Lessons from Descriptive Indexicals

Kjell Johan Sæbø
University of Oslo
k.j.sabo@ilos.uio.no

Two main methods for analysing de re readings of definite descriptions in intensional contexts coexist: that of evaluating the description in the actual world, whether by means of scope, actuality operators, or non-local world binding, and that of substituting another description, usually one expressing a salient or 'vivid' acquaintance relation to an attitude holder, prior to evaluation. Recent work on so-called descriptive indexicals suggests that contrary to common assumptions, both methods are needed, for different ends. This paper aims to show that there is indeed a division of labour between the two methods of analysis and to identify criteria for choosing among alternative ways to model the second, substitutional method.

1   Introduction

There are two general theories of so-called *de re* readings of definite descriptions in modal contexts. In one theory, the description is evaluated at the actual world. In the other, it is replaced by another description which is evaluated at the possible worlds under consideration. There is little tension between the two theories, partly because the first tends to be favoured by linguists while the second tends to be favoured by philosophers, but also because often, if the second theory is at all applicable, its predictions almost coincide with those of the first. Sometimes, however, definite descriptions have readings which can only be captured by the second theory. These readings correspond to the interpretations of indexicals and demonstratives that have been identified under the label 'descriptive indexicals' (Nunberg 1993, 2004, Recanati 1993, 2005, Elbourne 2008, Hunter 2010). Thus, in principle, a definite description has three readings, across all major classes of modal contexts: one *de re* under the first theory, one *de re* under the second theory, and one *de dicto*. The primary goal of this paper is to make a convincing case that the second reading exists and thus to show that the second theory is indispensable.

Consider, as a case where both theories are applicable, the following scenario, slightly adapted from a news story in the *Anchorage Daily News* of June 2, 2006:

(1)   A large black bear broke into an Anchorage home early this morning, rummaged around like a burglar and feasted on a box of chocolates before

the homeowner shot him dead with a Glock.
The bear entered the house around 2:30 a.m., according to police.
The owners were asleep when their rottweiler started barking wildly. The bedroom door was closed. Outside, the couple could hear things being knocked over. Police spokesman Lt. P. Honeman said the owners initially thought the bear was a burglar.

My intuition, shared by many, is that in the context described, (2) means something like (3), more precisely, *the bear* has an interpretation that more or less coincides with a plain, *de dicto* interpretation of the underlined term in (3).

(2)   The owners initially thought the bear was a burglar

(3)   The owners initially thought the creature setting the dog barking and knocking things over was a burglar

As I will try to show, in displaying this sort of interpretation, (2) is representative of a range of examples of definite descriptions in intensional contexts.

This reading of *the bear* is not a *de dicto* reading. Is it a *de re* reading? That depends—on the first type of theory mentioned above, it is not. There, the description is evaluated in the actual world. The classic is the scopal analysis due to Russell (1905) (Neale 2005 offers a recent restatement), where the DP (determiner phrase) is really outside the intensional context. Today, what Schwager (2011) calls the **transparent evaluation** approach dominates, especially the formulation in terms of 'free index-binding' (a key source is Percus 2000), the theory that von Fintel and Heim (2011, pp. 102ff.) refer to as the 'standard solution':[1] the variable for the world of evaluation for the description may be bound by a non-local world variable abstractor, to be, eventually, instantiated by the actual world. Applied to the example (2) above, this analysis predicts that the argument of *thought* is the proposition outlined in (4):[2] that the individual denoted by *the bear* in the actual world v (the context world, or world of utterance) is a burglar.[3]

(4) $\{\, w \mid \text{the bear}_v \text{ is a burglar}_w \,\}$

I will subsume these theories under the label **Transparency Theory** (TT), and call the reading according to TT the **transparent interpretation**.

As it turns out, the TT analysis of (2) and the hypothesis that (2) in the

---

[1] See Heim (2011): 'it is broadly agreed nowadays that something essentially like the Index-Binding account must be available.' An early source is Bäuerle 1983.

[2] It is assumed in this paper that propositions are sets of worlds; however, although for reasons of simplicity many formulations may suggest otherwise, the key points do not require that the semantic values of declarative sentences be equated with such sets, as warned against by Soames (1987). The necessary assumption about the semantic values is that propositions as sets of worlds or other truth-supporting circumstances be at least recoverable from them, as on the theory of structured meanings developed by Cresswell (1985). (Note that this is not to deny the view, defended by Elbourne (2010), that the semantic value of a declarative sentence may be a set of truth-supporting circumstances after all.)

[3] I will, following Elbourne (2013), Heim (2011), and Barwise and Perry (1983), take definite descriptions to denote individuals and to express (partial) individual concepts.

context described in (1) has an interpretation that coincides with the *de dicto* reading of (3) give different predictions. We get two different readings where one can be false while the other is true. Most notably, it may be that (2) is false on TT while (3) is true. And this case is representative of a general pattern, indeed, in several cases, the difference between the observed reading and the reading according to TT is clearer, and TT itself is more counterintuitive, than here.

This needs to be argued carefully, and that will be done in section 2. Here, I will defend the tripartite hypothesis that (i) there are 'descriptive' interpretations of all referential terms, that is, there are readings of proper names and of definite descriptions that correspond to those of 'descriptive indexicals'; (ii) these interpretations cut across all the major classes of modal contexts (attitudes, modals, counterfactuals); and (iii) they are neither 'de dicto' readings nor are they 'de re' readings in the sense of TT, that is, they are not transparent interpretations.

I will argue in section 3 that the theory that this reading calls for falls within an alternative type of theory where *de re* descriptions are strictly not evaluated as they stand. This type of theory, which can be traced back to Quine (1956) via Kaplan (1968), Lewis (1979), and Cresswell and von Stechow (1982) (recent work in this tradition includes Aloni 2005 and Maier 2009a), has largely been motivated by cases where rigid designators, names or indexicals, give rise to puzzles if they are taken 'at face value', that is, if their meaning is not in some way enriched by, for example, some relation that the designated individual stands in to the holder of a propositional attitude.[4] Crudely, the concept expressed is replaced by another, coextensional concept (one designating the same individual in the actual world), which is then given a *de dicto* reading. The term *de re* is thus not quite apt; we have a reduction of *de re* to *de dicto* under substitution. Applied to the example above, this type of analysis might, provided that *the bear* and *the creature setting the dog barking and knocking things over* denote the same individual in the actual world, predict that the argument of *thought* is the proposition outlined in (5):[5]

(5) $\{\,w \mid [\,\text{the creature setting the dog barking and knocking things over}\,]_w$
        is a $\text{burglar}_w\,\}$

I will subsume these theories under the label **Substitution Theory** (ST), and call the reading according to ST the **substitute interpretation**.

So we need a theory of *de re* in the sense of ST—more precisely, a formulation of ST which does not make reference to any particular type of modal operator. Certain versions of ST, notably theories of 'relational belief' (Kaplan 1968) and 'hidden-indexical theory' (Schiffer 1992), assume that certain intensional

---

[4]Cf. van Rooy and Zimmermann (1996, p. 131): 'it is not only important to know what object the belief attribution is about (as is assumed by a Russellian account), but also *the way* the believer thinks about that object.'

[5]To be sure, there is a vagueness involved in determining the description to be substituted; see the discussion in Sect. 4.2.1.

operators, notably epistemic attitude verbs, may introduce substitute concepts ('identifiers' based on 'acquaintance relations', or 'modes of presentation'). But because the interpretations we will be considering cut across all the major classes of modal contexts, it will not do to make them dependent on attitudes. Rather, they should be made to result from operations on what is expressed or denoted by the referential terms themselves.

After reviewing several recent versions of ST in this light in section 3, namely, Maier 2009a (*presupposed acquaintance*), Aloni 2005 (*conceptual covers*), van Rooij 2006 (*counterparts*), and Hunter 2010 (*intensional reconstruction*), in section 4 I sketch and critically discuss a formulation which can perhaps best be viewed as a compositional version of Aloni's theory, but which also incorporates features from the other representatives of ST. In this way, I show that a compositional formulation of ST can be given; I should stress, though, that my main concern does not lie with this particular implementation but in delineating a general conception of ST, distilling a synthesis from proposals that have been developed from different perspectives, all of which are close in spirit.

First of all, I discuss recent work on 'descriptive indexicals' by Elbourne (2008) and provide evidence to show, following Hunter (2010), that they are part of a larger picture than assumed there, a picture which includes names and descriptions, where the 'descriptive' interpretation is distinct from the transparent interpretation and amounts to a substitute interpretation.
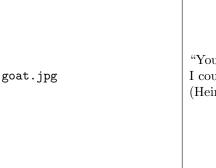
2   Descriptive designators

There is a class of cases of designators in intensional contexts where a substitute interpretation is commonly assumed to exist, beside the rigid reading, namely, the **descriptive** (uses of) **indexicals**. Here, there is a strong and widely shared intuition among linguists and philosophers that, for example, a pronoun like *you* has an interpretation on a par with a contextually evoked definite description. As the term implies, the assumption has been that these uses are limited to indexicals. Recent work suggests, however, that there are parallel uses of all potentially referential terms, including definite descriptions. Here, what corresponds to the straight, rigid reading of the indexicals is the transparent interpretation. Thus to the extent that the straight, rigid reading of the indexicals is judged counterintuitive, a transparent interpretation of the definite descriptions must be judged counterintuitive too. This already shows that Transparency Theory is insufficient. Moreover, in certain scenarios, the transparent reading can be argued to all but coincide with the *de dicto* reading and to be an absurd reading. I will take this to motivate a general notion of substitute interpretations.

2.1   Descriptive indexicals

The uses of pronouns like *I* (or other indexicals or demonstratives) in (6) or (8) were identified as 'descriptive' by Nunberg (1993) and Recanati (1993); more

recently, these cases have been studied by Nunberg (2004), Recanati (2005), Elbourne (2008), and Hunter (2010).[6]

(6)   Before the mother goat goes out, she instructs the little kids not to open the door to a stranger: "If somebody knocks, ask him to show his hoof in the window, and open the door only if you recognize the hoof as mine." But since she doesn't trust them, she decides to put them to the test. She returns and knocks, and the little kids open the door immediately. She chides them and says:

goat.jpg

"You shouldn't have opened the door! I could have been the Wolf!"
(Heim 2004)

(7)   I used to run the Davos summit program. [...] I said, in ten years' time, half of you will be women. (The 90% in the audience looked uncomprehendingly.)

(8)   As the challenged, I am traditionally allowed the choice of weapon.

The significant fact about (6) is that the interpretation that the speaker could have been identical to the Wolf is overshadowed by the one that the creature knocking on the door could have been him; by the same token, the only sensible interpretation of (7) is that in ten years' time, half of the participants at the Davos World Economic Forum Summit will be women, and what is meant by (8) is that in all situations in keeping with tradition concerning duels, the challengee is allowed to choose the weapon.[7] In the words of Bezuidenhout (1997, p. 384):

> Suppose you are alone in the house and not expecting your spouse to return for many hours. You hear a loud crash as the door bangs open, and you think that an intruder has entered your house. Just then your spouse walks into the room, and you say "I thought you were a burglar". Here it is plausible to say that what you thought was that the person

---

[6]The examples (7), (8), (18)–(23), (28), (29) and (34) are all authentic, but url source references are omitted for parsimony.

[7]A reviewer has noted that the adverb *traditionally*, which also figures in Nunberg's (1993) perhaps most well-known example (*I am traditionally allowed to order whatever I like for my last meal*), is not necessarily an adverb of quantification but might mean something like 'tradition dictates that', in which case the pronoun could be plainly referential. Unequivocal examples of adverbs of quantification (like *always* or *mostly*) are not easy to find attested, but see the constructed case (11) below, as well as Sect. 4.3.1.

who crashed through the door was a burglar. You did not entertain the thought that your spouse was a burglar.

On the same note, Elbourne (2008) states that (9) is potentially ambiguous: if John, an advocate of Intelligent Design, has been elected to the Kansas State Board of Education, known for changing hands between advocates of Intelligent Design and evolutionists, and says (9) to his fellow members, who are all like-minded, it has a reading equivalent to (the *de dicto* reading of) (10).

(9)   We might have been evolutionists

(10)   The members of the Board might have been evolutionists

> On one reading, he is saying of the actual members of the Board that they might have been evolutionists. On another reading, however, he is saying that there might have been other, more Darwin-friendly, people elected to the Board. This last reading, at least, seems to require *we* to be synonymous with the definite description *the current members of the Kansas State Board of Education* (Nunberg 1993, pp. 13–14). ... There is evidence, then, in favor of the contention that *we* can have readings equivalent to those of definite descriptions. (Elbourne 2008, p. 420)

Such a contention may not be uncontroversial; Recanati (1993), for one, argues for a more pragmatic approach. Still, the intuition that in these cases, the proposition expressed is not about one certain individual or group but about whatever individual or group fills a certain role seems to be quite robust.[8] Below, I review a way of modelling this intuition compositionally without positing an ambiguity in the pronoun, namely, the analysis proposed by Elbourne (2008).

Note that descriptive indexicals are not the same phenomenon as what is sometimes referred to as 'descriptive pronouns' or 'E-type pronouns' (see Evans 1977; Neale 1990; Elbourne 2001; van Rooij 2006, pp. 106–16; and Büring 2011): basically anaphoric third-person pronouns which seem to 'copy' or 'piece together' descriptive material from discourse antecedents, rather than acting as variables. Although there is a common core, descriptive indexicals only manifest themselves in intensional contexts, and they are not anaphoric, at any rate not in the sense that they depend on antecedent material in the preceding discourse.

2.2   A model of deferred reference

As part of his theory of demonstrative determiners and pronouns, Elbourne (2008, pp. 421f.) develops an analysis of descriptive indexicals based on the theory of deferred reference advanced by Nunberg (1993). The key idea is that deictic pronouns are definite articles with an underspecified 'relational component', encoded in a variable R, which together with a 'deictic component', encoded in a variable i which denotes what the pronoun naïvely denotes, expresses a property.[9] Usually, R is set to the identity relation, and this gives the

---

[8]Example (9) is perhaps complicated by an additional collective/distributive ambiguity; (6), however, is quite unambiguous in that regard.

[9]In this framework, properties are functions from individual concepts to propositions.

standard, referential interpretation of the pronoun when R combines with the deictic variable i and the result combines with the pronoun-as-definite-article. But let us say that I utter (11) while pointing at Benedict XVI, and it 'is obvious that I will be interpreted as saying that *the Pope* is usually an Italian'.

(11)   He is usually an Italian

Then the salient relation can be the function from an individual x to the property of holding the office held by x in the world of utterance w_0 at the time of utterance t_0. When this applies to the value of the individual variable i, b for Benedict XVI, it gives the property of holding the office held by b in w_0 at t_0, and when the meaning of the pronoun-as-definite-article applies to this property, the individual concept specified in (12) results.

(12)   the function from a situation s to:
        $\iota x.$ x holds in s the office held by b in w_0 at t_0

And 'Since Benedict XVI is, we can assume, the Pope in w_0 at t_0', that becomes, in effect, the individual concept specified in (13):

(13)   the function from a situation s to whoever is the Pope in s

This being what is expressed by the DP as a whole—what meets the eye or ear as *He*—in the context under consideration, what I am saying with (11) will in fact be that the Pope is usually an Italian.

Elbourne's analysis is compositional, and it avoids positing an ambiguity in the pronoun. Also, it is part of a larger theory where pronouns are interpreted as definite descriptions, supported, not least, by evidence from anaphoric pronouns which call for descriptive interpretations, the 'descriptive' pronouns referred to above. But the analysis can hardly be extended to proper names and definite descriptions, and for that reason, as we will see below, it is not sufficiently general to capture the full range of descriptive uses of referential terms.

2.3   Descriptive readings of names and descriptions

The intensional contexts where descriptive readings of indexicals have been noted in the literature fall into two classes: (i) contexts created by adverbs of quantification, involving intensionality with respect to times or 'cases', see (8) and (11), and (ii) contexts created by modal operators, involving intensionality with respect to possible worlds or situations, see (6) and (9). ((7) exemplifies a context created by a temporal adverbial which is not quantificational.) Hunter (2010, pp. 113ff.) is careful to make this distinction.

Now both Nunberg (2004) and Elbourne (2008) emphasize that only demonstratives and indexicals allow descriptive readings. Nunberg (2004, p. 269) asks: 'Why do only demonstratives and indexicals permit descriptive readings ... ? Or conversely, why can't descriptions have readings like these?' For Elbourne (2008), the fact that proper names do not allow these readings proves that the phenomenon is semantic and not 'merely' pragmatic in nature.

As pointed out by Hunter (2010, pp. 121ff.), however, the contexts adduced by Elbourne and Nunberg to show that only indexicals or demonstratives allow descriptive readings are all contexts of adverbs of quantification, not modal contexts in the narrower sense. Elbourne notes that no one saying (14) can be interpreted as saying that the Pope is usually an Italian:

(14)   (#) Benedict XVI is usually an Italian

And that is probably correct. But when one turns to modal contexts in the narrow sense, the picture changes: someone saying (15) might be interpreted as saying that many, perhaps having Cardinal Peter Turkson of Ghana in mind, felt that the successor of Pope Benedict XVI should have been from Africa.

(15)   Many felt that Bergoglio should have been from Africa

In section 2.3.1 below, I will present and discuss some authentic cases of names with descriptive interpretations in various kinds of modal contexts. In section 2.3.2, I will do the same in regard to definite descriptions.

Two conclusions could be drawn from the differential behaviour of proper names and definite descriptions in frequency contexts and modal contexts: (i) since they do not show descriptive readings in all intensional contexts, they do not have such readings, (ii) since they do show descriptive readings in some (arguably important) intensional contexts, they do have such readings. I think only the second conclusion is tenable. In section 2.4 I will return to the question what further and more general conclusions can be drawn from this.

2.3.1   Descriptive readings of names

Though agreeing with Nunberg and Elbourne that there is an asymmetry between names and indexicals as to the availability of descriptive readings, Hunter (2010, pp. 121ff.) also shows that the former can have such readings in one of the two classes of contexts where indexicals can have them, namely, in narrowly modal contexts (which she calls *intensional reconstruction* cases), for example, under counterfactual epistemic modals, as in (16):

(16)   Nancy Pelosi might have been a Republican

(17)   The Speaker of the House might have been a Republican

The descriptive reading is the one on which (16) is synonymous to (17). It is different from a properly *de re* reading which would take a world where, say, Nancy Pelosi has made a different choice of political party to be true; all it takes to be true is a world with a different congressional election result (to be accurate: the accessibility of such a world).

Below are a sample of attested cases, across a wide array of modal contexts, where a name is evidently intended to have a descriptive reading.

(18)   After all, for all he knew, <u>Michael</u> could have been a burglar, or a murderer.

(19)   Well, I heard a noise downstairs, and I came down, and I thought <u>Justin</u> was a burglar and I hit him with a baseball bat.

(20)   I had a feeling it would be a boy because I wanted it to be a girl. I wanted <u>Talea</u> to be a boy, and I got the opposite. At least I'll have one of each (boy and girl), so I'm pretty happy.

(21)   If <u>Mary</u> had been a boy then yes, I do believe England would have remained a Catholic country.

In each case, this reading is different from the *de re* reading and equivalent to one where the name is read as 'the entity playing the part played by *a*', if *a* is the name; specifically, these descriptions take the forms *the sudden intruder*, *the person making the noise downstairs*, *the baby in my womb back then*, *the child of Henry VIII by Catherine of Aragon*, respectively.

Note that all three major classes of modal contexts are represented here: modals, attitudes, and counterfactuals. True, in (18) the modal *could* is modified by *for all he knew* and in (21) the counterfactual is in the scope of the attitude verb *believe*, but that does not seem in any way essential. While epistemic modals, in particular epistemic possibility modals like *could* or *might*, can probably be analysed as a kind of attitude operator (as broadly proposed in von Fintel and Gillies 2011), the same does not hold for counterfactuals, and (22) appears a clear case of a name with a descriptive interpretation in such an 'impersonal' modal context:

(22)   I will tell you what would have been a true miracle: if the President was elected a Pima, Yaqui or even a Choctaw maybe. [...] Like I said it would have been really a true miracle if <u>Obama</u> had been a native American Indian.

If we, somewhat sloppily, interpret the consequent clause *it would have been really a true miracle* in terms of closeness or similarity between worlds, as on a Stalnaker (1968) or Lewis (1973) theory of counterfactuals, this counterfactual says that the closest worlds where Obama is a native American Indian are quite far away. It seems clear that what is judged to only hold true in a distant world is the proposition that the first minority to be elected President of the United States is a native American Indian; although the proposition based on the straightforward, rigid interpretation of the name *Obama* may also require a significant departure from the actual world, this is not what is intended to count as 'really a true miracle'.

It turns out to be less easy to find cases where the modality is factual than where it is counterfactual, but they do seem to exist; in (23), the first occurrence of *Jesus* must be read 'the man standing before Mary Magdalene at the tomb'.

(23)   It wasn't until He called her by her own name—'Mary'—that she knew <u>Jesus</u> was Jesus

Cumming (2007) discusses similar cases, inter alia, (24):

(24)   Biron thinks Katherine is Rosaline

This sentence is intuitively true in the following scenario, based on Shakespeare's *Love's Labour's Lost*: 'Rosaline, Maria and Katherine are en route to a masked ball. Each lady anticipates the attentions of a different suitor, and each suitor has given his lady a favour to wear, by which he hopes to recognize her during the masque. The ladies ... decide to swap favours and so cause the suitors to "woo contrary"' (Cumming 2007, p. 83). Under Cumming's analysis, (24) states that there is some discourse referent that actually refers to Katherine but which Biron thinks co-refers with the discourse referent for Rosaline. 'In the situation described, the witness for this existential is presumably the [discourse referent] denoted by the deictic mental symbol ... that Biron is using to track Katherine at the ball.' Although Cumming calls this reading of *Katherine* a *de re* reading, it seems clear that it coincides with the type of interpretation that Hunter and I call descriptive; the description will in this case be something like *the lady with the Turquoise favour*. And it is *de re* in a sense shared by all the cases cited above: the description has to actually co-refer with the term used, here *Katherine*.

The observations made above appear to run against Kripke's (1972, p. 270) dictum that 'no one other than Nixon might have been Nixon'; in fact, they indicate that in principle, a sentence like *Humphrey could have been Nixon* could be true. Note, however, that they do not challenge the thesis that names are rigid designators as far as their literal meaning is concerned, only that their literal meaning is the only interpretation they can have. Note also that the descriptive readings of names at issue here are different from 'descriptive names' in the sense of Evans (1982) and from names under a 'descriptivist' theory; any name can have a descriptive reading in the right circumstances, and its literal meaning may or may not be treated in a descriptivist way.

2.3.2   Descriptive readings of definites

Hunter (2010) goes on to show that 'intensional reconstruction' is possible not only with indexicals and names but with definite descriptions too, as in (25).

(25)   You should've checked the peephole. <u>Your mom</u> could've been a burglar.

Hunter remarks on this case that 'I get an IR [(intensional reconstruction)] style reading of [(25)] according to which it's not literally true that John's mom could have been a burglar, but that it might have been that someone else was knocking on the door and that person was a burglar' (Hunter 2010, p. 123).

It is not difficult to construct natural correspondences to the cases of names rendered above ((18)–(22)), spanning a range of modal contexts, in the form of definites. (26) (based on Ingstad (1933), *The Land of Feast and Famine*) sums up a situation where the protagonist has been trying to lure a moose to his post by imitating a moose's love call with a birchbark trumpet; the response calls he hears and attributes to a moose turn out to originate from an Athabascan Native in an approaching canoe imitating a moose's love call to attract a moose.

(26)   I had thought <u>the Indian</u> was a moose, and he had thought that I was

On the prominent and intended reading, what I had thought was that the creature emitting the love calls—whichever entity that was—was a moose.

Similar cases have been cited in the literature, independently of the debate about descriptive indexicals. Bonomi (1995), for instance, discusses (27):

(27)   Swann wants to kill the chief of the army

> Swann has come to the conclusion that his wife Odette has a lover, but he has no idea who his rival is ... he decides to kill his wife's lover, and he confides his plans to his best friend, Theo. ...
> Odette's lover is Forcheville, the chief of the army, and Theo is a member of the security staff which must protect Forcheville. During a meeting of this staff to draw up a list of the persons to keep under surveillance, Theo ... says [(27)] (Bonomi 1995, p. 167)

In this scenario, the *de dicto* reading is clearly not intended, but nor is the transparent reading, on which Swann intends to kill Forcheville; on the intended interpretation, what he wants to do is kill Odette's lover, whoever that may turn out to be. Note how the transparent and the substitute interpretation (*Odette's lover* read *de dicto*) come apart here; Swann's (belief-based) desire worlds may include many where Odette's lover is not Forcheville but somebody else.

Consider next a counterfactual:

(28)   A 2004 case heard by the Kansas Supreme Court had civil libertarians and gay rights groups protesting the existence of a double standard. Matthew Limon was a mentally disabled 17-year-old when he had consensual sex with a 14-year-old boy. Under the Romeo and Juliet law enacted in Kansas in 1999, Limon would have been sentenced to 15 months in prison if <u>the boy</u> had been a girl.

The closest worlds where the individual denoted by *the boy* in the actual world is a girl instead of a boy are not relevant; what we mean is that in the closest worlds where the 14-year-old person Limon had consensual sex with was a girl, his sentence was much less harsh than in reality.

Consider, finally, the counterfactual (29) (the context is a discussion of a snowsled accident where a boy, driving in snowdust and straying from the trail, has hit a tree; the debate is about whether or not the tree was to blame and should be cut down, this discussant arguing that no, the driver was responsible):

(29)   What if <u>the tree</u> had been a Moose, a deer or another sled, would this still have happened? I believe it would have, ...

We are not interested in possible worlds where the individual that is actually the tree is a moose or a deer or another sled instead, we are interested in possible worlds where what obstructed the course of the snowsled driven by so-and-so at such-and-such a time was not, as actually, the tree but a moose or a deer

or another sled. That is to say, the relevant reading is here again a descriptive reading of the definite description.

In sum, Hunter (2010) seems right: indexicals are not so special after all; there are clear signs that the intuitions that motivate a theory of descriptive indexicals provide motivation for a theory of descriptive referential terms in general—at least as far as modal contexts are concerned. Quantificational adverb contexts are different, a fact to which I will return in section 4.3.1.

2.4 Preliminary conclusions

The point I want to make now is that the 'plain', implausible, and unwanted interpretation of the indexicals in the relevant contexts corresponds to the *de re* interpretation of the definite descriptions under one theory, namely, Transparency Theory (TT), while the plausible 'descriptive' interpretation corresponds to the *de re* interpretation under Substitution Theory (ST). The transparent and the substitute interpretation are distinct in these contexts. In addition, there is throughout a 'normal' *de dicto* reading of the definite description—though this reading is often implausible, or even absurd, as in (26), (28), or (29) above.

Let me illustrate the parallel between indexicals and definites with a simple pair of examples, (30) and (31):[10]

(30)   I could have been a burglar

    Reading 1: plain (transparently *de re*);
    'my brother is a burglar, so I could easily have become one myself' or
    'I'm so deft at picking locks, I would have made a good burglar'

    Reading 2: descriptive, *de dicto* under substitution;
    'the person at the front door could have been a burglar instead of me, so you shouldn't have opened it without first checking who was there'

(31)   The pedlar could have been a burglar

    Reading 1: transparently *de re*;
    'his brother is a burglar, so he could easily have become one himself' or
    'he's so deft at picking locks, he would have made a good burglar'

    Reading 2: descriptive, *de dicto* under substitution;
    'the person at the front door could have been a burglar instead of a pedlar, so you shouldn't have opened it without first checking who was there'

There is also in principle a third, *de dicto* reading of (31), saying that there is some possible world w such that the pedlar in w is a burglar in w as well, a 'pedlar-burglar'. This reading is rather unnatural, though, and it may seem only marginally different from Reading 1. Note, however, that the *de dicto* reading

---

[10]Admittedly, the two readings of (30) or (31) will hardly coexist in the same context; the distinction will correlate with distinct settings for the modal's contextual parameters. The point is that one modalized clause can have both a natural transparent interpretation and a natural substitute interpretation.

is the only natural reading of (32), which provides a paraphrase of Reading 2 of example (31):

(32)   The person at the front door could have been a burglar

To see the distinction between Reading 1 and Reading 2 of (31) more clearly, consider a rough formalization of the argument proposition of the modal *could* in two versions, (33 i) explicating Reading 1 and (33 ii) explicating Reading 2.

(33)   i. $\{\, w \mid \mathrm{burglar}_w(\iota x.\, \mathrm{pedlar}_v(x)) \,\}$

    ii. $\{\, w \mid \mathrm{burglar}_w(\iota x.\, \mathrm{person\text{-}at\text{-}door}_w(x)) \,\}$

       provided $\iota x.\, \mathrm{person\text{-}at\text{-}door}_v(x) = \iota x.\, \mathrm{pedlar}_v(x)$

Here v is the actual world (the unshiftable context world, or world of utterance). The proviso that $\iota x.\, \mathrm{person\,at\,door}_v(x) = \iota x.\, \mathrm{pedlar}_v(x)$ means that when you use *the pedlar* in the sense of 'the person at the front door', that is on the presupposition that these two individual concepts are coextensional.

It should be clear that the two readings of (31) spelt out in (33 i) and (33 ii) differ from each other in the same way as the two readings of (30) and thus that the 'descriptive reading' of the indexical I corresponds directly to the substitute interpretation of the definite description *the pedlar*. Hence, if the plain interpretation of the indexical is counterintuitive, then so is the transparent interpretation of the definite description.

The case against Transparency Theory can still be strengthened. It has been noted that the 'third', *de dicto* reading, while always available in principle, is sometimes absurd (see (26) and (28)). Now in certain cases of attitude reports, the transparently *de re* reading can be argued to coincide, as far as the truth conditions of the reports are concerned, with the absurd *de dicto* reading and so to be equally absurd. Consider the following:

(34)   The experienced hunter [Mary Beth Harshbarger] has maintained for years that she mistakenly thought <u>her husband</u> was a black bear when she shot him through the chest with a hunting rifle in the dim light of a late-summer evening in the central Newfoundland woods on 9/14/06.

On the *de dicto* reading, the argument proposition of *thought* is the set of worlds w such that Mary Beth Harshbarger's husband in w was a black bear in w, and to think (in the sense of believing) that is irrational. On the transparent reading, the argument proposition of *thought* is the set of worlds w such that Mary Beth Harshbarger's husband in v, the actual world, namely, Mark Harshbarger, was a black bear in w—which is different. But, for Mary Beth Harshbarger to think that is arguably just as irrational, since arguably, in all the worlds w compatible with what she thought in v (in all of which, following Hintikka's (1969) semantics for propositional attitudes, the argument proposition of *thought* must be true for the attitude report to be true), Mark Harshbarger was the value of the term *her husband* at w; hence the transparent interpretation and the *de dicto* interpretation arguably coincide as far as the truth conditions of the complex

sentence are concerned, and both are absurd: one and the same individual must be both her husband and a black bear in Mary Beth Harshbarger's thought worlds. The only sensible reading of the embedded clause is (roughly) that the dark shape she saw moving out of the forest toward the pickup truck where she was waiting, although actually her husband, was a black bear.

Actually, to show that TT gives implausible predictions it suffices to assume that in *some* of Mary Beth Harshbarger's thought worlds Mark Harshbarger was her husband; conversely, to avoid ascribing an irrational thought to Mary Beth Harshbarger, the Transparency theorist is committed to assuming that in *none* of her thought worlds w, she knew her husband for whom he was in the sense that Mark Harschbarger was the denotation of *her husband* at w—and that seems like an unrealistic commitment to make.

In other words, substitute interpretations and a Substitution Theory are indispensable, not just for pronouns, but also for names and definites—indeed, for referential terms in general.

Having established that, the question is how to formulate such a theory so as to derive a result like (33 ii) in the right circumstances. The theory must be sufficiently general—not tailored to a specific class of referential terms, or a specific class of intensional contexts—and it should be compositional. Existing proposals tend to fall short of one or both of these two criteria. The theory should, moreover, not overgenerate, and it should be plausible. These needs and ways to meet them form the topic of the next section.

3   Substitution Theory

There are several recent formulations of the type of theory of *de re* interpretation that I call Substitution Theory: Aloni (2005), Maier (2009a), Schwager (2011), *inter alia*; Elbourne (2008) (see Sect. 2.2) can also be associated with this type of theory. All these approaches provide valuable insights, but none is both sufficiently general regarding the relevant terms and contexts and compositional. Below I review the proposals of Maier (2009a) (Sect. 3.2) and Aloni (2005) (Sect. 3.3), concluding that the former is not general enough and the latter, while general, is not compositional. I also assess, though it is not strictly a version of Substitution Theory, the recent formulation of counterpart theory by van Rooij (2006) (Sect. 3.4), as well as the analysis of descriptive readings given by Hunter (2010) (Sect. 3.5). First of all, though, I discuss a kind of analysis that may be thought to solve the problem of descriptive indexicals, namely, 'diagonalization' in the sense of Stalnaker (1978), concluding that it is not sufficiently general regarding the relevant terms, and that even concerning indexicals, this method has serious limitations.

3.1   Diagonalization?

The reinterpretation procedure known as *diagonalization* in a two-dimensional semantics, as introduced by Stalnaker (1978), is widely held to be a good tool

for dealing with sentences where indexical terms (in a wide sense) threaten to cause triviality (see Zimmermann 2012, pp. 2374ff. for a survey), like (35).[11]

(35)   (Open the door, my dear little goats!) I am your mother!

Diagonalization transforms a meaning which depends on the context of utterance to one which only depends on the context of evaluation, by making the latter play the role of the former; thus effectively, *I* is reinterpreted as *the speaker*. So, although as it stands (if *your mother* is understood as a rigid term), (35) is either trivially true or trivially false, its diagonal proposition is informative.

Cases like (35) are reminiscent of cases like (36), with descriptive indexicals: here, too, the intended interpretation of *I* is as something whose extension can vary across contexts of evaluation. Therefore, diagonalization might be thought to offer a solution to this problem as well.

(36)   "You shouldn't have opened the door! I could have been the Wolf! If I had been the Wolf, I would have eaten you all by now!"

While a treatment of this in the style of Elbourne 2008 would be based on the idea that *I* is interpreted as 'the creature knocking on the door a moment ago', the point of a diagonalization treatment would be to interpret *I* as 'the speaker a moment ago'. After diagonalization, the role of the situation of utterance is played by the situation of evaluation, so that the situation the pronoun depends on will be shifted in the modal context. Thus the argument proposition of the modal *could* and that of *if* will become the set of situations s such that (the time is a moment before the utterance time and) the speaker of s is the Wolf— not very different from the proposition resulting from a treatment in the style of Elbourne: the set of situations s such that (the time is a moment before the utterance time and) the creature knocking on the door in s is the Wolf.

The chief problem with this strategy is that there will not always be any speaker in the relevant possible situations. It just so happens that in the possible worlds witnessing the truth of the second clause of (36), the Wolf was speaking, saying (35). But far from all situations (with worlds and times defined for them) are utterance situations (with speakers defined for them as well); the burglar of (30) (reading 2) will typically not be likely to utter anything.[12]

This approach has another decisive drawback: diagonalization only makes a difference for terms whose intensions vary with the situation of utterance— indexicals, demonstratives, proper names perhaps too; but hardly or at least not unconditionally for definite descriptions. For the bulk of these, if they do not contain indexicals, demonstratives, or names, diagonalization will not affect the content, leaving us with the ordinary *de dicto* reading in the modal context.

---

[11]The case is presented by Heim (2004) as one where diagonalization is indispensable. The context is, again, the fairy tale where the seven little kids cannot know whether the creature knocking on the door is their mother or the Wolf.

[12]There may also be conceptual problems of the sort voiced by Büring (1998): 'modals . . . do not use the diagonal of their argument's meaning . . . '

In other words, the diagonalization strategy is too narrow in scope regarding the terms that display descriptive readings in intensional contexts. The next approach to be considered is sufficiently wide in scope in that regard, but not regarding the range of intensional contexts where descriptive readings arise.

3.2 Maier: presupposed acquaintance

The main line of tradition for Substitution Theory is the theory of 'relational (as opposed to 'notional', *de dicto*) belief' going back to Quine (1956) and Kaplan (1968). Here the aim has been (at least since Kaplan 1968) to modify the interpretation of a term embedded in an attitude report, for example, replacing a constant individual concept by a 'vivid name', an 'identifier', or a 'mode of presentation', typically a nonconstant individual concept based on an 'acquaintance relation' between the holder of the attitude and the value of the original individual concept, the so-called *res*.

To attain this, it has seemed necessary to somehow gain access to the term, and this has proven difficult to do compositionally without assuming either a syntactic movement violating island constraints or a nonstandard logical type for the attitude verb (different from $\langle\langle s,t\rangle,\langle e,t\rangle\rangle$). Thus Cresswell and von Stechow (1982) assume the type $\langle e,\langle\langle s,\langle e,t\rangle\rangle,\langle e,t\rangle\rangle\rangle$, where the *res* term's denotation is the first argument, while Percus and Sauerland (2003) assume the type $\langle\langle\langle e,\langle s,e\rangle\rangle,\langle s,t\rangle\rangle,\langle e,t\rangle\rangle$ (where $\langle e,\langle s,e\rangle\rangle$ is the type of a 'concept generator').

Maier (2009a), using a standard type for *believe*, analyses the *de dicto*/*de re* ambiguity in terms of presupposition resolution (binding or accommodation) in a DRT (Discourse Representation Theory) framework. A presupposition triggered in a belief report can be resolved globally to derive a *de re* reading, or locally to derive a *de dicto* reading.

> I adapt the presupposition resolution algorithm to incorporate the idea that a *de re* representation involves a context-dependent acquaintance relation/mode of presentation. ...The adapted resolution algorithm specifies that whenever a presupposition moves out of a belief, it triggers the introduction of such an acquaintance presupposition. (pp. 471f.)

Compositionality and a uniform logical type are preserved because all terms that can have *de re* interpretations can be taken to trigger presuppositions, presuppositions that in turn trigger acquaintance relation presuppositions iff they are resolved non-locally. Let us look at one of Maier's examples.

(37)   John believes the president of PepsiCo is rich

> [(37)] could mean two very different things. First, it might mean that John believes the proposition that whoever is the president of PepsiCo is rich, on the basis of, say, his general belief that presidents of large corporations are always rich. This is the *de dicto* reading ...But there is an additional reading of the sentence. John believes of a certain individual who is the actual president of PepsiCo that she is rich. On this *de re* construal, the

report is felicitous if John just talked to Indra Nooyi at a party about her yacht, and believes "wow, this woman is rich". (Maier 2009a, pp. 432f.)

If the presupposition of the definite description *the president of PepsiCo* is not resolved locally (which would result in the *de dicto* reading), this triggers a presupposition with, in addition to a discourse referent y for the president of PepsiCo, a discourse referent R for an acquaintance relation between x for John and y. R may be bound to (38), if the previous discourse provides this relation.

(38) $\lambda x \lambda y$ y is a woman that x just talked to at a party about her yacht

This results in a relational *de re* interpretation where what John is claimed to believe is, in effect, that the woman that John just talked to at the party is rich.

The analysis is unified in the sense that it generates uniform, but under-specified, preliminary representations. It is pragmatic in the sense that the pragmatic presupposition resolution mechanism does all the work; the syntax and semantics are straightforward. But the price to be paid is a relatively complex definition of presupposition binding and accommodation making explicit provision for contexts of belief; here is, one might contend, a residue of noncompositionality.[13]

As a representative of Substitution Theory, Maier's theory is general with respect to the relevant terms: pronouns, names, definite descriptions. However, it is not general when it comes to the relevant intensional contexts: in line with traditional work on *de re*, it confines itself to attitude reports. This is natural; the notion of an acquaintance relation supposedly underlying relational *de re* readings only makes sense in the setting of attitudes, where the relation can hold between the attitude holder and the *res*.

However, Kaplan's (1968) notion of acquaintance has been criticized, by, for example, Bonomi (1995) and Salmon (2009), as too strict; often the attitude holder's acquaintance with the *res* is very slight. It may in fact be that the presence of an acquaintance relation is more a matter of observational fact about *de re* readings in belief contexts than a precondition for such readings in general. Note that on Maier's analysis, though x (the holder of the belief) is one of the two arguments of R (the *res* y is the other), the key factor is that the R referent can find an antecedent; the notion of acquaintance is thus in a sense shifted to discourse level, where it reappears as *familiarity*. Note also that the 'relational component' in the theory of Nunberg (1993) and Elbourne (2008) does not depend on acquaintance but on *salience*.

In any case, to account for a substitute interpretation of a term in a modal context which is not an attitude context but an 'impersonal' modal context, say, a counterfactual like (22), (28), or the last sentence in (36), acquaintance on the part of a person referred to there cannot be a criterion, since there is not always any person referred to there, or if there is it may not be clear that it is that person or group that acquaintance is to be attributed to—indeed, if

---

[13] A proposal which is based in Transparency Theory but none the less resembles Maier's in using presupposition resolution to derive the two readings is given in Romoli and Sudo 2009.

Maier's theory is to accommodate such substitute interpretations, his dyadic R relation must be supplanted by a monadic P property, or an individual concept.

To be sure, this P must be constrained in some way; below, a method is discussed which moves the focus from the agents of reported attitudes to the agents of reports or, more generally, utterances, that is, to speakers, in a fully general (though noncompositional) theory of substitute interpretations.

3.3 Aloni: conceptual covers

Aloni's theory of conceptual covers (Aloni 2005) can be considered a theory of substitute interpretations, and as such it is general both in regard to intensional contexts and in regard to referential terms.[14] A conceptual cover is a set of individual concepts which for each individual d and each world w contains exactly one concept c such that $c(w) = d$. The key idea is that what we evaluate *in situ* is a contextually and pragmatically constrained individual concept which coincides extensionally with the *res* term.

Aloni's theory is not a compositional one. The central interpretation rule 'Quantification over contextually selected conceptual covers' (Definition 4) assumes a logical representation where a type $\langle s, e \rangle$ variable identified with the *res* term is existentially quantified. For (39), the representation is (40):

(39)   Al believes that the boss is a spy   (Aloni 2005, p. 507)

(40) $\exists x_n (x_n = b \,\&\, \mathcal{B}_a(spy(x_n)))$ (the notation is slightly altered)

In words: there is an individual concept x_n whose value at the actual world coincides with that of the concept expressed by *the boss* and which is such that in all Al's belief worlds, x_n is a spy.

Now if a standard interpretation rule is applied to (40) it will be too weak and overgenerate *de re* readings. Therefore, existential quantification over individual concepts is restricted to contextually selected conceptual covers.

(41) $M, w, \wp \models_g \exists x_n \phi$ iff $\exists c \in \wp(n) : M.w, \wp \models_{g[x_n/c]} \phi$

To paraphrase: (40) is true relative to the 'conceptual perspective' $\wp$ and in w if and only if there is an individual concept c in the conceptual cover $\wp(n)$ such that $c_w = b_w$ and $c_{w'}$ is a spy in all belief worlds of Al's w'.

> *De re* attitude reports are analysed as quantified modal sentences . . . x_n
> is taken to range over the set of . . . concepts pragmatically selected for n,
> rather than over the domain of individuals. In this way their interpretation
> is made dependent on the conceptualization of the universe of discourse
> which is contextually operative. (Aloni 2005, p. 512)

In fact, it follows from the definition of conceptual covers that any cover $\wp(n)$ will contain exactly one concept whose value at the actual world coincides with

---

[14]Though she actually only considers propositional attitudes and questions, there is nothing in the theory to bar substitute interpretations in other intensional contexts.

the value of the intension of a given *res* term, so that the proposition expressed by (39) relative to $\wp$ and n could be formalized as (42):

(42) $\{\, v \mid \mathcal{B}_v(\{\, w \mid spy_w([\iota x.\, x(v)\!=\!b(v) \,\&\, \wp(n)(x)]_w)\})(a)\}$

The object of belief can be put into words as: the set of worlds w such that

> the value of <u>the pragmatically selected concept whose value at v equals that of the one expressed by *the boss*</u> at w is a spy in w

The underlined definite description is a referential term, hence the need for existential quantification is eliminated. In fact, this move, not made by Aloni, makes a compositional theory appear more likely.

Beside a substitute interpretation, the transparent interpretation of a term like *the boss* ensues if the coreferential concept in $\wp(n)$ is a rigid designator, and the *de dicto* interpretation results if the concept is the one expressed by the same definite description; in general, the literal interpretation is preserved if the coreferential concept in the cover is the one expressed by the original term.

The theory's central element is the contextual 'conceptual perspective' $\wp$, which encodes a variety of factors. When asked how a substitute concept makes itself felt, or, in Aloni's words, how different perspectives are selected on different occasions, it is easy to think in terms of salience; but Aloni puts her main emphasis on pragmatic maxims and principles, such as consistency, informativity, relevance, and interpretative as well as generative economy.

Consider (27), from Bonomi 1995 (see Sect. 2.3.2):

(27)   Swann wants to kill the chief of the army

Bonomi points out that in the perspective of the reporter, Swann's attitude is directed towards the chief of the army, although he is aware that in Swann's perspective, his attitude is directed towards Odette's lover. Aloni would add that one important reason for pairing the expression *the chief of the army* with the interpretation 'Odette's lover' is that the expression *Odette's lover* would be uninformative or irrelevant in the utterance situation. Possibly, moreover, the pairing of the expression *the chief of the army* with the literal interpretation 'the chief of the army' would be inconsistent with the common ground. All this will help shape the conceptual perspective and ultimately determine the concept in the cover that coincides with the chief of the army in the actual world.[15]

3.4   van Rooij: counterpart functions

One more theory merits consideration as a version of Substitution Theory in a wide sense, namely, the recent formulation of Counterpart Theory due to van Rooij (2006, pp. 57ff.). This is a theory where, to put it simply, individuals, not individual concepts, supplant each other when the world of evaluation changes.

---

[15]This is much in the spirit of Schwager (2011, p. 411): 'in addition to the semantic conditions on when a *de qualitate*-report is possible, we need a proper pragmatic theory to explain when and why speakers choose to rely on the replacement rule.'

Belief attributions and other sentences are interpreted with respect to (a context and index world and) a **counterpart function**, which maps an individual d and a world w to an individual (d's counterpart, or 'representative', in w). Like Aloni's conceptual perspective, the counterpart function is a contextual parameter, reflecting aspects of the conversational situation, and it is designed to do much the same job.[16]

Note that the counterpart function applies freely to the denotations of all individual terms, and independently of the context the term is in, whether it is intensional or not; if not, the function boils down to the identity function. Hence, the theory is compositional. The key definition is (p. 65):

(43) $\quad \llbracket \underline{t} \rrbracket^{v,w,c} = c_w(\llbracket \underline{t} \rrbracket^{v,w})$

where $\underline{t}$ is an individual term token, v is the context world and w is the index world in a two-dimensional semantics à la Stalnaker (1978, 1981), and c is the counterpart function. Let us see how the relational *de re* reading of the term *the president of PepsiCo* in (37) as sketched in section 3.2 might be derived.

(37) $\quad$ John believes the president of PepsiCo is rich

Since this term, a definite description, can, in contrast to indexicals or names, be assumed to depend on the index world for its extension, the actuality operator @ (or *dthat*) should first be applied to it, making the context world play the role of the index world. (Otherwise, the counterpart function becomes the identity function and the definite description gets the ordinary *de dicto* interpretation.)

(44) $\quad [\underline{@(t)}]^{v,w} = [\underline{t}]^{v,v}$ $\quad$ (see van Rooij 2006, p. 232)

Then, the argument proposition of *believes* becomes, for a context world v,

(45) $\quad \{ w \mid c_w([\underline{thepresidentofPepsiCo}]^{v,v}) \in [\underline{rich}]^{v,w} \}$

Now all depends on, in the words of van Rooij (p. 66), which 'representation or counterpart we refer to', and this, according to von Stechow (1984) and Stalnaker (1988), in turn depends on which representation 'is most salient in the relevant conversational situation'. In the scenario of Maier (2009a, p. 432) (see Sect. 3.2), this representation is 'a woman John just talked to at a party about her yacht'. The counterpart function at play here can thus be spelt out as (46):

(46) $\quad c = \lambda w \lambda x$ the counterpart to x that is a woman John just talked to at a party about her yacht in w

and for $x = [\underline{thepresidentofPepsiCo}]^{v,v}$, for each world w in (45) the value of (46) is the counterpart to the president of PepsiCo in the actual world that is a

---

[16] According to van Rooij (p. 67), pragmatically there will be a unique most salient counterpart function relative to which a sentence is to be interpreted, but semantically we should existentially quantify over such functions; in a general sense this is readdressed in Sect. 4.2.

woman John just talked to at a party about her yacht in w.[17]

Though this theory may seem rather different from that of Aloni (2005) or Maier (2009a), it appears to produce approximately the same results, and it has the advantage of being general regarding both the referential terms and the modal contexts involved—like Aloni's theory—and also compositional.[18]

There is reason, though, to continue searching for a compositional version of Substitution Theory. The reason is that Substitution Theory can be understood, and has so far been tacitly understood, in the strict sense that the substitute interpretation is the ordinary, *de dicto* interpretation of some (possibly complex, possibly partially covert) expression in the language (like Maier's R, Aloni's x_n), so that the substitution is effected in the language. In this strict sense, since the counterpart function is not in the language, van Rooij's counterpart theory is not a version of Substitution Theory; it could be changed into one, however, by defining an object language counterpart functor, and such a theory would probably, like van Rooij's own, be compositional and in fact close to the one I will sketch in section 4.

3.5   Hunter: intensional reconstruction

Hunter (2010) is alone in offering an analysis expressly designed to capture descriptive readings of indexicals, proper names, and definite descriptions alike in modal contexts.[19] Her analysis bears resemblance to van Rooij's theory of counterparts, but here the counterpart function, or relation, is not a parameter of interpretation but a component of a shifted logical form. In a case like (47a), uttered by the most recently elected Texas Supreme Court Justice, a conflict between a literal interpretation of the consequent and the interpreter's world knowledge will trigger the shift described in (47b), so that the logical form for the consequent of (47a) becomes (47c) (x_a denotes the agent of the utterance).

(47)   a.   If a Democrat had won the last gubernatorial election, then I might have been Mexican
        b. $\lambda P(P(x_a)) \Rightarrow \lambda P \lambda x(x \sim_c x_a \rightarrow P(x))$ (P is any property)
        c.   *Might* $\forall x(x \sim_c x_a \rightarrow Mexican(x))$

$\sim_c$ is the relation obtaining between individuals sharing the contextually salient property c—here that of being the most recently elected Texas Supreme Court

---

[17]Note that there is no principled division between transparent and substitute interpretations of definite descriptions in this theory; since domains of individuals are world-specific, transparent interpretations will be emulated by counterpart functions that use intrinsic and relational similarity (see Lewis 1968, p. 114: 'your counterparts resemble you closely in content and context'), applied to individuals that satisfy the description in the context world.

[18]There is one generalization that it may seem not to capture: the constraint that the 'old' and the 'new' individual concept be coextensional; here, that the woman John talked to at the party is actually the president of PepsiCo. It is possible to argue, though, that this follows from the general constraint on counterpart functions that they do not affect their individual argument when this argument is in the individual domain of their world argument.

[19]For descriptive readings of indexicals in quantificational contexts, Hunter offers a slightly different analysis involving type coercion.

Justice—across worlds, more exactly, between individuals having c in a possible world and individuals having c in the actual world:

> The consequent will be true just in case there is a world w in the modal background provided by the antecedent which is such that all of the individuals who have the property c in w that x_a has in the actual world, are Mexican in w. (Hunter 2010, p. 135)

Though it is not made quite clear how, the shifted interpretation thus requires that the actual world is accessible in the modal context; already in the non-modal right-hand expression in (47b), there must be implicit reference to a shiftable beside an unshiftable world variable, or something to the same effect.

Note, also, that what corresponds to Aloni's 'cover' individual concept is here the contextually salient property c. If c can only hold of one individual (as is the case for the property of being the most recently elected Texas Supreme Court Justice), then to have a certain property c in a world w that an individual y has in a world w_0 is to be the value of the individual concept c', assigning y to w_0, at w, where $c' = \lambda w' \, \iota x \, c_{w'}(x)$. Since it seems generally to be the case that the contextually salient property c can only hold of one individual (though perhaps a plural one, see (9)), the difference between Hunter's and Aloni's proposals is in reality smaller than it may appear.

In turn, the property c corresponds rather closely to the contextually salient individual concept ε figuring in the compositional version of Substitution Theory sketched in section 4. Indeed, Hunter's theory could be spelled out in formal detail in such a way that the result would not be very different from that version. Given that the property c will systematically assign a singleton set of (perhaps plural) individuals to a world, it is not strictly necessary to abstract over the variable x on the right-hand side of (47b) or to universally quantify over it in (47c). The right-hand side of the shift arrow could be written as $\lambda P \, P(\iota x \, c(x))$ or, simply, $\iota x \, c(x)$, resulting in (47d) for (47c), if the shift is made contingent on (here) the identity between $\iota x \, c(x)$ and x_a in the actual world.

(47)   d.   *Might Mexican($\iota x \, c(x)$)*

Like Aloni, Hunter does not offer a compositional account—the shifted interpretation is brought on by a special operation which is not rooted in a word or other constituent of the linguistic input. But this, in turn, could be remedied by defining the shift of interpretation as the content of an optional operator, like the one introduced in section 4.1.

3.6   Interim summary

With a view to a general and compositional theory of substitute interpretations, the four proposals presented in the last subsections all have definite merits. The theories of Aloni (2005) and Hunter (2010) are general in covering the whole range of relevant terms and contexts. They also ascribe pragmatics and world knowledge a significant role in determining whether and how a substitute

interpretation arises. So does the theory of Maier (2009a), but, in addition, this theory comes closer to compositionality.

Aloni does not specify how the logical representations with existential quantification into modal contexts come about. As noted, however, this quantification may be viewed more as an artifact of the theory than an intrinsic property of it. Since it follows from the definition of a conceptual cover that one and only one coreferential concept will belong to the operative cover in any given case, the conditions are met for a conception of the substitute interpretation in terms of a function. Hunter describes substitute interpretations in terms of a shift operation resulting in logical forms with universal quantification over individuals in modal contexts. This universal quantification can also be seen as an artifact of the theory, however, as in the general case, the contextually salient property will yield a singleton set, allowing the substitute interpretation to be represented with an iota term.

The theory of counterpart functions developed by van Rooij (2006) can point the way towards a compositional and general version of Substitution Theory in the narrower sense (where the substitute interpretation is the ordinary interpretation of some expression so that the substitution is effected in the language): the coreferential individual concept could enter into the interpretation in a bottom-up way, irrespectively of whether the context grows to become an intensional context or stays extensional. If it stays extensional, the substituted concept will denote the same individual as the one it was substituted for.

There is a special reason to take a cue from Hunter (2010), who is the only one to directly address the descriptive uses of referential terms in general and who describes them in terms of an operation of reconstruction, or coercion, which only applies under pragmatic pressure. It seems a sensible move to conceive a substitute interpretation as an optional operation, to be applied only when, as it were, need and opportunity arise—where 'need' corresponds, in the hearer's perspective, to reasons to disbelieve an intention to communicate a literal interpretation, while 'opportunity' corresponds to reasons to believe an intention to communicate an alternative interpretation.

These ideas are developed further in the next section.

4   A simple Substitution Theory

In this section, I outline a compositional formulation of Substitution Theory which seeks to incorporate advantageous aspects of Aloni's (2005), van Rooij's (2006), Maier's (2009a), and Hunter's (2010) versions of the theory.

From Aloni comes the idea that the substitution relation is to hold between two coextensional individual concepts (functions from worlds to individuals) and the built-in uniqueness of its second argument which allows us to conceive it as a type $\langle\langle s, e\rangle, \langle s, e\rangle\rangle$ function where the substitute concept is the value: the operation which maps an individual concept x to the contextually selected substitute which assigns the same individual to the context world.

Maier (2009a) contributes the notion that the substitute is presuppositional;

in his framework, this means that there should be a discourse referent which has been previously introduced and can bind it. As a subsidiary option, it can be accommodated. From van Rooij (2006) I borrow the idea that the substitution function can in principle be applied to what is expressed by any referential term. To Hunter (2010), finally, I owe the consideration that the operation is optional and only brought on by pragmatic pressure.

Let us reflect for a moment on what a compositional version of Substitution Theory entails, where Substitution Theory is understood in the strict sense that the substitute interpretation is the normal interpretation of some (possibly complex) expression in the language. For that interpretation to be a function of the meanings of the parts of this expression, the substitution itself will have to be encoded in one of the parts, where the original term is one, not in the larger context or in a special rule or principle. This requirement is not fulfilled in the analyses by Aloni and Hunter as they stand, which makes them prima facie non-compositional. Aloni does not go below full formulae. Hunter does, but her shift from the original type $\langle\langle e, t\rangle, t\rangle$ logical form to a substitute type $\langle\langle e, t\rangle, \langle e, t\rangle\rangle$ logical form (see (47b)) is not anchored in any constituent.

One way of fulfilling the compositionality requirement while predicting the readings that Aloni and Hunter predict is to define a substitution operation as the denotation of a covert substitution operator. The key factor is that the operator is posited in the object language, as an adjunct to the original term phrase, so that the meaning of the whole is still a function of the meanings of the parts, although the term phrase (or determiner phrase, DP) does not ultimately have the meaning it would appear to have on the basis of its overt material.

In section 4.1 I define the substitution operator and show how the operation works in a concrete case. A critical issue is how it is contextually constrained, and in section 4.2 I discuss ways to refine the analysis in this respect. In section 4.3 I address some subsidiary topics and loose ends.

4.1 The substitution operator

Let us assume a covert operator $\mathcal{S}$ which can adjoin to any type e node (i.e. any individual term) at the level of logical form (LF).[20] Its meaning, defined in (48), is partial: it is only defined for a context of utterance c and an individual concept δ if c provides a uniquely salient individual concept ε such that $\varepsilon_c = \delta_c$— that is, the contextually supplied 'new' concept ε and the 'old' concept δ must assign the same individual to the context of utterance, in particular its world (the 'actual world').

To formalize this condition, I use a two-dimensional semantics similar to that used by van Rooij 2006 (see Sect. 3.4), with semantic values depending on (i)

---

[20]I assume that DPs, determiner phrases, are either of type e (names, pronouns, definites) or $\langle\langle e, t\rangle, t\rangle$ (quantifiers); if logical types are not distinguished at LF, $\mathcal{S}$ can adjoin to any DP node, but its adjunction to a quantifier will cause a type conflict and not be interpretable. Technically, bound variables are also of type e, but these individual terms are inappropriate as arguments to the substitution operator for other reasons—reasons too complex, however, to be explained in the present paper.

the context of utterance, consisting of a time and a world, and more, and (ii) the context of evaluation. In this picture, familiar from Kaplan (1977), the meaning (or character) of an expression m maps the former to m's intension (or content), which subsequently maps the latter to m's extension (or denotation).[21]

(48) **The substitution operator $\mathcal{S}$**

$$\llbracket \mathcal{S} \rrbracket^{c,j} = \lambda \delta_{\langle s,e \rangle} \begin{cases} \varepsilon_j & \text{if there is a uniquely salient } \varepsilon \text{ in } c \\ & \text{such that } \varepsilon_c = \delta_c \\ \text{undefined} & \text{otherwise} \end{cases}$$

Let me explain: c is the context (or situation) of utterance and j is the context (or situation, or circumstances) of evaluation (in this choice of symbols I follow von Stechow and Zimmermann 2005). The extension of the operator $\mathcal{S}$ at these two indices is that function from individual concepts δ which is defined iff c supplies a uniquely salient individual concept ε whose value at the context of utterance (in effect, the world and time of utterance) is the same as that of δ, and which then yields the value of ε at the context of evaluation.

The condition that the new concept ε and the old concept δ are coreferential in the context of utterance is what makes it necessary to work with a two-dimensional semantics (alternatively to using two indices, a designated variable for the context of utterance could be employed).

It may be useful to observe the substitution operation $\llbracket \mathcal{S} \rrbracket^{c,j}$ 'in operation', to see what it does with an argument and how its value interacts with other elements in the construction of the meaning of a sentence. In an appendix, I give a compositional derivation of the meaning of a modalized sentence under the substitute interpretation of the indexical I (see appendix).

The individual concept ε may have an unusual logical type for an entity which must be supplied by the context for a meaning function to be defined, but otherwise it is not significantly different from the contextual parameters that have been proposed for encoding, say, standards for relative adjectives, conversational backgrounds for relative modals, domain restrictions for determiners, or Questions Under Discussion (QUDs) for focus-sensitive expressions. It could also be introduced as a free variable in LF; other differences aside, it would then parallel the relation variable R in the theory of Elbourne (2008).[22] The 'metasemantic' issue (see Glanzberg 2007) of how an individual concept becomes uniquely salient in a context is left open for the moment, to be discussed in some more depth in section 4.2. But note that typical instances of

---

[21] One advantage of working with full contexts instead of just worlds is that it allows a uniform treatment of intensional cases whether they are modal (as they mostly are) or temporal, like (7) in section 2.1; another is that the context dependence of indexicals does not require a separate context index.

[22] See Sect. 2.2. In fact, a complex substitution operator consisting of a functor with a free variable for ε as its first argument could be considered an amalgam of Elbourne's R and the definite article expressed by the pronoun in his theory; in this way, the two analyses are not dissimilar as far as indexicals are concerned. It remains hard, though, to see how Elbourne's analysis could be extended to definite descriptions.

descriptive indexicals appear as clear cases: in (6), the salience of the concept of the individual who knocked on the door is enhanced by the utterance of the first sentence, which topicalizes the event of opening the door in reaction to an event of somebody knocking.

(6)   You shouldn't have opened the door! I could have been the Wolf!

It is also unusual for a covert element to be an optional operator—unusual, but not unique to $\mathcal{S}$; some versions of the covert genericity or habituality operator commonly assumed for the analysis of generic or habitual sentences are cases in point, for example, the habituality operator proposed by Reyle, Rossdeutscher, and Kamp (2007).

Note that $\mathcal{S}$ is an optional covert operator, meaning that in principle, any surface individual term will have two readings, one with it and one without it. Now the reading with it will be constrained; for one thing, it will not always be defined, and second, it will only be (taken to be) the intended reading when there is a reason. What reason that may be is addressed below.

4.2   Constraints and refinements

In connection with the definition of the $\mathcal{S}$ operator in (48), the question of what 'triggers' it, in particular, that of how a concept becomes uniquely salient in a context, was left open. In section 4.2.1, I consider more closely the ways in which a concept may or may not be determined by a context, and in section 4.2.2, I discuss how pragmatic factors influence both the choice to apply the operator, from the speaker's or the hearer's point of view, and the ensuing interpretation.

4.2.1   A cline of salience

At one end of a scale, a concept ε is so salient in a situation of utterance that it is almost as if the term actually used, expressing δ, were an anaphor and ε were its antecedent. Cases in point are (6), (7), (18)–(23), (26), (28), and (29), repeated here (recall that the context is a discussion of a snowmobile accident where a boy, driving in snowdust and straying from the trail, has hit a tree; the debate is about whether or not the tree was to blame and should be cut down):

(29)   What if <u>the tree</u> had been a Moose, a deer or another sled, would this still have happened? I believe it would have, . . .

Though there is no simple way to formulate a definite description expressing ε, its existence and uniqueness presupposition would be verified, and it is clear to both interlocutors that ε and δ ('the tree') are coreferential.

Quite often in these cases, the ε concept is event-based in the sense that it designates the agent or another participant of some given or inferable event, so that a definite description expressing it will involve a stage-level property; for instance, in the case of (29), this might be the property of obstructing the

course of the snowsled driven by so-and-so at such-and-such a time. In the case of (26), it would be the property of emitting that other moose love call.

And in many cases, the situations described are of the 'turn-out-to-be' type, where the content of a container, say, is originally unknown but turns out to be a certain individual; the Wolf and Burglar examples are of this kind, as is (20), in a literal sort of way. The author has an ultrasound and finds that her second baby will be a boy. The last time, when her first baby was on the way, she wanted it to be a boy, but it turned out to be a girl—Talea.

(20)   I had a feeling it would be a boy because I wanted it to be a girl. I wanted <u>Talea</u> to be a boy, and I got the opposite. At least I'll have one of each (boy and girl), so I'm pretty happy.

At this end of the scale, it is natural to regard the salient concept as something the context supplies.

In many cases, though, there is an amount of vagueness or indeterminacy. For example, in section 1 I simplistically assumed that the substitute for *the bear* in (2) is 'the creature setting the dog barking and knocking things over'.

(2)   The owners initially thought <u>the bear</u> was a burglar

But in fact, a range of slightly different concepts offer themselves as alternatives, *inter alia*:

— the being knocking things over
— the being setting the rottweiler barking wildly
— the being causing the commotion in the hall outside the bedroom

While it can be argued that the choice between these options does not matter, since relative to the common ground built up by the preceding sentences, if (2) is true under one substitution it is true under the others as well, the fact remains that the alternatives are distinct individual concepts, so that the 'uniquely salient' condition is strictly not defined. But there is at any rate a most salient family, though fuzzy at the edges, of closely similar concepts, among which the hearer can select one, or from which, alternatively, she can form a compound.

At the other end of the cline of salience, there are cases where the hearer is not expected to identify the substitute concept the speaker has in mind. Here, attitude contexts prevail, and (49), discussed by Zimmermann (1991), may be a case in point:

(49)   The Godfather believes that our infiltrated agent is totally reliable

Actually, whether it is a case in point depends on the context of utterance, as a concept like 'the Godfather's new chauffeur' may in one context be salient while in another not (the *de dicto* reading can be disregarded, as it would imply the attribution of a contradictory attitude). Both interlocutors may be aware that that is the way the Godfather is acquainted with their agent, so that it is

immediately clear to the hearer that the speaker is alluding to that concept.[23] But it seems that I can also use (49) to report the Godfather's attitude to you in a context where our agent's cover is known to me but not to you, and still get across that this attitude concerns a concept I have in mind. The same holds for (27) in section 2.3.2 (*the chief of the army—Odette's lover*).

But then two questions arise. First, how does it come across to you that I have another concept in mind than the one I express? In this particular case, the answer seems evident: the *de dicto* reading is absurd. In general, in the absence of 'positive' evidence, 'negative' pragmatic evidence of the kind described by Aloni (2005, pp. 520ff.) seems important. Second, what does the substitute interpretation amount to when the hearer has no clear idea of the substitute? Both questions are addressed below.

4.2.2 Pragmatic factors

In the theory of Aloni (2005) (see Sect. 3.3), a conceptual cover $\wp(n)$ is determined on the basis of a variety of factors, where pragmatic pressures play a major role. A non-literal reading always carries a certain interpretive cost, but if a literal reading would make the utterance inconsistent, uninformative, or irrelevant, this helps select a cover yielding such a reading. In a similar vein, Recanati (1993) sees descriptive readings as resulting from a procedure of pragmatic enrichment. This is echoed by Hunter (2010), whose operation of reconstruction, or coercion, is only brought on in response to a discrepancy, like an inconsistency.

Reasons to distrust the literal meaning thus provide reasons to be receptive to a substitute. Further, according to Aloni, since there are independent reasons to prefer short and rigid terms like indexicals and names to long, nonrigid ones, speakers will more often (be taken to) intend substitute readings of such terms. Conversely, a concept is more likely to be the intended substitute if it is costly to express and process. This factor, given the term *the tree* in (29), makes the meaning 'the entity obstructing the course of the snowsled driven by your son on Pitre Trail Friday' (provided the hearer is able to hypothesize it) probable.

There seem to be two lessons to be drawn from such considerations regarding the substitution operator $\mathcal{S}$ defined in (48) in section 4.1.[24]

(i) Pragmatic pressure is necessary for $\mathcal{S}$ to be generated and interpreted, so that even in a case where $[\![\,\mathcal{S}\,]\!]^{c,j}$ would be defined $\mathcal{S}$ may not apply.

---

[23]Cf. van Rooy and Zimmermann 1996, p. 131: 'a *de re* belief attribution is already appropriate if there is a unique most salient way in which the agent is acquainted with the object the belief attribution is about.' Although acquaintance on the part of an attitude holder is on the present analysis not generally a criterion for a substitute concept, if the intrasentential context is a belief sentence, mutual knowledge of an acquaintance relation between the agent and the object will be a typical way for a concept to be salient.

[24]Note that in Aloni's theory, because conceptual covers are token-sensitive and depend on the occasion on which the term is used, the pragmatic constraints are built into the perspective and thus constrain the contextually operative cover, which will often contain just the 'old' concept but sometimes a concept giving a pragmatically optimal pairing of form and content.

(ii) Pragmatic pressure may be sufficient for $\mathcal{S}$ to be generated and interpreted, even in a case where $[\![\,\mathcal{S}\,]\!]^{c,j}$ is strictly not defined.

Another way of saying (i) is that salience is not enough; it is also essential that the substitute interpretation is plausible. This general constraint must be relied on to keep the operation from over-generating. For example, it is a matter of observed fact that certain information-structural configurations are more conducive to substitute interpretations than others. It may not have escaped notice that in practically all the examples of descriptive readings given in section 2, the critical referential term is in subject position of a copular sentence; generally, it would appear that substitute interpretations are facilitated when the term in question is a so-called topic and the predicate is 'individual-level'.[25] This is more a tendency than a clear criterion, though, and although the mechanisms at work are as yet not well understood, it is natural to conclude that those are the contexts where a difference in plausibility between a literal and a non-literal reading easily comes to the fore, making the necessary pragmatic pressure felt.

Regarding (ii), there are two cases to be distinguished. First, the context may not supply a uniquely salient concept, but pragmatic reasoning may cause there to be one. Thus described, the case is no different from that of various other contextual parameters, such as 'discretionary' indexicals (Perry 2001, pp. 58f.), anaphora antecedents (Bach 2005), or determiner domain restrictions (Glanzberg 2006). The 'metasemantics' of the substitute concept ε may be rather complex, but this is something it has in common with many or most contextual parameters.

Second, if the context fails to supply a uniquely salient concept, pragmatic reasoning may bring on an indefinite substitute reading: in the case of (49), for some ε coextensional with our agent, the Godfather believes that ε is reliable. (27) in section 2.3.2 is a clear case: the hearer is not intended to realize that Swann wants to kill whoever is Odette's lover, only that he wants to kill whoever matches some description which happens to be true of the chief of the army. This may be called a case of accommodation. Again, the case is paralleled by various other contextual parameters, such as determiner domain restrictions (Glanzberg 2006), conversational backgrounds (Kratzer 1981, p. 61), or internal arguments of relational nouns (Cresswell 1996, pp. 21ff.). Accommodation is a notion that covers a variety of phenomena and approaches (see Beaver and Zeevat 2004 and von Fintel 2008); mostly, it has been studied as a strategy of adjusting the context to justify a presupposition. A slightly different use of the term is found in connection with contextual parameters (see Glanzberg 2007), to indicate the indeterminate, or underspecified, interpretation that a hearer will construct when unable to identify the value that a speaker has in mind.

Note that the resulting interpretation comes very close to what Quine (1956) called the substitutional interpretation of existential quantification and what Kaplan (1968) sought to constrain to cases where a substitute is 'vivid', to steer clear of substitutions where the content threatens to be trivialized (the 'shortest spy' problem). Here, the indefinite interpretation will be restricted, by

---

[25]I am grateful to Ora Matushansky for bringing up this point.

pragmatic considerations in the spirit of Aloni (2005), to what the speaker can reasonably be taken to have in mind, for example, if the intrasentential context is a belief sentence, a concept based on vivid acquaintance; and the concept of 'the shortest spy' will be excluded on the grounds that the speaker must be assumed to have something informative in mind.

There is also a question of how to assess cases where the salience of a certain substitute ε is evident and the speaker is confident that ε and the term she utters are coextensional, but the fact of the matter is that they are not. On the one hand, it seems that the proposition based on the substitute interpretation may still count as defined, as long as it is clear that the speaker intends this interpretation and believes that the two individual concepts are coextensional. Thus, for example, at the point in the tale of the Wolf and the seven young kids where the kids finally open the door to the Wolf, after being shown the white paw, an utterance of (50) might make sense, even to hearers who know that the creature at the door was not the mother, and be judged true.[26] This, too, might then be considered as a type of accommodation.

(50)   They did not open right away because they feared their mother might be the Wolf

On the other hand, this is reminiscent of cases, brought up by Donnellan (1966), where definite descriptions are used referentially and the speaker manages to refer to an entity even though she misdescribes it. It is widely held that such cases of successful communication are more pragmatic than semantic in nature and that the description actually fails to denote anything (or not the intended thing); Schoubye (2012) offers a recent survey of the debate. The same might be said of a substitute interpretation arising in spite of the substitute and the original term denoting two different entities in the context world: it is undefined, but a charitable hearer can still be expected to go along with it.

4.3   Loose ends

Two issues have been left unresolved. First: while the generalization of 'descriptive indexicals' to all individual terms has a firm empirical basis in modal contexts, it does not appear to be warranted in quantificational contexts. This is addressed in section 4.3.1. Second: since transparent interpretations evidently exist, it would seem that Substitution Theory can only supplement, not supplant, Transparency Theory (TT); yet it is for the most part possible to emulate transparent interpretations in Substitution Theory. Reasons why there is still a need for TT are discussed in section 4.3.2.

4.3.1   Temporal and quantificational contexts

Recall from section 2.3 that while all type e expressions—indexicals, names, or definite descriptions—can have substitute interpretations in narrowly modal

---

[26]I am indebted to an anonymous reviewer for bringing up this point.

contexts, with respect to worlds, it seems that only indexicals can have them in temporal or quantificational contexts, with respect to times or 'cases'. The substitution operation $\mathcal{S}$ is defined over situations of evaluation generally, which makes sense as far as indexicals are concerned, as indicated by (7) and (51).[27]

(7)   I used to run the Davos summit program. ...I said, in ten years' time, half of you will be women.

(51)   The nurse changes from day to day, but has, up to now, always been a woman. One day a male nurse walks in and the old patient looks surprised. The nurse says, 'I know, I'm usually a woman. But don't worry. I'll take just as good care of you as my female colleagues.'

But if substitute interpretations of names and definite descriptions are only attested in world-shifting contexts, the situations should be split into worlds and times, and the two parameters should be treated separately and differently. It is highly problematic, however, to restrict the substitution operation to individual concepts expressed by indexicals when times or 'cases', not worlds, are at issue, or conversely, to restrict its effect to the world parameter when names or definite descriptions are concerned.

Note, however, that descriptively, this restriction is not sharp at the edges. Maier (2009b) asks us to consider an utterance of (52) in a context where the identity of Wuornos has been firmly established, say a documentary about women on death row (p. 310):

(52)   Aileen Wuornos is traditionally allowed to order whatever she likes for her last meal

Here a descriptive reading of the proper name is at least marginally possible. Judging from this, it would seem that, again, information structure plays a role: it helps if the term belongs to the background, counting as given information.

On the other hand, it is not easy to find authentic examples of descriptive indexicals in unequivocally non-modal contexts. As noted in section 2.1, the adverb *traditionally* is not necessarily an adverb of quantification but might mean something like 'tradition dictates that', making it a modal expression. King (2006) notes that 'the descriptive readings are so fragile and idiosyncratic', citing the contrast between (53a) and (53b):

(53)   a.   I am traditionally allowed to order whatever I like for my last meal
            b. *I am always allowed to order whatever I like for my last meal

On the same note, Santorio (2012), referring to the phenomenon 'that indexicals can have descriptive-sounding readings when occurring under adverbs of quantification', laments that 'the productivity of the phenomenon is quite limited'.

Judging from this, there may not be a clear contrast between indexical and non-indexical terms as to the availability of descriptive readings in temporal and quantificational contexts after all. Note that a certain bias for indexical

---

[27]This example was suggested by an anonymous reviewer.

terms is predicted under Aloni's pragmatic theory (see Sect. 4.2.2). It may be that the pragmatic pressure required for substitute interpretations to arise has more difficulty building up when the context is temporal or quantificational than when it is modal. But if so, an explanation of this asymmetry must await further study.

4.3.2 Would we still need a Transparency Theory?

Would a Substitution Theory like the one sketched in this section render a Transparency Theory (like the one defended by Percus 2000, Heim 2011, or, somewhat differently, Keshet 2010 or Schueler 2011) redundant? At a superficial level, the answer is no: as noted in section 2.3, substitute and transparent readings may coexist (though maybe not in the same context), thus we would still need a Transparency Theory to derive the former. As suggested in section 3.3, however, one can emulate transparent readings on a Substitution Theory by letting the context select the rigid concept $\lambda j\, \delta_c$ (where c is the context of utterance) as the salient one; one could even use this to define a 'transparency' version of $\mathcal{S}$. So at a deeper level, the question is whether this is general enough.

There are two arguments that it is not, based on two sources of evidence. First, *de re* readings have been argued to exist not just for individual terms but for predicate terms (as well or generally); for instance, the 'non-specific *de re*' interpretation of an indefinite like the one in Fodor's (1970) example (54) is commonly (see, for example, von Fintel and Heim 2011, pp. 102ff.) taken to result from the transparent evaluation of the NP *hat just like mine*.[28] Similarly, the noun *semanticist* in (55), adapted from Percus (2000) by Heim (2011), is commonly ascribed a *de re* reading arising from actual-world evaluation.

(54)   Mary wants to buy a hat just like mine

(55)   If every semanticist owned a villa in Tuscany (instead), the world would be a boring place

It is difficult to make these cases fit into the frame of a Substitution Theory geared to individual concepts.

The second class of evidence that 'something essentially like the Index-Binding account must be available' (Heim 2011) comes from 'intermediate' readings of examples like (56):

(56)   John thinks he would arrive earlier if he wasn't in the bus he is in

The advantage of such an account in the face of such cases is that it makes it possible for a term like *the bus he is in* to be read transparently in regard to a local modal context (the counterfactual), but yet not totally *de re*, not in regard to the next modal context (the attitude); here it is read *de dicto*.[29] Since such readings are not enabled by Substitution Theory, nor by a theory

---

[28] But see Schwager 2011 for counterarguments to this view.

[29] See also the examples given in Elbourne 2013, Ch. 7.

of counterpart functions, they strengthen the argument that the flexibility of Transparency Theory (in the form of free index-binding) is irreplaceable.

5   Conclusions

There are basically two ways to respond to cases like the definite description in the thought report in the last, key clause of (1), repeated here as (57):

(57)   The owners initially thought <u>the bear</u> was a burglar

One way is to ascribe to the term a **transparent** interpretation on which in the thought worlds of the owners of the ransacked house, the individual denoted by *the bear* in the actual world was a burglar, as sketched in (58):

(58)   the owners thought$_v$($\{\, w \mid$ the bear$_v$ is a burglar$_w \,\}$)

While this response may just be tenable in this particular case, intuitively there is a close parallel to the interpretation of a pronoun like *you* in cases like (59) (discussed by Bezuidenhout (1997, p. 384) and cited in Sect. 2.1):

(59)   I thought you were a burglar

Here, the normal—transparent—reading is widely held to be inadequate. There is broad recognition that the actual interpretation coincides with the *de dicto* reading of a substituted description:

(60)   I thought the person who crashed through the door was a burglar

Elbourne (2008) offers a formal analysis of descriptive indexicals predicting, effectively, the step from (59) to (60). Hunter (2010) shows, however, that descriptive interpretations extend to proper names and definite descriptions; transparent interpretation is as inappropriate for (61) as for (59).

(61)   She thought her husband was a burglar

This is the cue to the other response to cases like (57) above: the term has a **substitute** interpretation on which in the thought worlds of the owners, the creature setting the dog barking and knocking things over was a burglar:

(62)   the owners thought$_v$($\{\, w \mid$ [the creature setting the dog barking and knocking things over]$_w$ is a burglar$_w \,\}$)

So one key conclusion is that substitute interpretations exist and cut across both the different types of individual terms (pronouns, names, definite descriptions) and the different classes of modal contexts (propositional attitudes, modal auxiliaries, counterfactuals).

In particular, they are not confined to attitudinal contexts, where they have been studied under the label of *relational de re*, and they are not limited to rigid designators, as previous work on descriptive indexicals would suggest. In

fact, the theory advocated above can be seen as integrating relational *de re* interpretations and descriptive interpretations into a common notion.

This notion is not meant to replace the notion of transparent interpretation. Although substitute interpretation can emulate direct referentiality as far as individual terms are concerned, individual terms are not always concerned, nor is direct referentiality always at issue. But it is meant to take the load off transparent interpretations in those cases where they are implausible.

Substitute interpretation means extra machinery: Substitution Theory; and a general and compositional formulation is desirable. Most existing theories are either not general regarding both the range of terms and the range of contexts or not compositional. I have proposed the introduction of an optional covert operator over individual terms, $\mathcal{S}$, expressing an operation on individual concepts. The exact implementation of this approach may be questioned, but still, a substitution operation essentially like the meaning of $\mathcal{S}$ is the logical response to the array of descriptive readings of individual terms in modal contexts.

Compositionality carries the price that restrictivity can only be attained by non-structural means, contextually and pragmatically: if the freely available (though optional) substitution operation were left to itself, it would do much unnecessary, and potentially harmful, work. Context and pragmatic factors are therefore assigned a key role in restricting the operation; for instance, in non-modal contexts the necessary pragmatic pressure will mostly be absent. Such a restrictive role for context and pragmatic factors is plausible, and it is anchored in the work of Aloni (2005), Maier (2009a), and Hunter (2010).

It is still arguable that substitute readings of terms should not be described in semantic terms but rather be viewed 'as pragmatic phenomena that arise in order to repair . . . problematic predications' (Hunter 2010, p. 148). I have explored the ways to meet the criteria of generality and compositionality in regard to a semantic treatment of these readings, and if a conclusion as to whether a general and compositional semantic treatment is desirable cannot yet be drawn, the present paper has at least gone some way towards clarifying the premises and preparing the ground for drawing such a conclusion.[30]

References

Almog, Joseph and Paolo Leonardi (eds) 2009: *The Philosophy of David Kaplan*. Oxford: Oxford University Press.

Almog, Joseph, John Perry, and Howard Wettstein (eds) 1989: *Themes from Kaplan*. Oxford: Oxford University Press.

Aloni, Maria 2005: 'A Formal Treatment of the Pragmatics of Questions and Attitudes'. *Linguistics and Philosophy*, 28, pp. 505–39.

Bach, Kent 2005: 'Context Ex Machina'. In Szabó 2005, pp. 15–44.

Bäuerle, Rainer 1983: 'Pragmatisch-semantische Aspekte der NP-Interpretation'. In Faust et al. 1983, pp. 121–31.

Barwise, Jon and John Perry 1983: *Situations and Attitudes*. Cambridge, MA: MIT Press.

Beaver, David and Henk Zeevat 2004: 'Accommodation'. In Ramchand and Reiss 2004, pp. 503–39.

Bezuidenhout, Anne 1997: 'Pragmatics, Semantic Underdetermination and the Referential/Attributive Distinction'. *Mind*, 106, pp. 375–409.

Bezuidenhout, Anne and Marga Reimer (eds) 2004: *Descriptions and Beyond*. Oxford: Oxford University Press.

Bonomi, Andrea 1995: 'Transparency and Specificity in Intensional Contexts'. In Leonardi and Santambrogio 1995, pp. 164–85.

Büring, Daniel 1998: 'Identity, Modality, and the Candidate Behind the Wall'. In Strolovich and Lawson 1998, pp. 36–54.

—— 2011: 'Pronouns'. In Maienborn, Portner, and von Heusinger 2011, pp. 971–96.

Carlson, Greg and Jeff Pelletier (eds) 2005: *Reference and Quantification: The Partee Effect*. Stanford: CSLI Press.

Cole, Peter (ed) 1978: *Syntax and Semantics, vol. 9: Pragmatics*. New York: Academic Press.

Cormany, Ed, Satoshi Ito, and David Lutz (eds) 2011: *Proceedings of SALT 19*. Ithaca, New York: CLC Publications.

Cresswell, Max 1985: *Structured Meanings*. Cambridge, MA: MIT Press.

—— 1996: *Semantic Indexicality*. Dordrecht: Kluwer.

Cresswell, Max and Arnim von Stechow 1982: 'De Re Belief Generalized'. *Linguistics and Philosophy*, 5, pp. 503–35.

Cumming, Samuel 2007: *Proper Nouns*. Rutgers University Dissertation.

Davidson, Donald and Gilbert Harman (eds) 1972: *Semantics of Natural Language*. Dordrecht: Reidel.

Davis, John Whitney, Donald James Hockney, and Warren Kent Wilson (eds) 1969: *Philosophical Logic*. Dordrecht: Reidel.

Donnellan, Keith 1966: 'Reference and Definite Descriptions'. *Philosophical Review*, 77, pp. 281–304.

Egan, Andy and Brian Weatherson (eds) 2011: *Epistemic Modality*. Oxford: Oxford University Press.

Eikmeyer, Hans-Jürgen and Hannes Rieser (eds) 1981: *Words, Worlds, and Contexts*. Berlin: de Gruyter.

Elbourne, Paul 2001: 'E-Type Anaphora as NP-Deletion'. *Natural Language Semantics*, 9, pp. 241–88.

—— 2008: 'Demonstratives as Individual Concepts'. *Linguistics and Philosophy*, 31, pp. 409–66.

—— 2010: 'Why Propositions might be Sets of Truth-supporting Circumstances'. *Journal of Philosophical Logic*, 39, pp. 101–11.

—— 2013: *Definite Descriptions*. Oxford: Oxford University Press.

Evans, Gareth 1977: 'Pronouns, Quantifiers, and Relative Clauses (I)'. *Canadian Journal of Philosophy*, 7, 467–536.

—— 1982: *The Varieties of Reference*. Oxford: Oxford University Press.

Faust, Manfred et al. (eds) 1983: *Allgemeine Sprachwissenschaft, Sprachtypologie und Textlinguistik*. Tübingen: Gunter Narr.

von Fintel, Kai 2008: 'What is Presupposition Accommodation, Again?' *Philosophical Perspectives*, 22, pp. 137–70.

von Fintel, Kai and Anthony Gillies 2011: 'Might Made Right'. In Egan and Weatherson 2011, pp. 108–30.

von Fintel, Kai and Irene Heim 2011: *Intensional Semantics*. Manuscript, MIT (http://web.mit.edu/fintel/fintel-heim-intensional.pdf).

Fodor, Janet 1970: *The Linguistic Description of Opaque Contexts*, PhD thesis, MIT.

Glanzberg, Michael 2006: 'Context and Unrestricted Quantification'. In Rayo and Uzquiano 2006, pp. 45–74.

—— 2007: 'Context, Content, and Relativism'. *Philosophical Studies*, 136, pp. 1–29.

Grimm, Robert and Daniel Merrill (eds) 1988: *Contents of Thought*. Tucson: University of Arizona Press.

Heim, Irene 2004: 'Notes on Indexicality'. Manuscript, MIT.

—— 2011: 'Definiteness and Indefiniteness'. In Maienborn, Portner, and von Heusinger 2011, pp. 996–1025.

Heim, Irene and Angelika Kratzer 1998: *Semantics in Generative Grammar*. Oxford: Blackwell.

Heusinger, Klaus von and Urs Egli (eds) 1996: *Proceedings of the Konstanz Workshop Reference and Anaphorical Relations* (= Arbeitspapier Nr. 79 der Fachgruppe Sprachwissenschaft). University of Constance.

Hintikka, Jaakko 1969: 'Semantics for Propositional Attitudes'. In Davis, Hockney, and Wilson 1969, pp. 21–45.

Hintikka, Jaakko, Julius Moravcsik, and Patrick Suppes (eds) 1973: *Approaches to Natural Language*. Dordrecht: Reidel.

Hunter, Julie 2010: *Presuppositional Indexicals*. University of Texas at Austin Dissertation.

Kaplan, David 1968: 'Quantifying In'. *Synthese*, 19, pp. 178–214.

—— 1977: 'Demonstratives: An Essay on the Semantics, Logic, Metaphysics, and Epistemology of Demonstratives and other Indexicals'. In Almog, Perry, and Wettstein 1989, pp. 481–563.

Keshet, Ezra 2010: 'Split Intensionality: A New Scope Theory of De re and De dicto'. *Linguistics and Philosophy*, 33, pp. 251–83.

King, Jeffrey 2006: 'Singular Terms, Reference and Methodology in Semantics'. *Philosophical Issues*, 16, pp. 141–61.

Kratzer, Angelika 1981: 'The Notional Category of Modality'. In Eikmeyer and Rieser 1981, pp. 38–74.

Kripke, Saul 1972: 'Naming and Necessity'. In Davidson and Harman 1972, pp. 253–355.

Landman, Fred and Frank Veltman (eds) 1984: *Varieties of Formal Semantics. Proceedings of the 4th Amsterdam Colloquium*. Dordrecht: Foris Publications.

Leonardi, Paolo and Marco Santambrogio (eds) 1995: *On Quine. New Essays*. Cambridge, MA: Cambridge University Press.

Lewis, David 1968: 'Counterpart Theory and Quantified Modal Logic'. *Journal of Philosophy*, 65, pp. 113–26.

—— 1973: *Counterfactuals*. Oxford: Blackwell.

—— 1979: 'Attitudes De Dicto and De Se'. *The Philosophical Review*, 88, pp. 513–43.

Maienborn, Claudia, Paul Portner, and Klaus von Heusinger (eds) 2011: *Semantics: An International Handbook of Natural Language Meaning. Vol. 2*. Berlin: de Gruyter Mouton.

Maienborn, Claudia, Paul Portner, and Klaus von Heusinger (eds) 2012: *Semantics: An International Handbook of Natural Language Meaning. Vol. 3*. Berlin: de Gruyter Mouton.

Maier, Emar 2009a: 'Presupposing Acquaintance: A Unified Semantics for De Dicto, De Re and De Se Belief Reports'. *Linguistics and Philosophy*, 32, pp. 429–74.

—— 2009b: 'Proper names and indexicals trigger rigid presuppositions'. *Journal of Semantics*, 26, pp. 253–315.

Montague, Richard 1973: 'The Proper Treatment of Quantification in Ordinary English'. In Hintikka, Moravcsik, and Suppes 1973, pp. 221–42.

Neale, Stephen 1990: 'Descriptive Pronouns and Donkey Anaphora'. *Journal of Philosophy*, 87, pp. 113–50.

—— 2005: 'A Century Later'. *Mind*, 114, pp. 809–871.

Nunberg, Geoffrey 1993: 'Indexicality and Deixis'. *Linguistics and Philosophy*, 16, pp. 1–43.

—— 2004: 'Descriptive Indexicals and Indexical Descriptions'. In Bezuidenhout and Reimer 2004, pp. 261–79.

Percus, Orin 2000: 'Constraints on Some Other Variables in Syntax'. *Natural Language Semantics*, 8, pp. 173–229.

Percus, Orin and Uli Sauerland 2003: 'On the LFs of attitude reports'. In Weisgerber 2003, pp. 228–42.

Perry, John 2001: *Reference and Reflexivity*. Stanford: CSLI Publications.

Quine, Willard van Orman 1956: 'Quantifiers and Propositional Attitudes'. *Journal of Philosophy*, 53, pp. 177–87.

Ramchand, Gillian and Charles Reiss (eds) 2004: *Oxford Handbook of Linguistic Interfaces*. Oxford: Oxford University Press.

Rayo, Agustín and Gabriel Uzquiano (eds) 2006: *Absolute Generality*. Oxford: Oxford University Press.

Recanati, François 1993: *Direct Reference: From Language to Thought*. Oxford: Blackwell.

—— 2005: 'Deixis and Anaphora'. In Szabó 2005, pp. 286–316.

Rescher, Nicholas (ed) 1968: *Studies in Logical Theory*. Oxford: Blackwell.

Reyle, Uwe, Antje Rossdeutscher, and Hans Kamp 2007: 'Ups and Downs in the Theory of Temporal Reference'. *Linguistics and Philosophy*, 30, pp. 565–635.

Riester, Arnd and Torgrim Solstad (eds) 2009: *Proceedings of Sinn und Bedeutung 13*. University of Stuttgart.

Romoli, Jacopo and Yasutada Sudo 2009: '*De Re / De Dicto* Ambiguity and Presupposition Projection'. In Riester and Solstad 2009, pp. 425–38.

Rooij, Rob van 2006: *Attitudes and Changing Contexts*. Dordrecht: Springer.

Rooy, Rob van and Ede Zimmermann 1996: 'An Externalist Account of Intentional Identity'. In von Heusinger and Egli 1996, pp. 123–36.

Russell, Bertrand 1905: 'On Denoting'. *Mind*, 14, pp. 479–93.

Salmon, Nathan 2009: 'Quantifying into the Unquantifiable: The Life and Work of David Kaplan'. In Almog and Leonardi 2009, pp. 25–44.

Santorio, Paolo 2012: 'Reference and Monstrosity'. *The Philosophical Review*, 121, pp. 359–406.

Schiffer, Stephen 1992: 'Belief Ascription'. *Journal of Philosophy*, 89, pp. 499–521.

Schoubye, Anders 2012: 'Against the Argument from Convention'. *Linguistics and Philosophy*, 35, pp. 515–32.

Schueler, David 2011: 'World Variable Binding and Beta-Binding'. *Journal of Semantics*, 28, pp. 241–66.

Schwager, Magdalena 2011: 'Speaking of Qualities'. In Cormany, Ito, and Lutz 2011, pp. 395–412.

Soames, Scott 1987: 'Direct Reference, Propositional Attitudes, and Semantic Content'. *Philosophical Topics*, 15, pp. 44–87.

Stalnaker, Robert 1968: 'A Theory of Conditionals'. In Rescher 1968, pp. 98–112.

—— 1978: 'Assertion'. In Cole 1978, pp. 315–32.

—— 1981: 'Indexical Belief'. *Synthese*, 49, pp. 129–51.

—— 1988: 'Belief Attribution and Context'. In Grimm and Merrill 1988, pp. 140–56.

Stechow, Arnim von 1984: 'Structured Propositions and Essential Indexicals'. In Landman and Veltman 1984, pp. 385–404.

Stechow, Arnim von and Ede Zimmermann 2005: 'A Problem for a Compositional Account of De Re Attitudes'. In Carlson and Pelletier 2005, pp. 207–28.

Strolovich, Devon and Aaron Lawson (eds) 1998: *Proceedings of SALT 8*. Ithaca, NY: Cornell University Linguistic Publications.

Szabó, Zoltán (ed.) 2005: *Semantics versus Pragmatics*. Oxford: Clarendon.

Weisgerber, Matthias (ed) 2003: *Proceedings of Sinn und Bedeutung 7*. University of Constance.

Wunderlich, Dieter and Arnim von Stechow (eds) 1991: *Semantics: An International Handbook of Contemporary Research*. Berlin: de Gruyter.

Zimmermann, Ede 1991: 'Kontextabhängigkeit'. In Wunderlich and von Stechow 1991, pp. 156–229.

—— 2012: 'Context Dependence'. In Maienborn, Portner, and von Heusinger 2012, pp. 2360–407.

Appendix: building a substitute interpretation

Here, I show by an example how the substitution operator defined in section 4.1 can serve to modify the meaning of a sentence. Let the example be (6):

(6)   I could have been the Wolf!

I use a standard scheme of semantic composition, following Heim and Kratzer (1998), with functional application where arguments and values are either both extensions or the former are intensions when need arises (so-called intensional functional application). In the two-dimensional framework used here, these two composition principles can be formulated as follows (a, b are any logical types; s is the type of contexts of evaluation):

(63)   **Functional application (FA)**

$$ [\![\, m_{\langle a,b \rangle}\ n_a\ ]\!]^{c,j} = [\![\, m\,]\!]^{c,j}([\![\, n\,]\!]^{c,j}) $$

(64)   **Intensional functional application (IFA)**

$$ [\![\, m_{\langle \langle s,a \rangle, b \rangle}\ n_a\ ]\!]^{c,j} = [\![\, m\,]\!]^{c,j}([\![\, n\,]\!]^c) = [\![\, m\,]\!]^{c,j}(\lambda i\, [\![\, n\,]\!]^{c,i}) $$

When two expressions m and n merge and m's extension is a function from the kind of entity n has as its extension, the extension of the merge is simply the value of the function at n's extension; but when m's extension is a function from functions from contexts (or, worlds) to the kind of entity n has as its extension, the extension of the merge is the value of the function at n's intension.

Now in fact, since to allow the definedness condition on the substitution operation to be inherited to function values, function arguments must in practice be intensions throughout, IFA will be the only principle in use (corresponding, *mutatis mutandis*, to the method used in Montague 1973).

Consider this simplified logical form for (6):

(65)   [*could* [[$\mathcal{S}$ *I*] [*be the Wolf*]]]

When we are to model the interpretation of this using $\mathcal{S}$ as defined in (48), repeated below, our starting point will be the intension of the pronoun I, given in (66), because $\mathcal{S}$ looks for an individual concept, an individual in intension.

(48)   **The substitution operator $\mathcal{S}$**

$$\llbracket\,\mathcal{S}\,\rrbracket^{c,j} = \lambda\delta_{\langle s,e\rangle} \begin{cases} \varepsilon_j & \text{if there is a uniquely salient } \varepsilon \text{ in } c \\ & \text{such that } \varepsilon_c = \delta_c \\ \text{undefined} & \text{otherwise} \end{cases}$$

(66) $\llbracket\,I\,\rrbracket^c = \lambda j$ the speaker in c

This is a constant function from contexts of evaluation j to individuals. $\llbracket\,\mathcal{S}\,\rrbracket^{c,j}$ is defined for it iff there is a uniquely salient individual concept ε in c such that $\varepsilon_c = (\lambda j \text{ the speaker in } c)_c =$ the speaker in c. This definedness condition will be inherited, or projected, as a presupposition. The first step of composition takes us, by IFA, to an individual if the presupposition is met:

(67) $\llbracket\,\mathcal{S}\,I\,\rrbracket^{c,j} =$

$$\llbracket\,\mathcal{S}\,\rrbracket^{c,j}(\llbracket\,I\,\rrbracket^c) = \begin{cases} \varepsilon_j & \text{if there is a uniquely salient } \varepsilon \text{ in } c \\ & \text{such that } \varepsilon_c = \text{the speaker in } c \\ \text{undefined} & \text{otherwise} \end{cases}$$

The extension of the predicate [*be the Wolf*], a partial function from individual concepts to truth values, is given in (68), where $\mathcal{W}$ is the individual constantly denoted by *the Wolf* as a proper name:

(68) $\llbracket\,\textit{be the Wolf}\,\rrbracket^{c,j} = \lambda\zeta_{\langle s,e\rangle} \begin{cases} \text{defined} & \text{iff } \zeta_j \text{ is defined} \\ 1 & \text{if defined and } \zeta_j = \mathcal{W} \\ 0 & \text{if defined and } \zeta_j \neq \mathcal{W} \end{cases}$

In the second step of composition, where the predicate is applied to the subject, we use IFA again, instantiating ζ to $\lambda i\,\llbracket\,\mathcal{S}\,I\,\rrbracket^{c,i}$ (which is either defined for all j or undefined for all j, as the definedness condition only concerns the context of utterance c):

(69) $\llbracket\,[\,\mathcal{S}\,I\,][\,\textit{be the Wolf}\,]\,\rrbracket^{c,j} = \llbracket\,\textit{be the Wolf}\,\rrbracket^{c,j}(\lambda i\,\llbracket\,\mathcal{S}\,I\,\rrbracket^{c,i}) =$

$$\begin{cases} \text{defined} & \text{iff there is a uniquely salient } \varepsilon \text{ in } c \text{ such that} \\ & \varepsilon_c = \text{the speaker in } c \\ 1 & \text{if defined and } \varepsilon_j = \mathcal{W} \\ 0 & \text{if defined and } \varepsilon_j \neq \mathcal{W} \end{cases}$$

The counterfactual possibility modal *could* is defined in (70). (To be sure, this is a simplistic definition; the primitive relation $\simeq$ between two contexts j and j' is to indicate that the world of the latter is accessible from that of the former. For a more elaborate and realistic semantics for an epistemic modal like this, see von Fintel and Gillies 2011.)

(70) $\llbracket\,\textit{could}\,\rrbracket^{c,j} = \lambda\phi_{\langle s,t\rangle} \begin{cases} \text{defined} & \text{iff } \phi_j \text{ is defined} \\ 1 & \text{if defined and } \exists k \simeq j : \phi_k = 1 \\ 0 & \text{if defined and } \neg\exists k \simeq j : \phi_k = 1 \end{cases}$

This function from (partial) propositions applies to the intension of the clause $[[\; \mathcal{S}\; I\;]\; [\; \textit{be the Wolf}\;]]$ to yield a truth value or no value:

(71) $[\![\; \textit{could}\; [[\mathcal{S}\; I\,][\,\textit{be the Wolf}\,]]\;]\!]^{c,j} =$ (by IFA)

$[\![\; \textit{could}\;]\!]^{c,j}\,(\lambda i\; [\![\; [\mathcal{S}\; I\,][\,\textit{be the Wolf}\,]\;]\!]^{c,i}) =$ (by (70))

$$\begin{cases} \text{defined} & \text{iff there is a uniquely salient } \varepsilon \text{ in } c \text{ such that} \\ & \quad \varepsilon_c = \text{the speaker in } c \\ 1 & \text{if defined and } \exists k \simeq j : \varepsilon_k = \mathcal{W} \\ 0 & \text{if defined and } \neg\exists k \simeq j : \varepsilon_k = \mathcal{W} \end{cases}$$

Now assume that there is indeed a uniquely salient individual concept ε in c such that $\varepsilon_c = $ the speaker in c, and that it is 'the individual who knocked on the door (at the given time in the immediate past)'. Letting $\mathcal{K}$ symbolize this, the truth conditions then reduce to (72):

(72) $[\![\; \textit{could}\; [[\mathcal{S}\; I\,][\,\textit{be the Wolf}\,]]\;]\!]^{c,j} = 1$ if $\exists k \simeq j : \mathcal{K}_k = \mathcal{W}$; else false

Thus in result, (6) is true with respect to c and j if there is a j-accessible world where (there is exactly one individual who knocked on the door and) the individual who knocked on the door is the Wolf, otherwise it is false. In other words, we get the truth conditions we would get if the sentence were *the individual who knocked on the door could have been the Wolf* instead, on a *de dicto* reading.

To be exact, the truth condition deriving from (70) ($\exists k \simeq j : \phi_k = 1$) is that for some j-accessible world k, (i) $\mathcal{K}_c = $ the speaker in c and (ii) $\mathcal{K}_k = \mathcal{W}$; (i), however, is redundant, since it coincides with the definedness condition. But note that if this condition were not anchored to the unshiftable context but to the context of evaluation, as it would have to be in a one-dimensional semantics, it would not be redundant but would cause an unwelcome prediction. In fact, indexicals and names would be predicted to have a rigid interpretation after all, and definite descriptions would be predicted to have a *de dicto* interpretation, if the coreference condition were anchored to the shiftable context of evaluation. Thus a two-dimensional semantics, or something essentially like it, anchoring the coextensionality condition to the 'actual world', seems indispensable for a substitution theory of *de re* interpretation.