently, an important argument that dedicated de se LFs are needed was the existence of anaphoric expressions that are unambiguously read de se (Chierchia 1990). It was reasoned that if de se-as-de re is the only route to de se construal, and if furthermore the

---

not actually an acquaintance relation: this pen in my hand is identical to itself but it is certainly not *acquainted* with itself. The idea underlying the definition in (18) is that SELF is just that acquaintance relation that all sentient individuals hold towards themselves simply in virtue of being sentient. I thank Frank Sode for discussion of this point.

SELF acquaintance relation is contextually supplied, then we should not expect the grammar itself to generate attitude reports that can only report de se attitudes. Obligatory controlled PRO is the classic example of an expression that is unambiguously read de se (Morgan 1970; Chierchia 1990): in languages such as German, French and Italian where *believe* is a control predicate, the counterpart of (21) can only be heard as false in the mirror situation.

21. John believed PRO to be a loathsome idiot.

Furthermore, logophoric pronouns – dedicated pronominal forms that can only occur in the scope of an attitude verb and obligatorily denote the attitude holder – were taken to be overt counterparts of obligatory controlled PRO, and thus to provide evidence that some languages have overt expressions that are unambiguously read de se (Heim 2001, 2002; Schlenker 1999; von Stechow 2002, 2003).

Recently, two significant challenges to this view have arisen. Firstly, fieldwork with native speakers has shown that the assumption that logophoric pronouns cannot be read de re is incorrect, at least for the West African language Ewe (Pearson 2013, 2015). This suggests at least that the array of unambiguously de se expressions in natural language may be significantly more restricted than was previously thought. Secondly, Emar Maier and Idan Landau have both developed lines of work where the grammar itself rather than the context imposes the restriction to SELF-based concepts (Maier 2011; Landau 2015a,b). On these approaches, the unavailability of a de re reading for PRO can be derived as a presuppositional constraint on the kind of concepts that can mediate the attitude holder's de re belief about herself. If they are correct that the resulting interpretation is truth conditionally equivalent to that arising through de se via binding, then the argument for the latter view from obligatory control is undermined.

Furthermore, Landau argues that a de se-as-de re treatment of PRO is to be preferred on the grounds that it permits a return to the traditional approach to control, whereby PRO is bound by the controller. He points out that de se via binding breaks the traditional dependency between the controller and PRO, which begs the question of how PRO comes to bear the same phi-features as its controller. According to Landau, there is no adequate way of implementing feature transmission from controller to PRO if PRO is bound by an abstractor in the left periphery of the infinitive itself rather than being bound by the controller directly.

In response to the challenge presented by Maier and Landau's views, I shall argue that a de se-as-de re approach to control cannot in principle handle the complete array of data that we find with PRO. In particular, the approach makes incorrect predictions regarding the interpretation of PRO in the types of cases introduced in (5) above – namely counterfactual attitude reports with counter-identity. De se via binding is needed for PRO and, by parity of reasoning, for counterfactual reports with counter-identity that contain an overt pronoun and receive the same interpretation as their counterpart with PRO.

### 3. Counterfactual reports with counter-identity

#### 3.1 Background

The notion that the interpretations yielded by *de se* via binding and *de se-as-de re* systematically come apart in counterfactual reports with counter-identity originates in work by Dilip Ninan (2008). Ninan noticed that (22), which is an example of such a report, gives rise to a puzzle.

22. I'm imagining that I am Brigitte Bardot and that **I** am kissing **me**.

[Ninan 2008: 25, ex 5a; based on Lakoff 1972]

The puzzle is this: normally, when a first person pronoun occurs in an attitude report where the attitude holder is first person and the attitude predicate is present tense, there is no choice but to construe the pronoun *de se*. Thus in our mirror scenario Jonathan would not be disposed to utter (23), because if he were to do so he would report that he holds a belief that is about himself and that he is aware is about himself (otherwise, why would he use *I* in the first place?).

23. I believe that **I** am a loathsome idiot.

We can see that this restriction is confined to the present tense by noticing that once Jonathan has realized that he is looking in a mirror, he might well say,

24. I believed that I was a loathsome idiot.

We should therefore expect that in (22), both the subject and the object of *kiss* are read *de se*. From that it would follow that the two pronouns are interpreted as picking out the same individual. But intuitively, the sentence reports an imagined kissing event between two individuals – the kisser is Bardot, and the kissee is the speaker. How can that be?

Ninan's solution is to assume that the subject of *kiss* is construed *de se* via binding, and the object is construed *de se-as-de re*. He develops an argument that the *de se* via binding route results in the pronoun picking out Brigitte Bardot, while *de se-as-de re* returns the speaker. I shall describe the evidence for this view in detail in section 3.3, but the basic idea is this: the acquaintance-based concepts that mediate *de re* construal in the scope of counterfactual attitudes such as *imagine* are anchored to the belief worlds of the attitude holder, not the worlds that the attitude holder counterfactually entertains. If I imagine that I am Brigitte Bardot, then my *beliefs* about who I am remain unchanged. It is this individual that is returned by the *de se-as-de re* route; *de se* via binding gives back Brigitte Bardot. The insight that the two routes to *de se* yield distinct interpretations in a case like (22) enables us to retain the generalization that *I* is always construed *de se* in first person present tense attitude reports, while also explaining why in Ninan's example the kisser and the kissee are non-identical.

To show how the argument works in detail, in the next two sub-sections I will lay out the

derivations for counterfactual reports with counter-identity, firstly with respect to the de se via binding route, and then for de se-as-de re.

### 3.2 De se via binding in counterfactual reports

We will need a semantics for *imagine*. For now, I will simply give this verb a parallel treatment as that discussed for *believe*: it is a quantifier over elements of a certain set of centred worlds, which in this case I shall refer to as ‘imagination alternatives’.<sup>8</sup> Roughly speaking, an attitude holder *x*’s imagination alternatives are those world-individual pairs  $\langle w', y \rangle$  such that those things that *x* imagines to be true are true in  $w'$ , and *y* is the inhabitant of  $w'$  that *x* imagines herself to be. Here is the lexical entry.

$$25. \llbracket \text{imagine} \rrbracket^{c, g} = \lambda P: P \in D_{\langle e, \langle s, t \rangle \rangle} \cdot \lambda x: x \in D_e \cdot \lambda w: w \in D_s \cdot \forall \langle w', y \rangle \in \text{Imagine}_{x, w}, P(y)(w')$$

Where  $\text{Imagine}_{x, w} = \{ \langle w', y \rangle : \text{it is compatible with what } x \text{ imagines in } w \text{ for } x \text{ to be } y \text{ in } w' \}$

Suppose the de se LF in (27) is assigned to the sentence in (26).<sup>9</sup>

26. Sophia is imagining (that she is Michelle Obama and) she is married to Barack Obama.

27. Sophia is imagining [<sub>CP</sub>  $\lambda x_1 \lambda w_2$  [<sub>w<sub>2</sub></sub> she<sub>1</sub> is married to Barack Obama]]

The semantics assigns the LF in (27) the following truth conditions.

$$28. \llbracket (27) \rrbracket^{c, g} = \lambda w. \forall \langle w', y \rangle \in \text{Imagine}_{\text{Sophia}, w}, y \text{ is married to Barack Obama in } w'.$$

(26) is thus predicted to be judged true just in case at each of Sophia’s imagination alternatives  $\langle w', y \rangle$ , *y* is married to Barack Obama in  $w'$ . Given the material in parentheses in (26), any *y* that is the individual coordinate of one of Sophia’s imagination alternatives must be Michelle Obama. This captures the intuition that on the most salient interpretation of (26), what Sophia is imagining is herself *as Michelle Obama* being married to Barack Obama, not Sophia being married to Barack Obama. In general, if a pronoun is de se bound in a counterfactual report with counter-identity, it will be construed as picking out the individual that the attitude holder imagines (wishes, pretends etc.) that she is.

### 3.3 De se-as-de re in counterfactual reports

Ninan argues that when a pronoun is construed de se-as-de re in a counterfactual report, it picks out the individual that the attitude holder believes herself to be, not the individual that she counterfactually supposes herself to be. It takes a little work to see why this should be. While studies of the semantics of de re reports have typically focused on de re belief, it seems to have been tacitly assumed that the analysis can carry over to other attitude

<sup>8</sup> This semantics will be revised in section 3.3.

<sup>9</sup> The material in parentheses in (26) merely establishes the context as one where Sophia puts herself in Michelle Obama’s shoes. I do not give the semantics for this portion of the sentence, but a complete representation would have it that the subject of the first conjunct is also de se bound.

predicates. Suppose we adopt this strategy, and attempt to give schematic truth conditions for de re imagination reports, based on those for de re belief reports given in section 2.2 above:

29. *Truth conditions for de re imagination reports* (first attempt)

Let  $\phi$  be a report of an imagining by an attitude holder  $x$  that ascribes a property  $P$  to a res  $u$ .  $\phi$  is true in  $w$  with respect to a contextually-supplied concept  $C$  iff  $C$  is a concept of  $u$  that is suitable for  $x$  in  $w$  and for all  $\langle w', y \rangle$ :  $\langle w', y \rangle \in \text{Imagine}_{x,w}$ ,  $C(w', y)$  is  $P$  in  $w'$ .

We then have the following truth conditions for (26):

30. 'Sophia is imagining that she is married to Barack Obama' is true in  $w$  with respect to a contextually-supplied concept  $C$  iff  $C$  is a concept of Sophia that is suitable for Sophia in  $w$  and for all  $\langle w', y \rangle$ :  $\langle w', y \rangle \in \text{Imagine}_{\text{Sophia},w}$ ,  $C(w', y)$  is married to Barack Obama in  $w'$ .

If we pick a SELF-based concept, then the sentence will turn out to be true just in case at each of Sophia's imagination alternatives  $\langle w', y \rangle$ , the individual that is identical to  $y$  is married to Barack Obama in  $w'$ . These are just the truth conditions that were assigned to the de se LF discussed above.

On first blush, it therefore appears that the de se-as-de re approach *can* derive the correct truth conditions for counterfactual reports with counter-identity, contrary to Ninan's claim that in this environment, the two routes to de se systematically yield different interpretations. If so, then Ninan's explanation for the interpretation of the Brigitte Bardot sentence repeated in (31) is no longer available to us.

31. I'm imagining that I am Brigitte Bardot and that I am kissing me.

Recall that Ninan proposed that the subject of *kiss* is construed de se via binding, whereas the object is read de re under the SELF acquaintance relation. If (29) is the correct schema for the truth conditions of de re imagination reports, then it is mysterious how Ninan's approach could yield the desired result that the subject and object of *kiss* can be different individuals.

Ninan points out, however, that there is good evidence that (29) is *not* the correct analysis of de re imagination reports. The evidence comes from the so-called puzzle of counterfactual de re. Consider the following sentence.

32. Ralph is imagining that Ortcutt isn't at the beach.

[Based on Ninan (2008, 2012)]

Ninan sets out the puzzle as follows. Suppose that there is only one way in which Ralph is acquainted with Ortcutt: he sees him at the beach. It seems that we can conceive of perfectly sensible situations that have this property in which the sentence is true. Indeed, this may be the exact moment at which Ralph sees Ortcutt at the beach; if at that very moment, Ralph imagines that the guy he is looking at is at home watching TV, then we can judge the sentence as true. Yet given (29), the truth conditions for (32) should be as follows.

33. ‘Ralph is imagining that Ortcutt isn’t at the beach’ is true in  $w$  with respect to a contextually-supplied concept  $C$  iff  $C$  is a concept of Ortcutt that is suitable for Ralph in  $w$  and for all  $\langle w', y \rangle$ :  $\langle w', y \rangle \in \text{Imagine}_{\text{Ralph}, w}$ ,  $C(w', y)$  isn’t at the beach in  $w'$ . Since the only acquaintance relation that Ralph bears to Ortcutt is the ‘sees at the beach’ relation, then if (32) is to stand a chance of being true it must be interpreted with respect to a concept  $C_{\text{BEACH}}$  that has the following properties:

34a. For every world-individual pair  $\langle w', y \rangle$  that is in the domain of  $C_{\text{BEACH}}$ ,  $y$  sees  $C(w', y)$  at the beach in  $w'$ .

(Since if  $C_{\text{BEACH}}$  is a suitable concept of Ortcutt for Ralph, then it must be acquaintance-based for Ralph.)

34b.  $C_{\text{BEACH}}(w, \text{Ralph}) = \text{Ortcutt}$

(Since if  $C_{\text{BEACH}}$  is a suitable concept of Ortcutt for Ralph, then it must be a reliable concept of Ortcutt.)

Furthermore, if (32) is true in the situation that we are considering, then the following must be true.

35. For all  $\langle w', y \rangle$ :  $\langle w', y \rangle \in \text{Imagine}_{\text{Ralph}, w}$ ,  $C_{\text{BEACH}}(w', y)$  isn’t at the beach in  $w'$ .

But (34a) and (35) jointly entail that Ralph is imagining that the person that he sees at the beach is not at the beach – that is, that he is imagining something impossible. Yet intuitively, when Ralph sees Ortcutt at the beach and at the same moment imagines that Ortcutt is not at the beach, what he imagines is not impossible.

The lesson is that the truth conditions for de re imagination reports cannot be stated by simply adopting the truth conditions for de re belief reports and replacing doxastic alternatives with imagination alternatives as I did in (29). Instead, it seems that the concepts that mediate the truth conditions of de re imagination reports need to somehow be anchored to the doxastic alternatives of the attitude holder, not his imagination alternatives. Intuitively, ‘Ralph is imagining that Ortcutt isn’t at the beach’ is true in our scenario because Ralph is imagining that the man he *believes* that he is looking at the beach isn’t at the beach. If we grant that an individual can have the property of being at the beach in Ralph’s belief worlds, but lack that property in his imagination worlds, then no contradiction arises.

There are various implementations of this idea on the market (Anand 2011; Ninan 2008, 2012; Yanovich 2011). For concreteness, I shall adopt one of the two solutions considered by Ninan (2008). Ninan’s starting point is the assumption that acts of imagining are always relative to belief states, in the sense that what is entertained is counterfactual relative to the doxastic alternatives of the attitude holder. This is modeled by treating the attitudinal alternatives quantified over by *imagine* as pairs of centred worlds  $\langle \langle w', y \rangle, \langle w'', z \rangle \rangle$ , where the first member of the pair,  $\langle w', y \rangle$ , is a doxastic alternative of the attitude holder, and the second member of the pair,  $\langle w'', z \rangle$  is compatible with what the attitude holder imagines relative to  $\langle w', y \rangle$ .<sup>10</sup> The truth conditions for de re imagination reports can then be stated as

<sup>10</sup> I will refrain from examining in detail the question of what it means to imagine something *relative to* some doxastic alternative. The basic idea is that the interpretation of certain attitudes, such as *imagine*, *wish* and *want* is

follows:

36. *Truth conditions for de re imagination reports* (revised version, based on Ninan 2008):

Let  $\phi$  be a report of an imagining by an attitude holder  $x$  that ascribes a property  $P$  to a res  $u$ .  $\phi$  is true in  $w$  with respect to a contextually-supplied concept  $C$  iff  $C$  is a concept of  $u$  that is suitable for  $x$  in  $w$  and for all  $\langle\langle w', y \rangle, \langle w'', z \rangle\rangle$ :  $\langle\langle w', y \rangle, \langle w'', z \rangle\rangle \in \text{Imagine}_{x,w}$ ,  $C(w', y)$  is  $P$  in  $w''$ .

Where  $\text{Imagine}_{x,w} = \{\langle\langle w', y \rangle, \langle w'', z \rangle\rangle$ : it is compatible with what  $x$  believes in  $w$  for  $x$  to be  $y$  in  $w'$ , and it is compatible with what  $x$  imagines in  $w$  relative to  $\langle w', y \rangle$  for  $x$  to be  $z$  in  $w''\}$

[Based on Ninan (2008): pages 44-45]

The correct truth conditions for (32) are given in (37).

37. ‘Ralph is imagining that Ortcutt isn’t at the beach’ is true in  $w$  with respect to a contextually-supplied concept  $C$  iff  $C$  is a concept of Ortcutt that is suitable for Ralph in  $w$  and for all  $\langle\langle w', y \rangle, \langle w'', z \rangle\rangle$ :  $\langle\langle w', y \rangle, \langle w'', z \rangle\rangle \in \text{Imagine}_{\text{Ralph},w}$ ,  $C(w', y)$  isn’t at the beach in  $w''$ .

According to the revised truth conditions, the sentence is true with respect to  $C_{\text{BEACH}}$  only if for all of Ralph’s imagination alternatives  $\langle\langle w', y \rangle, \langle w'', z \rangle\rangle$ ,  $C_{\text{BEACH}}(w', y)$  isn’t at the beach in  $w''$  – that is, if the individual that Ralph believes to be the the guy he sees at the beach is not at the beach at those worlds in which what Ralph counterfactually entertains (relative to what he believes) is true. We have now succeeded in stating truth conditions that match our intuitions about the meaning of the sentence.

With this in place, we can return to Ninan’s account of his Brigitte Bardot example, repeated below.

38. I’m imagining that I am Brigitte Bardot and that **I** am kissing **me**.

Ninan’s claim is that the subject of *kiss* is de se bound, and the object is construed de se-as-de re. To check that this correctly derives that the subject and object pick out different individuals, I will first give a revised semantics for the de se variant of *imagine*, which incorporates the notion that this verbs is quantifier over pairs of centred worlds.

---

parasitic on a doxastic modal base (Anand 2011; Heim 1992; Maier 2015; Ninan 2008; Yanovich 2011). This property has been argued to play a role, for example, in the presupposition projection properties of the verbs in question (Heim 1992; Maier 2015). In the case of *imagine*, it suffices to say that the content of the embedded clause is required to be counterfactual with respect to the doxastic alternatives of the attitude holder. This does not seem to be adequate, however, for a verb such as *say* or *claim* which is a counterfactual attitude verb according to the definition in footnote 2, but does not *demand* counterfactuality with respect to the attitude holder’s belief state. Since as far as I can see nothing I say here hinges on the precise details of how the de re concept comes to be applied to the doxastic alternatives of the attitude holder in a counterfactual report, I will set aside the question of how de re speech reports should be analyzed. I thank Frank Sode (p.c.) for observing that *say* and *claim* should be considered among the set of attitude verbs that can give rise to the puzzle of counterfactual de re.

39. *De se* variant of imagine (revised version):

$[[\text{imagine}]]^{c,g} = \lambda P: P \in D_{\langle e, \langle s, t \rangle \rangle} . \lambda x: x \in D_e . \lambda w: w \in D_s . \forall \langle \langle w', y \rangle, \langle w'', z \rangle \rangle \in \text{Imagine}_{x,w}, P(z)(w'')$

Where  $\text{Imagine}_{x,w} = \{ \langle \langle w', y \rangle, \langle w'', z \rangle \rangle : \text{it is compatible with what } x \text{ believes in } w \text{ for } x \text{ to be } y \text{ in } w', \text{ and it is compatible with what } x \text{ imagines in } w \text{ relative to } \langle w', y \rangle \text{ for } x \text{ to be } z \text{ in } w'' \}$

For the purposes of the interpretation of a *de se* bound pronoun, the first member of the ordered pair of centred worlds quantified over by *imagine* is an **idle wheel**; the property expressed by the embedded clause is applied to the individual and world coordinates of the second member of this pair:

40. Sophia is imagining [<sub>CP</sub>  $\lambda x_1 \lambda w_2$  [<sub>w<sub>2</sub></sub> she<sub>1</sub> is married to Barack Obama]]

41.  $[[ (40) ] ]^{c,g} = \lambda w. \forall \langle \langle w', y \rangle, \langle w'', z \rangle \rangle \in \text{Imagine}_{\text{Sophia},w}, z \text{ is married to Barack Obama in } w''$

As before, a *de se* bound pronoun is interpreted as the individual that the attitude holder imagines herself to be. By contrast, Ninan claims that that a pronoun that is construed *de se-as-de re* will be interpreted as picking out the individual that the attitude holder *believes* herself to be, given the semantics for *de re* imagination reports in (36). Since I am setting aside the question of how the truth conditions for *de re* reports are obtained compositionally, I refrain here from giving a complete LF and derivation for the Brigitte Bardot example. The point can be made simply by inspecting (42), which states the truth conditions that are assigned to an LF for (38) where the subject is *de se* bound and the object is free, in a context where the speaker is George.

42. 'I'm imagining (that I am Brigitte Bardot and) I am kissing me.' is true in  $w$  with respect to a contextually-supplied concept  $C$  iff  $C$  is a concept of George that is suitable for George in  $w$  and for all  $\langle \langle w', y \rangle, \langle w'', z \rangle \rangle: \langle \langle w', y \rangle, \langle w'', z \rangle \rangle \in \text{Imagine}_{\text{George},w}, z$  is kissing  $C(w', y)$  in  $w''$ .

If we pick that SELF-based concept  $C_{\text{SELF}}$  of George that is suitable for George, then the sentence turns out to be true just in case each of George's imagination alternatives  $\langle \langle w', y \rangle, \langle w'', z \rangle \rangle$  is such that  $z$  is kissing  $y$  in  $w''$ . This is just to say that George is imagining that the individual that he is imagining himself to be (Brigitte Bardot) is kissing the individual that he believes himself to be (in the ordinary course of things, George himself).

This demonstrates that Ninan's proposal that *de se* via binding and *de se-as-de re* come apart in counterfactual reports with counter-identity can explain the puzzling fact that in those environments, two occurrences of the first person pronoun may both be construed *de se* but nonetheless pick out distinct individuals. This on its own is not enough to demonstrate that *de se* binding is necessary, however. In order to do that, one would have to show that the relevant interpretation for the Brigitte Bardot sentence cannot be derived by letting both the subject and the object to be construed *de re*. This is the goal of the next section.

#### 4. The need for de se binding

In this section, I consider three possible strategies that the proponent of de se-as-de re who wishes to dispose of de se binding might employ to accommodate Ninan's observations counterfactual attitudes with counter-identity. I argue that none of them is successful.

##### 4.1 Revisiting the lesson from counterfactual de re

Recall that in order to explain why (43) can be true in the situation discussed in the last subsection without it following that Ralph is imagining something impossible, Ninan proposed a semantics that assigns the truth conditions in (44) to de re imagination reports.

43. Ralph is imagining that Ortcutt isn't at the beach.

44. Let  $\phi$  be a report of an imagining by an attitude holder  $x$  that ascribes a property  $P$  to a res  $u$ .  $\phi$  is true in  $w$  with respect to a contextually-supplied concept  $C$  iff  $C$  is a concept of  $u$  that is suitable for  $x$  in  $w$  and for all  $\langle\langle w', y \rangle, \langle w'', z \rangle\rangle$ :  $\langle\langle w', y \rangle, \langle w'', z \rangle\rangle \in$

$\text{Imagine}_{x,w}, C(w', y) \text{ is } P \text{ in } w''$ .

Where  $\text{Imagine}_{x,w} = \{\langle\langle w', y \rangle, \langle w'', z \rangle\rangle$ : it is compatible with what  $x$  believes in  $w$  for  $x$  to be  $y$  in  $w'$ , and it is compatible with what  $x$  imagines in  $w$  relative to  $\langle w', y \rangle$  for  $x$  to be  $z$  in  $w''\}$

But nothing in the data described so far requires that (44) must be the only available reading for (43). What if in addition to the reading on which the concept that mediates the reported de re imagination is anchored to the attitude holder's doxastic alternatives, there is a second reading where that concept is anchored to the centred worlds that are compatible with what he imagines? That is, what if in addition to (44), the grammar also generates the following truth conditions:

45. Let  $\phi$  be a report of an imagining by an attitude holder  $x$  that ascribes a property  $P$  to a res  $u$ .  $\phi$  is true in  $w$  with respect to a contextually-supplied concept  $C$  iff  $C$  is a concept of  $u$  that is suitable for  $x$  in  $w$  and for all  $\langle\langle w', y \rangle, \langle w'', z \rangle\rangle$ :  $\langle\langle w', y \rangle, \langle w'', z \rangle\rangle \in$

$\text{Imagine}_{x,w}, C(w'', z) \text{ is } P \text{ in } w''$ .

If this were the case, then it would not be necessary to appeal to de se binding in order to explain why the subject of *kiss* is construed as picking out Bardot in 'I'm imagining that I am Brigitte Bardot and I am kissing me', or why the subject of *married* is Michelle Obama in 'Sophia is imagining that she is Michelle Obama and she is married to Barack Obama'. It would be enough to say that the pronouns in these examples are construed de re with respect to a SELF-based concept, on the reading for de re imagination reports given in (45). On this construal, the relevant properties are ascribed to that individual that the attitude holder imagines herself to be identical to – Brigitte Bardot and Michelle Obama in the examples under consideration.

The task, then, is to show that there is no reading of de re imagination reports where the concept is anchored to the attitude holder's imagined worlds rather than her belief worlds. If there were, then there would be a reading of (43) on which it reports Ralph's imagining

something impossible: on this reading, (43) would mean roughly that for every world  $w'$  in which what Ralph imagines is true, the man he sees at the beach in  $w'$  is not at the beach in  $w'$ . It is clear that there is no such reading. Furthermore, the unavailability of this reading cannot be explained away by appealing to an independent constraint that prevents the grammar from generating readings that lead to contradiction. While other corners of the grammar provide evidence that contradictions can lead to ungrammaticality, for example in the domain of negative polarity item licensing (see Chierchia 2013 and references therein), the grammar does generate sentences that ascribe contradictory attitudes. Compare for example (43) to (46).

46. Ralph is imagining that the man that he is looking at the beach is not at the beach.

There is a reading of this sentence on which it is true in the situation we have been discussing, but it also has a reading on which it reports that Ralph imagines something impossible: namely, that the man picked out by the concept  $C_{\text{BEACH}}$  is not at the beach. I conclude from this that there is no reading of *de re* imagination reports where the concept is anchored to the centred worlds that are compatible with what the attitude holder imagines, rather than what she believes. Consequently, a pronoun that is construed *de se-as-de re* must be interpreted as picking out the individual that the attitude holder believes herself to be, which in the cases we are considering is distinct from the individual she imagines herself to be.

#### 4.2 More than one concept

It is not difficult to construct attitude reports where two co-referential nominals are interpreted as picking out distinct individuals, without *de se* being involved. Here is one:

47. Ralph believes that Orcutt is spying on himself.

(47) has a reading on which it is true in a situation where Ralph believes that the man he saw in the brown hat is spying on the man he saw at the beach, and does not realize that the two individuals are one and the same. In this case, the sentence is true in virtue of the availability of two distinct concepts for one and the same individual. Can we play the same game with the Brigitte Bardot example?

48. I'm imagining that I am Brigitte Bardot and that I am kissing me.

Because Ninan employed *de se* binding to account for the interpretation of the subject of *kiss*, he only needed one *de re* concept – namely the SELF-based one that the speaker has of herself. But in typical circumstances we bear any many acquaintance relations to ourselves, and correspondingly there are many concepts that could in principle mediate a *de re* attitude that is about the self. Can one of these be recruited for the interpretation of *I* in (48), thereby letting us dispense with *de se* binding?

The problem is that for (48), only those concepts will be picked that when applied to the attitude holder's doxastic alternatives return the same individual as is returned by a SELF-based concept. The reason is that any time I use the first person pronoun in a first person present tense attitude report, I report an attitude that I believe (*de se*) that I hold towards

myself. This just follows from my being competent from the meaning of the first person pronoun. For the subject of *kiss* to be construed as picking out Brigitte Bardot, we would need a concept that picks out that individual when applied to my doxastic alternatives. Unless I believe that I am Bardot (which I don't, if the context is one in which (48) can be uttered felicitously), this will not be the case. I conclude that we cannot dispense with de se binding in (48) by appealing to distinct concepts for the two occurrences of the first person pronoun.

### 4.3 Methods of identification across worlds

Consider again the semantics for de re imagination reports that we are assuming.

49. Let  $\phi$  be a report of an imagining by an attitude holder  $x$  that ascribes a property  $P$  to a res  $u$ .  $\phi$  is true in  $w$  with respect to a contextually-supplied concept  $C$  iff  $C$  is a concept of  $u$  that is suitable for  $x$  in  $w$  and for all  $\langle\langle w', y \rangle, \langle w'', z \rangle\rangle$ :  $\langle\langle w', y \rangle, \langle w'', z \rangle\rangle \in \text{Imagine}_{x,w}$ ,  $C(w', y)$  is  $P$  in  $w''$ .

Where  $\text{Imagine}_{x,w} = \{\langle\langle w', y \rangle, \langle w'', z \rangle\rangle$ : it is compatible with what  $x$  believes in  $w$  for  $x$  to be  $y$  in  $w'$ , and it is compatible with what  $x$  imagines in  $w$  relative to  $\langle w', y \rangle$  for  $x$  to be  $z$  in  $w''\}$

An important property of this semantics that I have been ignoring up until now is that it entails that in the worlds  $w'$  in which what the attitude holder imagines is true, some property  $P$  holds of an individual that is obtained by applying a concept  $C$  to a centred world whose world coordinate is distinct from  $w'$ . We therefore need a method of identifying individuals across worlds. Yanovich (2011) proposed a solution to the puzzle of counterfactual de re that employed such a method of identification, in the form of a function  $g_{id}$ . Setting aside the technical details of Yanovich's full proposal, the crucial property of  $g_{id}$  for our purposes is that it maps an inhabitant of the attitude holder's doxastic alternatives (that is, an individual obtained by applying the de re concept to the doxastic alternative) to a correspondent of that individual in the relevant counterfactual worlds. (It is of type  $\langle e, \langle\langle s, e \rangle, e \rangle\rangle$ .) In principle, in a counterfactual report with counter-identity the function might, when applied to the individual that the attitude holder believes herself to be, return that individual that she counterfactually supposes that she is.<sup>11</sup> If  $g_{id}$  does indeed have this property, then it will yield a construal that is equivalent to that obtained by de se binding. To see this, take our Michelle Obama sentence as an example.<sup>12</sup>

50. Sophia is imagining (that she is Michelle Obama and) she is married to Barack Obama.

Incorporating  $g_{id}$  into the semantics yields the following truth conditions:

51. 'Sophia is imagining (that she is Michelle Obama and) she is married to Barack Obama' is true in  $w$  with respect to a contextually-supplied concept  $C$  and identification function  $g_{id}$  iff  $C$  is a concept of  $u$  that is suitable for Sophia in  $w$  and for all  $\langle\langle w', y \rangle, \langle w'', z \rangle\rangle$ :  $\langle\langle w',$

<sup>11</sup> I thank Igor Yanovich for discussion of this point.

<sup>12</sup> The same point could have been made with Ninan's Brigitte Bardot example, but since the distinct interpretation of subject and object in that sentence are irrelevant for this portion of the discussion, I set it aside in favor of a simpler case with only one pronoun.

$y\rangle, \langle w'', z\rangle\rangle \in \text{Imagine}_{x,w}, g_{id}(\mathbf{C}(w', y))(w'', z)$  is married to Barack Obama in  $w''$ .

Since I am continuing to consider strategies for defending the view that de se-as-de re can do anything that de se binding can do, I will focus on the interpretation that results from picking a SELF-based concept as the value of C. Then assuming that Sophia believes (de se) that she is Sophia, this is the individual that is supplied as the first argument of  $g_{id}$ . If in addition  $g_{id}$  has the latitude to map the individual that an attitude holder believes that she is to the individual that she imagines that she is, then plugging in Sophia's imagination worlds returns Michelle Obama. The result is indeed equivalent to that obtained by de se binding.

So far I have said very little about what a method of identification across worlds should look like. There are no doubt thorny issues lurking here; clearly, the set of pairs of individuals that can end up being identified with one another should be highly constrained. The question before us is whether there are identification functions  $g_{id}$  that have those properties that the account of (50) just described depends on. That is, we need to know whether there can be an identification function that has the properties stated in (52).

52.  $g_{id}(x)(w'', z) = u$

Where (i)  $\langle w'', z\rangle$  is the second member of the arbitrarily chosen ordered pair  $\langle \langle w', y\rangle, \langle w'', z\rangle\rangle$ , which is an element of the imagination alternatives of the attitude holder; (ii)  $x$  is the individual coordinate of some  $\langle w', x\rangle$ , which is an element of the doxastic alternatives of the attitude holder; and (iii)  $u$  is the individual coordinate of some  $\langle w'', u\rangle$ , which is a member of the attitude holder's imagination alternatives.

This is at least in part an empirical question. I will now set out evidence suggesting that there is no such identification function.

My starting point is this. In a report of Sophia imagining that she is Michelle Obama, an embedded pronoun that is anaphoric to the attitude holder can pick out either Michelle Obama or Sophia. So far, I have been focusing on the construal where it picks out Michelle Obama. But consider the following case.

53. *Scenario*: Sophia and Michelle Obama are always arguing about healthcare reform; Sophia holds a more rightwing view than Michelle Obama. Sophia wonders what it must be like for Michelle Obama to listen to her opinions on this topic.

Sophia is imagining that she is Michelle Obama and **she** is saying rightwing things about healthcare reform.

There is a reading of the sentence on which it is true in this scenario; it is the reading on which the bolded pronoun is construed as picking out Sophia, not Michelle Obama. If one were to adopt the strategy I am considering in this section for dispensing with de se binding, then one would have to say that identification functions come in at least two flavors. One type, call it  $g_{id1}$ , has the properties illustrated in (52): it maps an individual that the attitude holder believes herself to be to an individual that she imagines herself to be. A second type,

$g_{id2}$ , maps an individual that the attitude holder to believes herself to be to the *very same* individual in her imagination alternatives.

The data that we have seen so far suggest that either type of identification function can be recruited to calculate the truth conditions of de re reports with overt pronouns. But matters look quite different when one replaces the overt pronoun with controlled PRO:

54. Sophia is imagining PRO being Michelle Obama and PRO saying rightwing things about healthcare reform.

This sentence is unambiguously false in the scenario described above. What it reports is Sophia imagining, as Michelle Obama, saying rightwing things about healthcare reform.

One might suspect that the crucial difference between the control sentence and its counterpart with an overt pronoun is that on its control variant, *imagine* forces the sentence to be read as a report of an imagining ‘from the inside’ (Higginbotham 2003). But the pattern holds up systematically across the range of control predicates that can denote an attitude whose content is counterfactual with respect to the beliefs of the attitude holder. In the following sets of examples, I set up a pair of scenarios. In the first one, the counterfactual report is true if the embedded subject is construed as picking out the individual that the attitude holder believes herself to be; in the second, it is true if the embedded subject picks out whoever the attitude holder counterfactual supposes that she is. In each case, the sentence with the overt pronoun or reflexive is judged true in both scenarios, whereas the sentence with PRO is only judged true in the second scenario.

55. *Scenario 1*: Sophia wants to enhance her own reputation among Republicans. She disguises herself as Michelle Obama and tells Republicans that Sophia has been saying rightwing things about healthcare reform.

*Scenario 2*: Sophia wants to damage Michelle Obama’s reputation among Democrats. She disguises herself as Michelle Obama and ‘confesses’ that she has in the past said rightwing things about healthcare reform.

55a. Sophia is pretending that she is Michelle Obama and she said rightwing things about healthcare reform.

Scenario 1: True      Scenario 2: True

55b. Sophia is pretending PRO to be Michelle Obama and PRO to have said rightwing things about healthcare reform.

Scenario 1: False      Scenario 2: True

56. *Scenario 1*: Sophia wants to enhance her own reputation among Republicans. She disguises herself as Michelle Obama and says to Republicans, ‘I am Michelle Obama and Sophia has been saying rightwing things about healthcare reform’.

*Scenario 2:* Sophia wants to damage Michelle Obama’s reputation among Democrats. She disguises herself as Michelle Obama and ‘confesses’, ‘I am Michelle Obama and in the past I have said rightwing things about healthcare reform’.

56a. Sophia claimed that she was Michelle Obama and that she had said rightwing things about healthcare reform.

Scenario 1: True      Scenario 2: True

56b. Sophia claimed PRO to be Michelle Obama and PRO to have said rightwing things about healthcare reform.

Scenario 1: False      Scenario 2: True

Another way to make the point is to set up environments where the only plausible interpretation entails that the embedded subject picks out the belief self rather than the counterfactual self. Control sentences are degraded in these environments.

57. *Scenario:* Sophia thinks it would be great to be Michelle Obama, because of the status and political influence that it would bring. However, she thinks that if she were Michelle Obama, she would need an excellent political advisor. In fact, she thinks that that she herself would be the person best qualified for the job.

57a. Sophia wants to be Michelle Obama and for herself to be her political advisor. *True*

57b. #Sophia wants to be Michelle Obama and PRO to be her political advisor.

58. *Scenario:* Last night, Sophia had a dream in which she was Michelle Obama. In the dream, Sophia herself was her political advisor.

58a. Sophia dreamed that she was Michelle Obama and she was her political advisor. *True*

58b. #Sophia dreamed about PRO being Michelle Obama and PRO being her political advisor.

In the (b) examples above I use ‘#’ to indicate that the sentence is degraded on the reading where a single desire or dream is reported. Strikingly, to the extent that the sentences are acceptable, it is on a reading where being Michelle Obama and being Michelle Obama’s political advisor fulfill distinct desires, or events that take place in separate dreams. The reason is that in these sentences, PRO must be interpreted as picking out whoever Sophia wants to be or dreams that she is. If the content of the second conjunct is taken to describe the same desire or dream as the first conjunct does, then this individual must be Michelle Obama. The property of being Michelle Obama’s political advisor is then ascribed to Michelle Obama herself - a state of affairs that is at least implausible. By contrast, replacing PRO with overt pronouns yields felicitous sentences that are judged true in the relevant scenarios. This is because unlike PRO, overt pronouns can be read as picking out the attitude holder’s belief self in counterfactual reports with counter-identity.

What does this mean for our two flavors of identification functions  $g_{id1}$  and  $g_{id2}$ ? If we are to dispense with de se binding and instead use identification functions to obtain the mapping from the attitude holder’s belief self to her counterfactual self, then we shall have to stipulate

that overt pronouns can be construed with respect to either  $g_{id1}$  or  $g_{id2}$ , but PRO only has access to  $g_{id1}$ . If PRO could be interpreted with respect to  $g_{id2}$ , then it would be possible to interpret it as picking out the belief self in counterfactual reports with counter-identity, contrary to fact.

I expect that it is possible to formulate a constraint on a de se-as-de re analysis of PRO requiring that the mediating identification function must be of the  $g_{id1}$  type rather than the  $g_{id2}$  type. But it seems unlikely that this is the right way to go. It would mean that the semantics of control must have built into it a restriction that affects the interpretation only in a rarified set of contexts, and only for certain predicates. This restriction would not seem to follow from any other property of control, nor would its effects be felt elsewhere in the grammar.

For example, that PRO is unambiguously de se does not in itself force the restriction: as we have seen, *I* is also unambiguously de se in first person present tense attitude reports; if it is construed de re, then the mediating concept must be SELF-based. Yet it displays the same ambiguity that we see with other overt pronouns, but is excluded with PRO. Ninan's Brigitte Bardot example is a case in point; we can also construct a closer parallel with the control sentences that we have been examining:

59. *Scenario*: As in (53) above. Sophia describes what she is imagining as follows: 'I am imagining that I am Michelle Obama and I am saying rightwing things about healthcare report.'

If PRO is de se bound, then we can say what the crucial difference is between it and the first person pronoun: the former is obligatorily bound, which entails that it picks out the individual that the attitude holder imagines, pretends, wishes etc. that she is. *I*, like other overt pronouns, can but need not be bound. If it is free, then it is construed de re; if the mediating concept is SELF-based, then it is interpreted as picking out the belief self of the attitude holder.

A corollary of the approach that I am advocating is that an expression that denotes the attitude holder but cannot be de se bound cannot be construed as picking out the attitude holder's counterfactual self. Proper names look like a likely candidate with which to test this prediction. If we replace PRO with a proper name in our examples, the situations in which the resulting sentences are true should be the complement set of the situations in which the PRO sentences are true. This is exactly what we find.<sup>13</sup>

---

<sup>13</sup> The sentences are perhaps a little degraded owing to the repeated name penalty induced by use of the proper name in an environment where it would have been natural to use a pronoun. Nonetheless, it seems clear that there is a crisp boundary between the situations in which they are judged true and those in which they are judged false. The acceptability of the sentences is perhaps somewhat improved by adding a pronoun or reflexive, as in the following examples.

(i) Sophia is imagining that she is Michelle Obama and she, Sophia, is saying rightwing things about healthcare reform.

(ii) Sophia is imagining that she is Michelle Obama and Sophia herself is saying rightwing things about healthcare reform.

60. Scenario: Sophia and Michelle Obama are always arguing about healthcare reform; Sophia holds a more rightwing view than Michelle Obama. Sophia wonders what it must be like for Michelle Obama to listen to her opinions on this topic.

Sophia is imagining that she is Michelle Obama and Sophia is saying rightwing things about healthcare reform. True

61. *Scenario 1:* Sophia wants to enhance her own reputation among Republicans. She disguises herself as Michelle Obama and tells Republicans that Sophia has been saying rightwing things about healthcare reform.

*Scenario 2:* Sophia wants to damage Michelle Obama's reputation among Democrats. She disguises herself as Michelle Obama and 'confesses' that she has in the past said rightwing things about healthcare reform.

Sophia is pretending that she is Michelle Obama and Sophia said rightwing things about healthcare reform.

Scenario 1: True      Scenario 2: False

62. *Scenario 1:* Sophia wants to enhance her own reputation among Republicans. She disguises herself as Michelle Obama and says to Republicans, 'I am Michelle Obama and Sophia has been saying rightwing things about healthcare reform'.

*Scenario 2:* Sophia wants to damage Michelle Obama's reputation among Democrats. She disguises herself as Michelle Obama and 'confesses', 'I am Michelle Obama and in the past I have said rightwing things about healthcare reform'.

Sophia claimed that she was Michelle Obama and Sophia had said rightwing things about healthcare reform.

Scenario 1: True      Scenario 2: False

63. *Scenario:* Sophia thinks it would be great to be Michelle Obama, because of the status and political influence that it would bring. However, she thinks that if she were Michelle Obama, she would need an excellent political advisor. In fact, she thinks that that she herself would be the person best qualified for the job.

Sophia wants to be Michelle Obama and for Sophia to be her political advisor. True

64. *Scenario:* Last night, Sophia had a dream in which she was Michelle Obama. In the dream, Sophia herself was her political advisor.

Sophia dreamed that she was Michelle Obama and Sophia was her political advisor. True

One might be tempted to reply to these data by saying that names work differently: of course *Sophia* cannot refer to Michelle Obama. But in the scope of an attitude predicate it can – or rather it can invoke a concept that in the relevant attitudinal alternatives can be mapped to an individual that is identical to Michelle Obama. Suppose that Bill reads a speech that was written by Sophia but which he thinks was written by Michelle Obama.

Then the following sentence is true in virtue of the facts that (i) Bill believes that writer of the speech he read is Michelle Obama and (ii) the writer of the speech he read is in fact Michelle Obama.

65. Bill believes that Sophia is Michelle Obama.

Why, then, should it not be possible to interpret (66) as reporting that Sophia is imagining herself, in the guise of Michelle Obama, saying rightwing things about healthcare reform?

66. Sophia is imagining that she is Michelle Obama and Sophia is saying rightwing things about healthcare reform.

In fact, it should be, if there is a mechanism of mapping the belief-self (derived via *de se-as-de re*) to the counterfactual self by invoking the  $g_{id1}$  type of identification function. So defenders of this view would not only have to state what the difference is between pronouns and PRO such that PRO only tolerates the  $g_{id1}$  type of identification function, but they must also explain why proper names apparently only tolerate the  $g_{id2}$  type. These matters look murky for this type of view, but can be explained elegantly once one countenances *de se* binding: in a counterfactual report with counter-identity, the only route to having the pronoun pick out the counterfactual self of the belief holder is by letting it be bound from the left periphery of the embedded clause. Obligatorily controlled PRO is obligatorily bound by this operator, which explains not only why it is obligatorily *de se*, but also why it cannot pick out the belief self in this environment. Overt pronouns can either be bound or free, and hence can be mapped either to the counterfactual self or to the belief self of the attitude holder. Proper names cannot be bound by a lambda operator, and therefore can only pick out the belief self of the attitude holder. I therefore conclude that there are dedicated *de se* LFs.

## 5. Consequences of the proposal

In this section I describe two interesting consequences that follow from the view proposed in this paper.

### 5.1 De re blocking effects

The way of thinking about *de se* construal in counterfactual reports developed here sheds new light on so-called ‘*de re* blocking effects’. Percus and Sauerland (2003a) noticed that in dream reports with counter-identity, a *de se* pronoun cannot be c-commanded by a *de re* pronoun:

67. Sophia dreamed that she was Michelle Obama and she disliked her husband.

According to Percus and Sauerland, this sentence can report a dream in which Sophia’s dream self – Michelle Obama – dislikes Sophia’s husband. On this reading, *she* is construed *de se* and *her* *de re*. However, it cannot report a dream where Sophia dislikes Barack Obama.

The judgments are subtle, but have received initial experimental support from Pearson and Dery (2014).<sup>14</sup>

Several accounts of de re blocking effects have been offered in the literature (Anand 2006; Charlow 2010; Percus and Sauerland 2003a). All of them share the assumption that de re blocking is a form of syntactic intervention, which is expected if de se construal comes about via binding, but is unexplained otherwise. For example, Percus and Sauerland propose that a de se pronoun is not bound in situ, but rather behaves like a relative pronoun in that it moves (covertly) to the left edge of the embedded clause, triggering insertion of an abstractor that binds the trace. The resulting structure is assigned a property interpretation; in this sense, the proposal is a variant of the binding in situ approach.<sup>15</sup>

68a. Sophia dreamed [<sub>CP</sub> (she)  $\lambda x_1 \lambda w_2$  [<sub>w\_2</sub>  $t_1$  disliked her<sub>3</sub> husband]].

68b.  $[[CP]]^{c,g} = \lambda x \lambda w. x$  dislikes  $g(3)$ 's husband in  $w$

A principle of Superiority dictates that given two pronouns that can both undergo this movement operation, it is the structurally higher one should move. The reason why the sentence cannot be understood as reporting a dream where Sophia dislikes Barack Obama is that this would require movement of the possessive pronoun across the de re subject, which would violate Superiority.

To the extent that this proposal is successful it too provides an argument for de se binding; if there are no syntactic dependencies that can produce de se readings then it is difficult to explain why the array of de se and de re construals is apparently structurally conditioned. However, a sticking point for this approach is that de re blocking effects apparently only show up with certain predicates. Anand (2006) noticed that in the scope of *believe*, a de re pronoun *can* c-command a de se pronoun.<sup>16</sup>

69. Sophia believed that she was Michelle Obama and she disliked her husband.

Suppose that Sophia is suffering from amnesia, and confused about her identity. Various circumstances lead her to form the belief that she is Michelle Obama, the wife of the President. She comes across a newspaper article that she herself wrote, which is very critical of the President. She concludes, 'Wow, whoever wrote this must really dislike my husband'. According to Anand, a sentence like (69) is true in this type of situation. He concludes that de re blocking effects are found with *dream* but not with *believe*.

Since we should not expect the syntactic constraint which gives rise to de re blocking to be active in the scope of some verbs but not others, Anand takes his data to show that there is a second route to de se, in addition to de se binding. This route – de se as a special case of de

---

<sup>14</sup> The empirical picture is further complicated by a follow-up study to Pearson and Dery (2014) (Dery and Pearson to appear). This paper shows that participants' judgments in this domain are sensitive to both effects of exposure and task effects, which need to be better understood if the experimental data are to be used as a basis for robust conclusions about linguistically naïve speakers' judgments.

<sup>15</sup> I gloss over the details of the de re interpretation of the possessive pronoun, to keep the representations simple.

<sup>16</sup> Again, see Pearson and Dery (2014) for initial experimental support for these judgments.

re – is available for the possessive pronoun in the scope of *believe* but not for *dream*. Consequently, a de re pronoun can c-command a corresponding de se pronoun in a belief report without violation of syntactic constraints.

In this paper I align my support with Anand’s conclusion – both routes to de se are needed. In addition, we are now in a position to understand why the de se-as-de re route is apparently available for *believe* but not for *dream*. Anand answers this question by giving a semantics for *dream* that constrains the concepts that may mediate the interpretation of de re material in its scope: such concepts may not be SELF-based. This solution is obviously unsatisfying: no principled explanation is given of why de re construal should be constrained in this way in the scope of *dream* but not of *believe*.<sup>17</sup> Furthermore, we can construct cases where a pronoun in the scope of *dream* can be interpreted with respect to the SELF acquaintance relation. To do so, we need a dream report with counter-identity where the only acquaintance relation that the attitude holder bears to herself is SELF. Here is an example:

70. *Scenario*: Mary wakes up from a coma, and remembers nothing; the only acquaintance relation she bears to herself is SELF. She does not know, for instance, what she looks like, but when she drifts off to sleep she has a dream in which she sees herself from someone else’s perspective.

Mary dreamed that she was someone else and she was looking at herself. True

In this example, Mary’s dream self is looking at Mary’s belief self; the reflexive is construed de se-as-de re. So building a ban on SELF-based concepts into the semantics of *dream* is not the right move.

In fact, the view proposed in this paper has the consequence that nothing special needs to be said about the semantics of *dream* and *believe* in order to explain why the former shows de re blocking effects but the latter does not. The crucial difference between the two verbs is that whereas in the scope of doxastic predicates with counter-identity, the two routes to de se return the same individual, in the scope of *dream* they do not. As has now been argued at length, de se via binding returns the counterfactual self, whereas de se-as-de re returns the belief self. Consequently, the de se-as-de re route lets the possessive pick out Michelle Obama in (69), but not in (67).

Notice that this solution depends on the assumption that *dream* is like *imagine*, *pretend*, *want* and *claim* in denoting an attitude that an attitude holder can coherently hold towards a content that she believes to be false. This is at odds with the treatment of *dream* given in Percus and Sauerland (2003a), for example, where it is analyzed as meaning roughly ‘believe while asleep’. Instead, I propose that the verb should more properly be treated as ‘imagine while asleep’. In fact, there is a philosophical literature arguing in favor of this latter view of the nature of dreams (Ichikawa 2009). As linguistic evidence, consider the sentence in (71).

---

<sup>17</sup> Percus (2006) suggests that the difference between *dream* and *believe* may be that the former but not the latter involves ‘putting oneself in someone else’s shoes’ – precisely the type of construal that I am concerned with in this paper. The view developed in this paper provides an articulation of why this should matter for de re blocking.

71. Ralph is dreaming that Ortcutt never existed.

We can re-construct the puzzle of counterfactual de re with this case: (71) can report a de re dream that Ralph is having about Ortcutt (qua the man he saw at the beach, say) that does not involve an impossibility (that the man Ralph saw never existed). So we can run the arguments that we have developed for *imagine* on *dream*: in a report of a dream that Sophia has in which she is Michelle Obama, de se via binding gives back Michelle Obama, and de se-as-de re gives back Sophia. This provides a more principled explanation of the landscape of de re blocking effects than was available before now.

## 5.2 Counter-identity without de se?

In this sub-section, I consider the possibility that the counter-identity cases I have discussed are not specific to de se reports (Dever 2014).

I have argued that in a counterfactual report with counter-identity, the only way for a nominal expression that takes the attitude holder as its antecedent to be construed as picking out the counterfactual self is if it is de se bound. As is well known, a nice property of de se binding is that in counter-identity reports such as those in (72), it circumvents the problem of identity between distinct individuals that would arise on a propositional view: it does not matter that there is no world in which Sophia is Michelle Obama, or where Heimson is Hume, since the embedded pronouns are not interpreted as denoting those individuals, but rather are merely lambda abstracted variables.

72a. Sophia is imagining that she is Michelle Obama.

72b. Heimson believes that he is Hume.

But de se binding is not the only way of circumventing this problem, as we saw in example (65), repeated below:

73. Bill believes that Sophia is Michelle Obama.

Here, the trick is that *Sophia* and *Michelle Obama* are both construed de re; what is required of Bill's doxastic alternatives is not that Sophia and Michelle Obama be identical, but rather that the individuals that the relevant concepts return are. We should ask, then, what happens in cases where *believe* is replaced with a counterfactual attitude:

74a. Bill is imagining that Sophia is Michelle Obama.

74b. Bill wishes that Sophia were Michelle Obama.

Notice firstly that it is difficult to know what to make of these sentences on encountering them out of the blue. But with a bit of context, they can be heard as acceptable. Suppose, for example, that Bill used to work for Michelle Obama, and he now works for Sophia. He preferred Michelle Obama as a boss, so he imagines that he is working for her again, or wishes that he were working for her again. In these scenarios, I think that the sentences in (74) can be judged true.

This is initially surprising given what I have said about counterfactual de re: if Bill is not

mistaken about who Sophia and Michelle Obama are, then the sentences should require that in the relevant counterfactual worlds Sophia is Michelle Obama. Is this a problem for our overall picture?

I think that these cases need to be understood better, but that they do not undermine the view developed in this paper. Firstly, it is telling that they require a significant amount of contextual support: out of the blue, I do not know what it means to imagine that Sophia (or Hilary Clinton, or anyone else) is Michelle Obama, but I do know what it means to imagine that I am Michelle Obama. In the former case, there must be some contextually salient role that Sophia occupies (such as being Bill's boss); in the latter, this is not needed. In fact, this 'role-occupying' type of identity predication can arise without an attitude in sight. If I temporarily fill in for Sophia at her job (because she is sick, say), I might say to the staff:

75. For today, I am Sophia.

Thus one way of dealing with the cases in (74) that leaves our view in tact is to invoke something like the following lexical entry for the copula:

76.  $[[be]]^{c,g} = \lambda x \lambda y \lambda w. \text{ in } w, x \text{ occupies the role canonically occupied by } y$

The view that counterfactual de se with counter-identity is a distinct phenomenon from counter-identity about other individuals is supported by failure of inference from the first type of attitudinal content to the second type. Consider the following example:

77. Sophia's sister has the same desires as Sophia has.  
Sophia wishes that she were Michelle Obama.  
#Therefore, Sophia's sister wishes that Sophia were Michelle Obama.

This is quite different from the inference patterns that we find in counterfactual reports without counter-identity: in principle, there is nothing that blocks an inference from x desiring (de se) property P to y desiring that P hold of x:

78. Sophia's sister has the same desires as Sophia has.  
Sophia wishes that she were rich.  
Therefore, Sophia's sister wishes that Sophia were rich.

In sum, the data in this section support the view that there is something distinctive about counterfactual de se attitudes with counter-identity: they cannot be collapsed together with reports of imagined or desired counter-identity about third parties. This is to be expected if the former class of case is derived by a binding mechanism that is unavailable for the latter class.

## 6. Conclusion

The debate between proponents of de se binding and of de se as a special case of de re is premised on a mistaken assumption: namely, that the two routes yield equivalent truth conditions in all cases. I have argued that this assumption is incorrect, building on Ninan's

insight that the two routes come apart in counterfactual attitudes with counter-identity. The strong version of this view that I have proposed is that the two routes not only can yield distinct interpretations in this class of case, but in fact they must. There is no way for a pronoun that is construed de se-as-de re to be interpreted as picking out the attitude holder's counterfactual self, for principled reasons having to do with the workings of de re in counterfactual attitude reports. Where we find a pronoun that can be construed in this way, we should conclude that it can be de se bound. Where we find one that can *only* be construed in this way, we should conclude that it must be. Overt pronouns fall into the former category, and PRO falls into the latter one.

This way of reasoning about de se lets us be more precise about the facts that need to be derived by a semantics for obligatory control – a case that lies at the heart of the project of analyzing de se reports. The core semantic fact about obligatorily controlled PRO is not simply that it is unambiguously read de se, but that it unambiguously encodes a particular flavor of de se. In this respect, theories of control that appeal to de se binding rather than de se-as-de re enjoy a clear advantage.

[Acknowledgments to be added]

## References

- Anand, P. (2006). *De De Se*. Unpublished PhD dissertation, MIT.
- Anand, Pranav. (2011). Suppositional projects and subjectivity. Ms., UCSC.
- Boër, S.E. and Lycan, W.G. (1980) Who, me? *The Philosophical Review*, 89(3). 427-466
- Cappelen, H. & J. Dever (2013). *The Inessential Indexical: on the Philosophical Insignificance of Perspective and the First Person*. Oxford University Press.
- Charlow, S. 2010. Two kinds of de re blocking. Handout of talk given at MIT Ling Lunch.
- Charlow, S., & Sharvit, Y. (2014). Rethinking the LFs of attitude reports. *Semantics and Pragmatics*, 7(3). 1-43.
- Chierchia, G. (1990). Anaphora and attitudes de se. In R. Bartsch, J. van Benthem & van Emde Boas, (Eds.), *Semantics and Contextual Expression* (pp. 1-32). Dordrecht: Foris.
- Chierchia, G. (2013). *Logic in Grammar: Polarity, Free choice and Intervention*. Oxford: Oxford University Press.
- Cresswell, M.J. and von Stechow, A. (1982). 'De re' belief generalized. *Linguistics and Philosophy*, 5(4). 503-535.
- Dery, J. & Pearson, H. (To appear). On experience-driven semantic judgments: a case study on the Oneiric Reference Constraint. *Proceedings of the 6<sup>th</sup> Conference on Quantitative Investigations in Theoretical Linguistics*.
- Dever, J. (2014). Reply to Richards' and Pearson's comments on *The Inessential Indexical: on the Philosophical Insignificance of Perspective and the First Person*.
- Heim, Irene. (1992). Presupposition projection and the semantics of attitude verbs. *Journal of Semantics* 9(3). 183–221.
- Heim, I. (2001). Semantics and morphology of person and logophoricity. Handout.
- Heim, I. (2002). Features of pronouns in semantics and morphology. Handout.
- Higginbotham, James. 2003. Remembering, imagining, and the first person. In Alex Barber (ed.), *Epistemology of language*, 496–533. Oxford: Oxford University Press.
- Hintikka, J. (1969). Semantics for propositional attitudes. In J. W. Davis, & D. J. Hockney (Eds.), *Philosophical Logic* (pp. 21-45). Dordrecht: Reidel.

- Ichikawa, J. (2009). Dreaming and imagination. *Mind and Language*, 24(1), 103-121.
- Kaplan, D. (1968). Quantifying in. *Synthese*, 19(1-2), 178-214.
- Lakoff, G. (1972). Linguistics and natural logic. In D. Davidson & G. Harman, (eds.), *Semantics of Natural Language*. D. Reidel Publishing Company, Dordrecht.
- Landau, I. (2015a). *A two-tiered theory of control*. MIT Press: Cambridge MA.
- Landau, I. (2015b). Direct variable binding and agreement in obligatory control. Unpublished manuscript, Ben Gurion University.
- Lewis, D. (1979). Attitudes de dicto and de se. *The Philosophical Review*, 88(4), 513-543.
- Magidor, O. (2014). The myth of the de se. To appear in *Philosophical Perspectives*.
- Maier, Emar. 2006. *Belief in context: Towards a unified semantics of de re and de se attitude reports*: Radboud Universiteit Nijmegen dissertation.
- Maier, E. (2010). Presupposing acquaintance: a unified semantics for de dicto, de re and de se belief reports. *Linguistics and Philosophy* 32, 429-474.
- Maier, E. (2011). On the roads to de se. In *Proceedings of Semantics and Linguistic Theory 21*. 393-412.
- Maier, E. (2015). Parasitic attitudes. *Linguistics and Philosophy* 38. 205-236.
- Morgan, J. (1970). On the criterion of identity for noun phrase deletion. *Papers from the Sixth Regional Meeting of the Chicago Linguistics Society*, Chicago, IL. pp. 380-389.
- Ninan, D. (2008). *Imagination, Content and the Self*. PhD dissertation, MIT.
- Ninan, D. (2012). Counterfactual attitudes and multi-centred worlds. *Semantics and Pragmatics*, 5, 1-57.
- Pearson, Hazel (2013). *The Sense of Self: Topics in the Semantics of De Se Expressions*. PhD dissertation, Harvard University.
- Pearson, H. & Dery, J. (2014). Dreaming de re and de se: Experimental evidence for the Oneiric Reference Constraint. *Proceedings of Sinn und Bedeutung* 18, 322-339.
- Pearson, H. (2015). The Interpretation of the Logophoric Pronoun in Ewe. *Natural Language Semantics* 23, 77-118.
- Percus, O. (2006). Uninterpreted pronouns? Unpublished manuscript.
- Percus, O., & Sauerland, U. (2003b). Pronoun movement in dream reports. *NELS* 33:

*Proceedings of the 33rd Annual Meeting of the North East Linguistic Society*, MIT, Cambridge MA. pp. 347-366.

Percus, O., & Sauerland, U. (2003b). On the LFs of attitude reports. *Proceedings of the Conference "Sub7 - Sinn Und Bedeutung"*, University of Konstanz. pp. 228-242.

Quine, W. V. (1956). Quantifiers and propositional attitudes. *Journal of Philosophy*, 53, 177-187.

Reinhart, T. (1990). *Self-representation*. Unpublished manuscript.

Schlenker, P. (1999). *Propositional Attitudes and Indexicality: a Cross-categorical Approach*. Unpublished PhD, MIT.

von Stechow, A. (2002). *Binding by verbs: tense, person and mood under attitudes*. Unpublished manuscript.

von Stechow, A. (2003). Feature deletion under semantic binding. *Proceedings of the North East Linguistics Society*.

Yanovich, I. (2011). The problem of counterfactual de re attitudes. *Proceedings of Semantics and Linguistic Theory* 21, pp. 56-75.