Simon Charlow, *NYU*

De re anaphors. We motivate a binding theory for pronouns in intensional contexts with distributional facts about *de re*/*de se* ±anaphoric pronouns. We present new arguments that (a) *de se* readings are grammatically encoded and (b) that grammatical processes—e.g. Conditions A and B, Fox's (2000) Economy of variable binding—are *sensitive to* the *de re* vs. *de se* distinction, *pace* Anand (2006). An Economy condition on *de re* "introduction" is proposed and defended, with consequences for Zimmermann's (1991) argument that *de re* LFs are compatible with *de se* readings.

Heim (1994) observes that a *de se* pronoun (†) can "bind" a non-*de se* (‡) anaphor. For instance, (1) is judged true if Olympia doesn't realize the person she wants to criticize is herself:

(1)   Olympia wants PRO† to criticize herself‡!

The standard view has it that *de se* attitudes stem from self-ascription of a property—cf. Lewis (1979). However, if in an attitude ascription this property originates via λ-abstraction—cf. Chierchia (1989)—mixed readings as in (1) should be impossible since *herself* must be locally bound:

(2)   Olympia$_i$ wants λ_j PRO$_j$ to criticize herself$_{*i/j}$.

Heim proposes that an anaphor may be long-distance bound just in case its "normal" antecedent is *de se*. Using new data from nearly-free control—cf. Jackendoff & Culicover (2003)—we argue that long-distance binding cannot be the mechanism underlying mixed readings:

(3)   Olympia$_i$ talked to Susan$_j$ about PRO$^\dagger_{i+j}$ criticizing each other$^\ddagger_{i+j}$!

(4)   *Olympia$_i$ talked to Ben$_j$ about PRO$^\dagger_i$ criticizing himself$^\ddagger_j$!

Since reciprocal anaphors don't tolerate split antecedents, Heim's account predicts (3) to be ungrammatical with a *de re* anaphor. Absent stipulation, Heim's account also doesn't explain the contrast between (1) and (4). We conclude that a *de se* pronoun can bind a *de re* anaphor.

Sharvit (2009) argues that a *de re* pronoun cannot bind a *de se* reflexive and that if a *de re* pronoun binds a *de re* reflexive, the two must be construed relative to the same acquaintance relation:

(5)   *Olympia$_i$ thinks she$^\ddagger_i$ burgled herself$^\dagger_i$.

(6)   Olympia$_i$ thinks she$^{\ddagger_1}_i$ burgled herself$^{\ddagger_1/?2}_i$.

We argue that Sharvit is correct about (5) but mistaken about (6). In order: while an utterance of *Olympia thinks she burgled herself* seems *compatible* with a scenario in which Olympia sees the burgl-ee in the first-person way and and burglar in the third-person way, embedding in decreasing contexts—cf. Percus & Sauerland (2003)—reveals that this cannot be due to a *de re–de se reading* of (5), *per se*:

> Susan sees herself on TV without recognizing it's her. The woman on TV is defending "the junior senator from Maine"—i.e. herself. Susan thinks "How generous of her to stick up for her colleague." Mary sees herself defending Olympia without recognizing she's on TV. Mary, a suggestible amnesiac who's under the impression she's Olympia, thinks, "How generous of her to stick up for me."

(7)   #Mary and Susan are both confused, but only MARY thinks she's defending herself.

(8)   #Mary thinks she's defending herself, but Susan doesn't.

Context requires that only Mary has a *de se* belief about being defended by someone (who we take to be) identical to her. So if a *de re–de se* reading were available for *Mary thinks she's defending herself*, both (7) and (8) would be felicitous, contrary to fact.[1] As for (6), an utterance of e.g. *Olympia thinks she killed herself* is judged true if Olympia thinks x killed y, without taking them to be identical either to herself (though they are), or each other. Focal stress on both pronouns helps bring the reading out.

This is a surprising set of facts. It's been assumed since Zimmermann (1991) that the felicity of e.g. *everyone$_i$ thinks he$_i$ won re-election* in cases where some individuals have a *de se* belief but some only a *de re* belief entails that *de re* LFs are compatible with *de se* readings.[2] But if $[x^{\ddagger_1}_i[\ldots x^{\ddagger_2}_i\text{-self}]]$ is licit, what could rule out $x^{\ddagger_2}_i\text{-self}$ being construed *de se*—thereby contradicting $*[x^{\ddagger}_i[\ldots x^{\dagger}_i\text{-self}]]$?

We assume the following (fairly standard) centered-worlds semantics for attitudes ($\mathcal{A}$):

- Following Percus & Sauerland (2003), G is a variable over *concept generators* of type $\langle e, \langle \sigma, e \rangle \rangle$, with '$\sigma$' the type of centered worlds. For any G, attitude relation $\mathcal{A}$, x_e, w_s, and any $\kappa'_\sigma = \langle w', x' \rangle \in \mathcal{A}^w_x$,[3] $[\![\text{G } y_i]\!]^{g,\kappa'} := \iota y : R(x')(w')(y)$, with R an acquaintance relation constrained such that $\iota y : R(x)(w)(y) = g_{@}(i)$. This implements Kaplan's (1968) framework for *de re* attitudes.

- $\mathcal{A}(p)(x)(w) = 1 \Leftrightarrow \exists \text{G} \forall \langle w', x' \rangle \in \mathcal{A}^w_x : p(\text{G})(x')(w') = 1$. (cf. Anand 2006)

---

1 Crucially, parallel configurations with possessive pronouns are licit in belief ascriptions: $[x^{\ddagger}_i[\ldots x^{\dagger}_i\text{'s NP}]]$, cf. Anand 2006.

2 See (11) below for one example of how a *de re* LF can yield a *de se* reading.

3 $\langle x', w' \rangle \in$ e.g. $\text{Dox}^w_x \Leftrightarrow x'$ is someone x thinks she might be in w, and w' is a world compatible with what x believes in w.

- Bare (i.e. G-less) pronouns get bound by a syntactic λ operator—c.f. (2). So they're *de se*.

Additionally, we assume *de re* pronominal DPs are *derived* from "bare" (i.e. *de se*) pronominal DPs via a structure-building operation: i.e. $[_{\text{DP}}\, x_i] \rightsquigarrow [_{\text{DP}}\, \text{G}\, x_i]$. Fox (2000) argues that LF transformations should have semantic import. And plainly, if for any G, x_i, assignment g, world w, attitude relation $\mathcal{A}$, and doxastic state Dox_x^w, $\forall \kappa' \in \text{Dox}_x^w \,.\, \text{G}(\llbracket x_i \rrbracket^g)(\kappa') = g(i)$, the *de re* transform will be for naught. We propose that all G-introductions meet the following criterion: $\exists \kappa' \in \text{Dox}_x^w \,.\, \text{G}(\llbracket x_i \rrbracket^g)(\kappa') \neq g(i)$. Informally, Gs should *do something* in the information state of the belief ascribee. Given an LF $[[\text{G}_1\, x_i]\, [\ldots [\text{G}_2\, x_i\text{-self}]]$, this entails that neither pronoun can be intereprted *de se*. If the LF $[[\text{G}\, x_i][\ldots x_i\text{-self}]]$ can be ruled out (see below), we derive $*[x_i^{\ddagger}[\ldots x_i^{\dagger}\text{-self}]]$.

Of course, this proposal entails that *de re* pronouns of *all* sorts may never be construed *de se*. Isn't this inconsistent with Zimmermann's data? Evidence it's not comes from the following:

> Olympia, Susan, and Mary are watching TV. As it happens, some talking head is on the air criticizing Olympia, Susan, and Mary. Olympia and Mary, somewhat confused, realize only that the talking head is criticizing some politician. Each thinks, "I want to defend those poor souls." Susan, on the other hand, realizes what's going on and thinks, "I want to defend myself."

(9) Olympia wants to defend herself, and $\{^{\checkmark}\text{Mary}, {}^{\checkmark}\text{Susan}\}$ does too.

(10) $\checkmark$ Each of those senators wants to defend herself.

Since $*[x_i^{\ddagger}[\ldots x_i^{\dagger}\text{-self}]]$, we concluded *de re* anaphors couldn't be interpreted *de se*. Yet the reflexive is licensed in a mixed scenario (exactly as in Zimmermann's case)! Actually, the proposed Economy condition predicts this since only a "do-nothing" G can make *Susan wants to defend herself* true and thus yield true readings of (9) and (10) (parallel reasoning rescues Zimmermann's example):

(11) $\lambda \text{G}[\forall \kappa' = \langle w', x' \rangle \in \text{Boul}_{\mathbf{s}}^{w@} : \mathbf{defend}(\text{G}(x')(\kappa'))(x')(w')](\lambda y_e \lambda \kappa_\sigma . y)$
$\rightsquigarrow_\beta \forall \kappa' = \langle w', x' \rangle \in \text{Boul}_{\mathbf{s}}^{w@} : \mathbf{defend}(x')(x')(w')$ [**True in our scenario**]

Ruling $[[\text{G}\, x_i]_j\, [\ldots x_i\text{-self}]]$ out—while ruling $[x_i\, [\ldots [\text{G}\, x_i\text{-self}]_j]$ in—remains. We propose an account which (a) leaves "bare" objects in situ but (b) forces *de re* objects to raise at LF—à la object shift or focus movement. Taken together with the assumption that Binding Theory cares only about *extensional* covaluation, this derives the asymmetry:

(i) ✗$[[\text{G}\, \text{she}_i]_j\, [\lambda_j\, [_{\text{VP}}\, t_j\, [\text{defends herself}_i]]]]$ (ii) ✓$[\text{she}_i\, [\lambda_i\, [\text{G}\, \text{herself}_i]_j[\lambda_j\, [_{\text{VP}}\, t_i\, [\text{defends}\, t_j]]]]]$

In (i) *herself*$_i$ must, given Condition A, be bound by the trace t_j of $[\text{G}\, \text{she}_i]_j$. Since (given Economy of G introduction) $g(i) \neq g(j)$ this move is illicit. In (ii), by contrast, she$_i$ binds $[\text{G}\, \text{herself}_i]_j$. Since (by Percus & Sauerland's 2003 rules for G) $\llbracket \text{G}\, \text{herself}_i \rrbracket^{g,\kappa@} = g(i)$, this move is licit.

We thus maintain, *pace* Anand (2006), that Binding Theory "sees" *de re* vs. *de se*—a conclusion I think the asymmetry between (1) and (7)–(8) impels. Some more evidence this is on the right track:

(12) Olympia said she voted for her bill, and Susan did too.

(12) lacks a STRICT–SLOPPY reading, a fact known as Dahl's (1973) puzzle. Fox (2000) proposes an account based on locality of variable binding: the missing reading requires long-distance binding of *her* by *Olympia*; however, this is ruled out since local binding yields a truth-conditionally equivalent result. If this mechanism is sensitive to the truth-conditional difference between *de re* vs. *de se* readings, we should observe obviations of Dahl's effect if a *de re* pronoun intervenes between a *de se* pronoun x_i and the *de se* abstraction operator λ_i. In fact, this is precisely what obtains:

> One day Olympia and Susan return home to find that both of their computers have been stolen. They discover that a plant has been knocked over in the living room and surmise that whoever the thief was, (s)he must have been the one who knocked over that plant. In reality, Olympia knocked the plant over the previous evening in a drunken stupor, an incident she has completely forgotten.

(13) Well this is funny. Olympia thinks SHE stole her computer, and SUSAN does too.

Finally, we note that obviations of Condition B effects can occur if the binding pronoun is *de re*:

(14) Well this is funny. Olympia thinks SHE($_i$'s the one who t_i) robbed her.

Though (14) should induce a Condition B violation—see Sharvit (2009) on Condition B effects for *de re* pronouns—the construction is only slightly deviant, and the parenthetically given form is impeccable.

Anand, P. 2006. De De Se. **Chierchia, G.** 1989. Anaphora and attitudes *de se*. **Dahl, O.** 1973. On so-called 'sloppy identity'. *Synthese* 26. **Fox, D.** 2000. *Economy and Semantic Interpretation.* **Heim, I.** 1994. Puzzling reflexive pronouns in *de se* reports. **Jackendoff, R. & P.W. Culicover.** 2003. The Semantic Basis of Control in English. *Language* 79. **Kaplan, D.** 1968. Quantifying in. *Synthese* 19. **Lewis, D.** 1979. Attitudes *De Dicto* and *De Se. Phil Review* 88. **Percus, O. & U. Sauerland.** 2003. On the LFs of Attitude Reports. *Proceedings of SuB 7.* **Sharvit, Y.** 2009. Covaluation and Unexpected BT Effects. **Zimmermann, T.E.** 1991. Kontextabhängigkeit.