

# Transparent Evaluation and Pronouns in Attitude Reports

Daniel Tiskin

Department of Logic, Saint Petersburg State University  
5 Mendeleevskaya line, 199034 St. Petersburg, Russia

Department of the Humanities, NRU Higher School of Economics—St. Petersburg  
16 Soyuza Pechatnikov st., 190008 St. Petersburg, Russia  
daniel.tiskin@gmail.com

**Abstract:** The paper outlines an account for Simon Charlow’s data concerning *de se* and *de re* readings of pronouns and anaphors in attitude contexts. Using Arnim von Stechow’s binding technique as well as the insights about the internal structure of pronouns (due to Rose-Marie Déchaine and Martina Wiltschko) and about transparent readings of predicates (due to Yasutada Sudo) I treat *de se* readings as primitive and *de re* ones as derived. An additional assumption about the semantics of reflexives is used to explain why, as shown by Charlow, a *de se* anaphor cannot be bound by a *de re* subject.

Next, I show another direction within the problem of anaphora one might proceed in with the treatment of pronouns found in Déchaine and Wiltschko’s paper.

Finally, comparing my proposal with its predecessors, I touch upon the issue of the extent to which a semantic theory should be philosophically laden.

**Keywords:** propositional attitude reports, anaphora, *de se*, *de re*, transparent readings

## 1. Introduction

The problem of propositional attitudes and attitude reports *de se* has a fairly long history, (Castañeda, 1966; Lewis, 1979) being common early references in philosophy. Semantically put, the gist of the problem is this: do sentences like (1) have a single reading or two different readings, depending on whether John realises that it is himself who is doing so well in his election campaign (*de se* attitude) or he fails to realise this (perhaps due to being drunk; *de re* attitude), and if there are distinct readings, what should the semantics for the *de se* one look like?

John<sub>*i*</sub> believes that he<sub>*i*</sub> will win the election. (1)

Contrary to e.g. Emar Maier (Maier, 2009), Simon Charlow (Charlow, 2010) has recently argued that there can be no uniform treatment. He invokes an older argument from (Percus & Sauerland, 2003, 234) centering around sentences like (2):

Only John<sub>*i*</sub> believes that he<sub>*i*</sub> will win the election. (2)

Imagine a scenario where John, Tom and Mary are all candidates watching themselves on TV speaking during their election campaign. Each of them is impressed by his/her own speech, so each believes s/he will be elected; however, only John is sober enough to recognise himself. In such a setting (2) has a true reading whereas its *de re* construal is clearly false, for all three candidates (not *only* John) hold relevant *de re* beliefs. Therefore, as Charlow puts it, “we must be able to generate [special] *de se* truth conditions.” His own way of doing this is to restrict the domain of quantification over *concept generators*<sup>1</sup> (functions from individuals, i.e. attitude holders, to individual concepts, i.e. to individuals-in-intension towards which the attitude is aimed) to those generators that, for a given attitude holder, return himself. As the reader may easily note, this is all fairly complex as it stands.

---

<sup>1</sup>The reason for introducing concept generators is that they formalise the notion of an acquaintance relation; for the original presentation of the latter, see (Lewis, 1979).

Now, Charlow poses a new challenge for any theory of *de se* ascriptions. Consider an attitude report containing a reflexive in the object position coreferential with the subject as in (3).

John<sub>i</sub> fears that he<sub>i</sub> will hurt himself<sub>i</sub>. (3)

As Charlow shows, only three of the four logically possible combinations of *de re* and *de se* construals are attested in such cases:

1. *De se—de se*: e.g. if John, occupied with some dangerous work, thinks with fear, “I will hurt myself.”
2. *De se—de re*: e.g. if John is about to drop a heavy hammer on his own feet without realising they are his and thinks, “I will hurt that guy.”
3. *De re—de re*: e.g. if John is watching a video of himself dropping a hammer on his feet, again without realising his own starring role in the video, and thinks, “Lo! This guy will hurt himself.”

The remaining pattern, *de re—de se*, is unattested: (3) cannot be uttered felicitously if John’s thought is, “This guy [actually John himself] will hurt me”. This is problematic since what generally matters for anaphora is mere extensional coindexing, and in this respect *de se* is not different from *de re*.

Charlow’s own solution for this problem again makes use of concept generators, resulting in tricky logical forms. Is there a way to account for his puzzle at lesser expense? In what follows, I will sketch a possible alternative and provide an excursus into another problem that might once be solved with the help of the proposed technique.

Here is the outline of the paper. Section 2 presents the ingredients of the proposal. In Section 3, I give more justification for using transparent evaluation techniques in the study of pronouns in attitude reports. Section 4 concludes with some remarks on the scope of a semantic theory based on the comparison between the proposed account and its predecessors.

## 2. The Apparatus

### 2.1. Pronouns Decomposed

Rose-Marie Déchaine and Martina Wiltschko (Déchaine & Wiltschko, 2002) proposed a tripartite taxonomy of pronouns across languages as well as within a given language. There is no need to invoke the whole theory here; suffice it to say that they treat such third-person pronouns as *he* and *she* as  $\phi$ Ps, that is, *he* is a complex syntactic constituent of the form  $[\phi_P \text{ he } [\text{NP } \emptyset]]$ . The title “ $\phi$ P” comes from the number, gender and person features that pronouns bear. This accounts for their usability as predicates as well as arguments. Now, I suggest that pronominal  $\phi$ Ps take the null determiner<sup>2</sup> (cf. (Heim & Kratzer, 1998, 290 ff.)), resulting in the overall DP structure  $[\text{DP } \emptyset [\phi_P \text{ he } [\text{NP } \emptyset]]]$ . While in Déchaine and Wiltschko’s original formulation such a pronoun is semantically merely an individual variable, I propose to deviate from their analysis at this point and treat pronominal DPs as definite descriptions containing a free variable, of the form  $\iota x.x \simeq y$ .

<sup>2</sup>Here is another reason to hold such a view: there are so-called “paycheck sentences” (Karttunen, 1969) where the pronoun borrows its descriptive content from its antecedent DP but definitely denotes a different object.

The man who gave his paycheck to his wife is wiser than the man who gave **it** to his mistress. (\*)

Assuming that the pronoun itself is not a complete DP but rather a  $\phi$ P, we can allow for there being a separate quantifier corresponding to the null determiner that precedes the pronoun.

## 2.2. Transparent Evaluation

A crucial element of the present proposal is the mechanism of transparent evaluation for predicates and quantifier restrictors. In addition to the well-known *de re/de dicto* ambiguity, yet another reading is attested (Fodor, 1979), initially called “non-specific transparent”; thus (4) can be true if John wants to buy a (non-specific) Mercedes-Benz car and still has no idea as to whether it will cost him much whereas the speaker knows that all too well.

John wants to buy an expensive car. (4)

As long as the scopes of the quantifier corresponding to the indefinite article and of its restrictor fall apart (by stipulation, the car is non-specific, but the predicate *expensive* is not part of John’s attitude), a fair deal of ingenuity is needed to provide a satisfactory analysis for (4). There have been several attempts at its formulation, but their evaluation exceeds the limits of the present paper. I only mention a recent insight due to Yasutada Sudo (Sudo, 2013) that generalises to main predicates as well. In his semantics Sudo allows for free substitution of an expression with another one “contextually equivalent” to it in the course of semantic evaluation. Contextual equivalence between  $\alpha$  and  $\beta$  in the set of worlds  $C$  is here understood along the lines of (Schwager, 2009) as equivalence, for all  $w \in C$ , in all  $w'$  (with  $\alpha$  and  $\beta$  non-empty) maximally similar to  $w$ .

Granted this (and actually weakening the equivalence requirement up to entailment in all maximally similar possible worlds: if at least one such substitution comes out true, the initial report does as well), we can easily substitute *expensive car* for *Mercedes-Benz car*—a predicate that entails *expensive car* in all worlds compliant with our knowledge. This very device will be used below to account for *de re* readings of pronouns in subject position, as opposed to their corresponding *de se* readings.

## 2.3. Reflexives

As long as the puzzling examples from (Charlow, 2010) involve reflexives, we need a theory thereof. The choice I have made will be justified below by the fact that it fares well on the data in question, so here I just give the idea. This idea we find in (Lubowicz, 1999). As Lubowicz claims, the (genderless) analogue of *himself/herself* in Polish represents a *detransitivising* operator on predicates. An analogous clause for *himself* (simplified compared to Lubowicz’s full account) would look like

$$\| \textit{himself} \|_{\langle \langle e, et \rangle, et \rangle} = \lambda R \lambda x. R(x, x) \quad (5)$$

Therefore, the resulting predicate will effectively demand only one argument value—namely, the value of its subject. Note that on this approach, there is *no pronoun to bind* and no DP to separately interpret in the object position separately from binding the subject; my approach rests crucially on this feature.

As it stands, this definition might cause trouble in the case of three- (and four-, etc.) place predicates, but in the present paper we will simply restrict ourselves to two-place verbs.

## 2.4. Binding

As for the binding mechanism for the subject pronoun in the *de se* case, I make use of Arnim von Stechow’s proposal (von Stechow, 2002). He assumes the device of “centred worlds”: those are world-individual pairs, the individual intuitively conceived of as the “first person,” or the “I,” from whose viewpoint the world is given. The accessibility relation is likewise defined on pairs. In von Stechow’s semantics, variables corresponding to pronouns in an embedded clause may

be bound by the attitudinal verb c-commanding the clause. That is, for him attitudinal verbs are (ignoring time) universal quantifiers over individual-world pairs:

$$\|believes\| = \lambda P_{\langle e, st \rangle} . \lambda x_e . \lambda w_s . \forall \langle x', w' \rangle \in \text{DOX}(x, w) : P(x', w') \Leftrightarrow \top \quad (6)$$

So if there is a pronoun within the embedded clause coindexed with the lambda abstractor over the clause,<sup>3</sup> this pronoun will be interpreted as coreferential with the attitudinal verb's subject, yielding the *de se* interpretation.

## 2.5. The Treatment

As indicated in 2.1, I take pronominal DPs as definite descriptions containing a free variable. So the embedded clause of (1) will look like

$$\dots \lambda y . [\text{the } x . x \simeq y] x \text{ will win the election,} \quad (7)$$

and the embedded clause of (3)—with *himself* already applied to *will hurt*—like

$$\dots \lambda y . [\text{the } x . x \simeq y] x \text{ will hurt } x. \quad (8)$$

Now in (8) there are two predicates. One, *will hurt*  $x$ ,<sup>4</sup> can be interpreted, according to Sudo's rule, as any predicate equivalent to it. Obviously, this equivalence relation has now to be made assignment-relative, for *will hurt*  $x$  contains a free variable at this stage of semantic composition. This done, *will hurt*  $x$  may be replaced in the course of interpretation by, say, *will hurt the guy whose feet John sees*—the predicate contextually equivalent to *will hurt*  $x$  with  $x \mapsto \mathbf{John}$ . This replacement gives us the desired *de se*—*de re* pattern.

What about the remaining one, *de re*—*de re*? To get it, what has to be interpreted transparent is no longer the main predicate *will hurt*  $x$  but rather the subject restrictor  $[\simeq y](x)$ . On any assignment mapping  $y$  to John,  $\{x \mid x \text{ John sees } x \text{ in the video}\} = \{x \mid x \simeq y\}$  in the *de re*—*de re* scenario listed above. Of course if this restrictor predicate is interpreted transparent, *himself* has no chance of getting the *de se* interpretation: as noted above, under the adopted semantics for *himself*, there is just no DP that would get its interpretation independently of the subject DP. So the unattested reading is avoided.

## 3. Another Case for Transparent Evaluation

In this section I will bring to the reader's attention another case in the semantics of attitude reports where transparent evaluation may come in handy. Here I go only half-way and do not present a ready-made solution, but rather show that the line of thought that has been long taken as doomed to failure, is in fact not hopeless.

Peter Geach is credited for bringing into prominence a problem he labelled as *Intentional Identity*. In (Geach, 1967), the following example is discussed:

$$\begin{aligned} &\text{Hob thinks a witch has blighted Bob's mare,} \\ &\text{and Nob wonders whether she (the same witch) killed Cob's sow.} \end{aligned} \quad (9)$$

What is problematic for the traditional scopal analysis is that, somewhat like in Fodor's case above, neither *de re* nor *de dicto* construal of *a witch* captures the intended reading. The former

<sup>3</sup>This abstraction operation turns the clause from a proposition (type  $\langle s, t \rangle$ ) into a property ( $\langle e, \langle s, t \rangle \rangle$ ), as has become traditional for *de se* ascriptions after (Lewis, 1979). To handle *de re*, von Stechow introduces an additional piece of machinery again having to do with the "modes of presentation" of the individual who is the *res*.

<sup>4</sup>To simplify things, I abstract from tense and treat *will hurt* as tenseless.

fails for the prosaic reason that there are presumably no witches whereas the quantifier would commit us to their existence; the latter fares no better because the quantifier would be unable to bind the occurrence of its variable corresponding to *she* in the second subordinate clause.

Different responses have emerged to this problem. Nicholas Asher (Asher, 1987) uses “reference markers” and “referential chains” keeping track thereof; Walter Edelberg (Edelberg, 1992) employs quantification over “[intermediate]-singular-terms” standing for “ideas”; Ahti-Veikko Pietarinen (Pietarinen, 2001), working within the tradition of Independence-Friendly Logic, changes the notion of scope so that the pronoun ends up bound. Jeffrey King (King, 1993) proposes to treat pronouns as standing for *context-dependent quantifiers* whose force and range depend on their antecedent<sup>5</sup> quantifiers (e.g. on *a witch* in (9)).

What all those attempts have in common is that they unanimously reject (Geach, 1967; Edelberg, 1986; King, 1993; Pietarinen, 2001) an apparent way of circumventing the difficulty, namely to let the pronoun *she* be, in Geach’s parlance, a “pronoun of laziness” going proxy either for *the witch that has blighted Bob’s mare* or for *the witch that Hob believes has blighted Bob’s mare*. The answer typically given is that this would impose on Nob something he does not entertain in the scenario provided by Geach—a belief about Hob (and his attitudes) or at least about Bob’s mare, neither of whom Nob has to know.

Such rejection is, however, too quick. Assuming as we did before that pronouns form full-fledged DPs containing a restrictor predicate, we can again appeal to the possibility of transparent evaluation.<sup>6</sup> Just like in (4), where the predicate *expensive car* is not part of John’s attitude but rather the reporter’s insertion, here in (9) *witch who Hob believes has blighted Bob’s mare* is in no way attributed to Nob as a constituent of his attitude. This is not to say that the problem is solved: what is needed is an essential part of the solution, namely the theory of how a pronoun of laziness inherits (a part of) the meaning of its antecedent (cf. a similar difficulty in Evans—McKinsey’s and King’s analyses).

#### 4. Semantics and Epistemology of *De Re* and *De Se*

Many accounts of *de re* and *de se* within different frameworks, including (Maier, 2009; Percus & Sauerland, 2003; Charlow, 2010), employ in this or that form the notion of an *acquaintance relation* stemming from Lewis’s seminal paper (Lewis, 1979, 542). Apart from complicating the apparatus, sometimes to a considerable extent, such a technique has another characteristic that might be viewed as its disadvantage; namely, it brings *epistemological* considerations into semantics. In fact what the aforementioned theorists do is not merely outline a model of how a competent speaker understands given sentences but also, *volens nolens*, a theory of what is required (in some philosophical sense) for an agent to hold a *de re* (respectively, *de se*) attitude. In the same vein writes Nicholas Asher on Intentional Identity: “The really important factor ... is how ... reference markers were introduced and what they stand for in the minds of Hob and Nob. What is really needed here is a general theory of belief formation and communication” (Asher, 1987, 153). The point, however, is that the study of attitude *reports* need not be the study of their corresponding *attitudes*, if there are any in whatever sense of being. After all it is the

<sup>5</sup>The structural conditions on the antecedence relation are, however, different from those for anaphor binding.

<sup>6</sup>I submit that my proposal follows the spirit of Michael McKinsey’s paper (McKinsey, 1986), an exception to the aforementioned sceptical tradition. McKinsey recognises the possibility to determine the reference of the pronoun by a description recovered from the preceding discourse, just like Gareth Evans (Evans, 1977) proposed for donkey sentences, yet another die-hard puzzle. This is yet technically different from the approach I defend.

Another hint at a treatment closer to what I speculate on here is due to Robert van Rooy (van Rooy, 2000, 178, note 1) who complained that the only option for Evans-style analysis would be to give *she* wide scope in (9) but this option is unavailable due to the non-existence of witches. Recall that wide scope effects for the restrictor, not for the quantifier itself, are characteristic of the “non-specific transparent” reading.

speaker who uses the pronoun *she* in Geach's example, not Nob, so the speaker is the one to take responsibility for the coordination of the objects of thought across agents.<sup>7</sup> So there is nothing specifically *mental* in the phenomenon of Intentional Identity, *pace* McKinsey's (McKinsey, 1986) "mental anaphora" talk and Geach's original intuition that "[w]e have intentional identity when a number of people, or one person on different occasions, have attitudes with a common focus, whether or not there actually is something at that focus" (Geach, 1967, 627).

Now, whereas the ways people (or more generally, agents) use to acquire epistemic attitudes are by all means important for the verification of attitude reports, they *a priori* have virtually nothing to do with their truth conditions. Moreover, all the concept or acquaintance relation parlance may very well have nothing in common with how attitudes are ascribed in ordinary language. I can envisage no sound reason for a semantics to prefer such theoretically-laden metalanguage over a simpler one, where *de re*'ness would be taken as a primitive structural characteristic of an attitude report. What I attempted to formulate in the present paper is precisely a sketch of such a "minimal" approach to a restricted set of phenomena.

## References

- Asher, N. (1987). A typology for attitude verbs and their anaphoric properties. *Linguistics and Philosophy*, 10(2), 125–197.
- Castañeda, H.-N. (1966). 'He': A study in the logic of self-consciousness. *Ratio*, 8(December), 130–57.
- Charlow, S. (2010). *De re* anaphors. In *Proceedings of SALT 20*.
- Déchaine, R.-M., & Wiltschko, M. (2002). Decomposing pronouns. *Linguistic Inquiry*, 33(3), 409–442.
- Edelberg, W. (1986). A new puzzle about intentional identity. *Journal of Philosophical Logic*, 15(1), 1–25.
- Edelberg, W. (1992). Intentional identity and the attitudes. *Linguistics and Philosophy*, 15(6), 561–596.
- Evans, G. (1977). Pronouns, quantifiers, and relative clauses (I). *Canadian Journal of Philosophy*, 7(3), 467–536. doi: 10.1080/00455091.1977.10717030
- Fodor, J. D. (1979). *The linguistic description of opaque contexts*. Garland.
- Geach, P. T. (1967). Intentional identity. *Journal of Philosophy*, 64(20), 627–632.
- Heim, I., & Kratzer, A. (1998). *Semantics in generative grammar*. Oxford: Blackwell.
- Karttunen, L. (1969). Pronouns and variables. In *Chicago linguistic society* (Vol. 5, pp. 108–16).
- King, J. (1993). Intentional identity generalized. *Journal of Philosophical Logic*, 22(1), 61–93.
- Lewis, D. (1979). Attitudes *de dicto* and *de se*. *Philosophical Review*, 88(4), 513–543.
- Lubowicz, A. (1999). Two views of Polish reflexives. In *Proceedings of WCCFL 18* (Vol. 18, p. 337).
- Maier, E. (2009). Presupposing acquaintance: a unified semantics for *de dicto*, *de re* and *de se* belief reports. *Linguistics and Philosophy*, 32(5), 429–474.
- McKinsey, M. (1986). Mental anaphora. *Synthese*, 66(1), 159–175.
- Percus, O., & Sauerland, U. (2003). On the LFs of attitude reports. In *Proceedings of Sinn und Bedeutung* (Vol. 7, pp. 228–242).

---

<sup>7</sup>Cf.: "[I]f Arsky and Barsky are engaged in a conversation, it is Arsky and Barsky *themselves* who are *responsible* for their use of pronouns and presupposition triggers; but that when a speaker *attributes* beliefs to Arsky and Barsky or to Hob and Nob, it is *the speaker* who is *responsible* for the anaphoric and presuppositional links, and not the agents that the belief attribution is about" (van Rooij, 2006, 143).

- Pietarinen, A. (2001). Intentional identity revisited. *Nordic Journal of Philosophical Logic*, 6(2), 147–188.
- Schwager, M. (2009). Speaking of qualities. In *Proceedings of SALT* (Vol. 19, pp. 395–412).
- Sudo, Y. (2013). On *de re* predicates. In *Proceedings of WCCFL 31*.
- van Rooij, R. (2006). *Attitudes and changing contexts* (Vol. 332). Springer.
- van Rooy, R. (2000). Anaphoric relations across attitude contexts. In *Reference and anaphoric relations* (pp. 157–181). Springer.
- von Stechow, A. (2002). *Binding by verbs: Tense, person and mood under attitudes*. Retrieved from <http://www.sfs.uni-tuebingen.de/~astechow/Aufsaetze/BindingbyVerbs.pdf>