3 Domains and orientation

In chapter 1, section 1.2.5 we arrived at a general format for Binding Conditions, which I repeat here:

(3.1) An NP of **class** must **(not)** be coindexed with a **commanding** NP within its **domain**.

We saw that the grammar of English contains three conditions of this form, pertaining to three disjoint **classes** of NPs. In the discussion up until now we have assumed that the relevant binding **domain** in which reflexives need to be bound, and non-reflexive pronouns need to be free, is the local clause.

In this chapter we are first going to refine this notion of binding domain. The discussion will start with English, again, showing that the correct description of the relevant domain should be something like 'smallest category containing a subject,' rather than just 'smallest clause.'

Second, we will introduce an additional parameter into the Binding Conditions, *orientation*. Whereas orientation doesn't seem to be central for the description of English pronouns (but see chapter 11), it is very important in many other languages. We will introduce three kinds of orientation: subject-orientation, anti-subject orientation, and logophoricity.

Third, we will then examine a range of data from other languages that requires Binding Conditions to make reference to different domains (both different from English, and different among the classes of NPs within a given language).

Fourth, we will briefly discuss so-called long-distance reflexives, that is, cases in which reflexive pronouns lead a 'double-life' as either locally bound (like in English) or bound in the sentence domain, like subject oriented pronouns and logophors.

With all these refinements in stock, then, we will fifth and finally look at a few complete pronominal systems in languages other than English.

3.1 Binding domains in English: governing category

3.1.1 Exceptional case marking (ECM)

Consider first so-called *exceptional case marking* (ECM) constructions. The hallmark of these is that a certain class of *ECM-verbs* case-mark an NP

which is thematically the argument to an embedded verb. Examples are the verbs *believe* and *want* in English. The thematic subject of the lower verbs, *Georgina*, behaves like a grammatical object to the higher ECM-verbs: it is marked with accusative case, (3.2b), and gets promoted under passivization of the matrix verb, (3.2c). We will henceforth refer to these phrases, italicized in the following examples, as the *ECM-subjects*:

(3.2) (a) O'Leary wants/believes *Georgina* to lie.
(b) O'Leary wants/believes *her* to lie.
(c) *Georgina* is believed to lie.

Turning to binding, now, we observe that the ECM-subjects behave like they are clausemates to the matrix subject: coreference requires a reflexive and prohibits a pronominal:

(3.3) O'Leary$_6$ believes himself$_6$/him$_{*6}$ to deserve the crown of England.

At the same time the ECM-subjects behave like clausemates with respect to NPs in the embedded clause: they trigger reflexivization in the lower clause and do not tolerate coreference with a non-reflexive object pronoun:

(3.4) O'Leary wants Georgina$_8$ to protect herself$_8$/her$_{4/*8}$.

These two findings are yet unproblematic. They seem to suggest that all NPs in an ECM-construction, whether they thematically belong to the embedded verb or the matrix verb, populate the same binding domain in the sense relevant to Binding Conditions A and B. This, however, predicts that the matrix-subject should count as a clause-mate to the embedded non-subjects, too, which it patently does not:

(3.5) O'Leary$_{12}$ wants Georgina to protect him$_{12}$/*himself$_{12}$.

At this point we are in a bind: whatever constituent is the binding domain for the embedded object must include the ECM-subject (because of [3.4]), but exclude the matrix subject (because of [3.5]). But then that constituent, even though containing the ECM-subject, cannot be the binding domain for the ECM-subject (because of [3.3]). In other words, the pertinent binding domains must *overlap* as in (3.6):[1]

(3.6) $\overbrace{[_{S^m}\,\text{SUBJ}^m \ldots V \underbrace{[_{S^e}\,\text{SUBJ}^e\,[\,\text{V OBJ}\,]]}_{\text{binding domain for OBJ}}}^{\text{binding domain for SUBJ}^e}$

These domains are correctly computed by the definition of *governing category* (GC) in (3.7):

[1] The schema in (3.6) assumes that the ECM-subject resides within the embedded clause. Essentially the same conclusion would hold, however, if we assumed that it raises to the matrix clause (i.e. if these sentences are to be analyzed as *raising to object*).

(3.7) γ is the *governing category* (GC) for NP if and only if (iff) γ is the smallest clausal category (S, $\bar{S}$, IP, CP, TP…) which dominates
(a) NP
(b) NP's case assigner

(3.8) Binding Conditions (preliminary):
(A) A reflexive pronoun must be bound in its governing category.
(B) A non-reflexive pronoun must be free in its governing category.

The hypothesis expressed by (3.8), then, is that the binding domain for both reflexive and non-reflexive pronouns in English is their governing category. (3.7) can perhaps be understood best by looking at (3.6): the embedded object, OBJ, receives its case from the embedded V, hence its GC is the embedded clause. Accordingly, the embedded object must be reflexive if coreferent with the ECM-subject, but not if coreferent with the matrix subject. The ECM-subject, on the other hand, receives its case from the *matrix*-verb (exceptionally, as it were), which means that the embedded clause is *not* its GC: even though it contains the ECM-subject itself, it doesn't contain its case assigner. Indeed the smallest clausal category containing both the ECM-subject and its case assigner is the matrix clause. Accordingly, coreference of the ECM-subject with the matrix subject requires the former to be reflexive.

3.1.2 Infinitival clauses

Next, let us briefly look at infinitival complements other than ECM. At first, pairs of examples like those in (3.9) seem to provide contradictory evidence. (3.9a) suggests that the binding domain for the embedded object is the matrix clause, but (3.9b) suggests that it is something smaller than that, e.g. the infinitival clause/VP:

(3.9) (a) John_3 tried to educate $\text{himself}_3/*\text{him}_3$.
(b) Ana_1 told John to educate $\text{her}_1/*\text{herself}_1$.

The mystery is resolved, however, once we recognize that *try* is a *subject-control verb* (the understood subject of the embedded verb is the matrix subject), whereas *tell* is an *object-control verb* (the understood subject of the embedded verb is the matrix object). We can then establish the following generalization:

(3.10) An infinitival clause functions as a G(overning) C(ategory), with the understood subject acting as a binder.

This generalization receives confirmation from sentences like (3.11) (compare to [3.9b]), and the structurally parallel Marathi example (3.12):[2]

[2] From Dalrymple (1993):17.

(3.11) Ana told John$_3$ to educate himself$_3$/*him$_3$.

(3.12) John$_1$ ne Jane$_2$ laa swataahlaa$_{2/*\,1}$ maraaylaa saangitle (Marathi)
John ERG *Jane* DAT *self*-ACC *hit* *told*
'John told Jane to hit self (=Jane).'

In both examples the embedded reflexive corefers with the controller, which happens to be the matrix object. What makes the Marathi case interesting is that the reflexive *swataah* is subject oriented; it cannot ever be bound by a non-subject argument (see section 3.2 below). If the matrix object *Jane laa* were the actual binder in (3.12), this sentence would constitute a mysterious exception to the subject-orientation of the reflexive. But assuming, as we did, that the logical subject counts as a binder in these cases provides an immediate explanation.

How can we implement (3.10)? If we assume that phrase structural c-command is the relevant notion, we will have to assume an actual empty NP as the subject of the embedded infinitival clause, commonly called *PRO*. The index on *PRO* depends on the matrix verb (subject, object, or arbitrary control), but *PRO* is the inevitable binder for a reflexive or reciprocal in the embedded clause. Representative structures are given in (3.13):

(3.13) (a) John$_4$ tried [*PRO*$_4$ to educate himself$_4$]. (subject control)
(b) Ana$_2$ told John$_4$ [*PRO*$_{4/*2}$ to educate himself$_4$/*herself$_2$]. (object control)

It bears emphasizing that none of the binding principles discussed so far is responsible for the indexing of *PRO* in (3.13), e.g. that it must be indexed 4 rather than 2 in (3.13b). BT as discussed here only governs the indexing of the embedded pronouns.[3]

Alternatively, if we pursue an argument-structure based account of binding, things are even more straightforward. All that needs to be assumed is that verbs like *try*, *tell*, and adjective-based predicates like *be dangerous* select for a VP, and encode that the index of the (unsaturated) subject-argument of that VP be unified with the pertinent element on the a(rgument)-list of that verb:

(3.14) (a) a-list for *try*: ⟨NP$_1$, VP$_{to}$[SUBJ[⟨1⟩]]⟩
(b) a-list for *tell*: ⟨NP, NP$_{acc,1}$, VP$_{to}$[SUBJ[⟨1⟩]]⟩
(c) a-list for *promise*: ⟨NP$_1$, NP$_{acc}$, VP$_{to}$[SUBJ[⟨1⟩]]⟩

The indices indicate that NP$_1$ is on the a-list of both the matrix verb and the embedded verb (by virtue of being the SUBJ of the embedded VP). Accordingly, more oblique elements on the a-list of the embedded verb can be bound by NP$_1$, regardless of the position of NP$_1$ on the matrix a-list; for more details see Sag and Pollard (1991); Sag and Wasow (1999).

[3] The interested reader is referred to Harbert (1995) and the references therein for attempts at subsuming the choice of (co)reference for *PRO* under BT as well; for a recent discussion of control see Landau (2001a,b).

Exercise 3.1

The Marathi sentence in (3.12) made an argument that the reflexive is not bound by an NP in the matrix clause, but by the (invisible) subject of the infinitival clause. Make an argument to the same effect using the English sentences in (3.15):

(3.15) (a) It is embarrassing to see yourself in the newspaper headlines like that.
(b) It is dangerous to trust each other like that.
(c) John agreed with Mary to bring each other's pictures to the meeting.

3.1.3 NPs as binding domains

In this section we will briefly consider another modification of the definition of G(overning)C(ategory).

NPs with and without subjects

So far we have almost exclusively looked at NPs in verbal argument position (subject, object), when examining Binding Conditions. If we expand our data base, we find that NP positions in adnominal argument positions seem to behave as predicted:

(3.16) John_5 saw $[_{NP}$ a picture of $\left\{ \begin{array}{c} \text{himself}_5 \\ *\text{him}_5 \end{array} \right\}$]

The pronoun, functioning as an argument to the noun *picture*, receives its case from the preposition *of* within the NP. But the only clausal node dominating the pronoun is the matrix S, which also dominates the subject. Therefore the subject is in the GC of the nominal argument, which therefore must be reflexive if coreferent with the subject.

Things change, however, if we consider NPs with a prenominal argument (a *possessor*); now the non-reflexive is permitted, and, according to many authors, only the non-reflexive:

(3.17) John_5 saw $[_{NP}$ Mary's picture of $\left\{ \begin{array}{c} ^{??}\text{himself}_5 \\ \text{him}_5 \end{array} \right\}$]

Assuming these judgments, it seems that an NP functions like a (finite) clause with regard to BT if and only if it has a possessor. Accordingly, the following revision of the definition of GC in (3.7) suggests itself:

(3.18) γ is the *governing category* for NP, iff γ is the smallest category **that has a Subject** and dominates
(a) NP
(b) NP's case assigner
(where a Subject is either a clausal subject or a possessive)

Applied to all cases in sections 3.1.1 and 3.1.2, this definition will yield the same results as before, given that clauses contain subjects (and clausal subjects are Subjects in the technical sense). If a pronoun is contained in an NP, however, that

NP, rather than the clause minimally containing it, can be the GC, if it contains a possessor. This is the case in (3.17): the GC for *him(self)* is the NP *Mary's picture of him(self)*, which implies that a reflexive has to be bound within that NP (which it evidently can't, given the gender mismatch). The NP in (3.16), on the other hand, is not the GC for *him(self)* because it contains no Subject, so the entire clause is. Accordingly, a pronoun must be free within the entire clause, and coreference with the clausal subject requires reflexivization.

This is probably the right time to note that at least one part of the generalization that motivates the formulation in (3.18) is not borne out by the facts, namely that post-nominal reflexives in NPs *with* a possessor only allow binding to the possessor. In a magnitude estimation experiment with 52 English speakers, Keller and Asudeh (2001) found that reflexives and non-reflexives are judged equally acceptable in a sentence like (3.19a) (in contradistinction, a reflexive in a sentence like *Joan's father respects herself* was clearly rejected). Similarly, Runner *et al*. (2002) found in an eye-tracking experiment that almost 25 percent of their subjects interpreted the reflexive to denote Ken in a sentence like (3.19b); similarly, examples of reflexives bound from outside a possessive NP are widely attested, e.g. (3.19c):[4]

(3.19) (a) *Hanna* found Peter's picture of *her(self)*.
(b) Have *Ken* touch Harry's picture of *himself*.
(c) "C.B.'s father had fared better in this respect than most of his forebears, but still resented his wife for *her low opinion of himself*, of the Whiting mansion, of Empire Falls, of the entire backward state of Maine . . . "

Asudeh and Keller's experiment also revealed that post-nominal non-reflexives bound to a possessor or a local sentential subject in possessor-less examples (essentially the him_5 variant of [3.16]) – both of which are predicted to be ungrammatical by the account given here – while not fully acceptable, are significantly better than non-reflexive object pronouns bound to verbal coarguments.

It is a well-acknowledged fact that the data in this area of BT are complex and hard to judge (see e.g. Kuno [1987]:section 4.3; Reinhart and Reuland [1993]:683, 690). Experiments like Runner *et al*. (2002) and Keller and Asudeh (2001) are of utmost importance in that they provide a way of establishing a reliable data base even where individual speakers' introspective judgments are insecure or inconsistent. For a theoretical interpretation of some of these findings, see Asudeh and Keller (2002).

Non-complementary environments

Let me finally discuss an additional complication regarding pronouns within NPs, and briefly sketch a way of addressing them that roughly follows the proposals in Huang (1983) and Chomsky (1986):164ff. Although these proposals

[4] From Richard Russo, *Empire Falls (Kampf)*, p. 4, found by C. Potts (italics added).

have received a lot of attention in the literature, I believe they are superseded by the simpler and more accurate treatments afforded within the proposals discussed in chapter 11, so I will skip details wherever possible.

The complementarity between reflexive and non-reflexive pronouns has been at the heart of the definitions we have provided so far. It turns out, however, that while complementarity is observed in the vast majority of cases, it isn't always. (3.20) is a case in question:

(3.20) John$_7$ believes [that [$_{S^e}$ pictures of $\left\{\begin{matrix}\text{him}_7\\ \text{himself}_7\end{matrix}\right\}$ are on sale]]

According to our definition (3.18) the GC for *him(self)* in (3.20) should be the embedded clause S^e: it contains the pronoun, its case assigner *of*, and a subject *pictures of him(self)*. This result yields the correct prediction for the non-reflexive *him*, which is free in S^e and can thus corefer with the matrix subject. By the same token, however, it blocks the reflexive, which is not bound in S^e.

It seems that reflexives and non-reflexives part company here: the GC for *himself* appears to be the matrix clause; that for *him* the embedded clause. Suppose this generalization is correct, then the question emerges if there is any way to define GC so as to get these two different domains for *him* and *himself* in (3.20). One ingenious attempt at that is found in Chomsky (1986).

Perhaps the best way to illustrate the gist of Chomsky's proposal is this: the GC for the reflexive in (3.20) doesn't contain any c-commanding NP (note that the embedded subject itself doesn't c-command an NP it contains). It is thus *in principle* impossible for the reflexive to meet the Binding Condition pertinent to it, Binding Condition A, within that GC. This is different for the non-reflexive, because a GC without any c-commanding NP is just a special case of a GC without a binder, so a non-reflexive can, and in fact always will, meet its Binding Condition, Binding Condition B, within such a GC. The idea then is that the GC for a given NP must be chosen 'mercifully,' in such a way that NP can at least in principle meet its Binding Condition in that GC. Consider the following revision of (3.18) (cf. Chomsky 1986:171f.):

(3.21) γ is the *governing category* for NP iff γ is the smallest category that has a Subject and dominates
- (a) NP
- (b) NP's case assigner
- (c) an NP$'$ c-commanding NP, if NP needs to be bound

Take sentence (3.20) again, with a reflexive (*John believes that pictures of himself are on sale*). The GC for *himself* according to (3.21a–3.21b) would be the embedded clause S^e, but that clause doesn't contain any c-commanding NP, which could function as the binder of *himself*, as required by Binding Condition A. Accordingly, (3.21c) mercifully 'broadens' the GC to the next clause up, which indeed contains a binder for the reflexive.

In the case of a non-reflexive (*John believes that pictures of him are on sale*), (3.21a–3.21b) again determine S^e as the GC, but this time (3.21c) doesn't change anything about that because Binding Condition B, the Binding Condition pertinent for non-reflexive pronouns, doesn't require any binder at all. So S^e is the ultimate GC for *him*, in which it is free, as required.

It remains to verify that even the relaxed definition of GC in (3.21) doesn't rule in cases like (3.22):

(3.22) (a) $*\text{Mary}_3$ said that [$_{S^{e1}}$ John believes that [$_{S^{e2}}$ [pictures of herself$_3$] are on sale]].
(b) $*\text{John}_1$ believes that [[Mary's pictures of himself$_1$] are on sale].

The GC for *herself* in (3.22a) is S^{e1}, not the matrix clause, because S^{e1} already contains an NP which could bind the reflexive, meeting Binding Condition A; the fact that *John* cannot be the actual antecedent to *herself* is irrelevant to (3.21c). Likewise, *himself* in (3.22b) has the subject NP *Mary's picture of himself* as its GC, given that that NP contains a c-commanding NP, *Mary's*, that could serve as the binder (were it not for the gender mismatch).

The definition of GC in (3.21) predicts non-complementarity for two more positions: possessives and clausal subjects, as both of these, being the Subject in the sense of (3.21), do not have a c-commanding NP in their 'original' GC. This prediction turns out to be correct for the former case, but incorrect for the latter (we use a reciprocal here, since English doesn't have reflexive possessives):

(3.23) (a) They_6 love [$\left\{ \begin{array}{l} \text{their}_6 \\ \text{each other}_6\text{'s} \end{array} \right\}$ pictures].
(b) $*\text{They}_3$ think [that each other$_3$ will win].[5]

An independent reason why (3.23b) is unacceptable has been proposed in Rizzi (1989), namely that reflexives universally cannot occur in agreeing positions.[6] But even if Rizzi's generalization is correct, reasonable doubts about the validity of the 'mercy-condition' on the definition of GC have been voiced, and alternative and more comprehensive accounts have been proposed (see chapter 11).[7]

[5] Lebeaux (1983) and many following him have claimed that sentences like (3.23b) are slightly better than full-blown Binding Condition A violations, and become virtually acceptable in the context of *wh*-extraction such as ??*They don't know what each other are doing*. I am not aware of any coherent account of this contrast.

[6] A very different attempt at explaining the ungrammaticality of (3.23b), involving movement of the reflexives, is found in Chomsky (1986).

[7] It is also instructive to note that binding of reflexives and reciprocals in embedded finite subjects is by far not universally allowed. Languages as closely related as Dutch and German strictly prohibit this:

(i) Martell hofft, dass eine Reportage über $\text{ihn}_1/*\text{sich}_1$ im Radio gespielt wird. (German)
M. hopes that a report about him/himself in-the radio played becomes
'Martell is hoping that a report about himself is going to be aired.'

Likewise, Kannada, Italian (Yang, 1983); Polish (Reinders-Machowska, 1992); Russian (Rappaport, 1986) do not allow this kind of binding for reciprocals and/or reflexives.

Exercise 3.2

Assume we replace (3.21c) by a clause that says 'enough material for NP to meet its Binding Condition (ignoring mismatches in person, gender and number),' while maintaining the rest of definition (3.21) and the Binding Conditions in (3.8). This would seem to express the idea of 'be merciful where appropriate' even better. Yet it is haunted by a fatal formal problem. Which?

3.1.4 PPs as binding domains

According to the definition of GC in (3.18), a GC needs to contain a Subject. Whatever the details of the technical Subject notion, it seems clear that PPs don't contain a subject and should therefore not constitute a GC for their complement NP. In other words, the prediction is that in *[$_{PP}$ P NP]*, NP can be a reflexive bound from outside of the PP; by the same token, if NP is non-reflexive, it should have to be free within the next higher domain containing a subject (e.g. the clause of which PP is an immediate constituent). What are the data? Consider (3.24):

(3.24) (a) John$_1$ sent a letter to him$_{*1}$/himself$_1$.
(b) John$_1$ always relies on him$_{*1}$/himself$_1$.

These sentences are as expected. But they contrast with the superficially parallel (3.25):

(3.25) (a) John$_1$ looked around him$_1$/himself$_1$.
(b) John$_1$ pulled the blanket over him$_1$/himself$_1$.
(c) Muhammad$_1$ hid the book behind him$_1$/himself$_1$.

Here it seems as if the binding domain for *him* must be smaller than that for *himself*. For example Hestvik (1991), following unpublished work by Joan Bresnan, proposes that the binding domain for *him* is the PP, while it is the clause for *himself*; accordingly, complementarity between *him* and *himself* breaks down, as the former is free within PP, while the latter is bound within S.

If we adopt this kind of analysis, it means we give up on the assumption that there is *one* binding domain, the governing category, that is relevant for both Binding Condition A and Binding Condition B (we effectively gave up that assumption in section 3.1.3 above, but now it seems less likely that there could even be a uniform *formulation* of GC). While the binding domain for reflexives can remain what it was (the smallest category containing it, its case assigner, and a Subject), the binding domain for a non-reflexive pronoun should include the subject only if the pronoun is a complement of a verb (to block *He_4 *likes* him_4), but not if it is a complement of a preposition (as in [3.25]).

Let us define NP's *coargument domain* as the smallest XP that contains NP, NPs case assigner C, and all other arguments of C (cf. section 3.3 below). Since

subject and all objects are arguments of a verb, the coargument domain for any verb argument is its minimal clause. Given that a preposition has only a complement, but no subject, the coargument domain of an NP selected by a P is the PP. This is the distinction we're after:

(3.26) Binding Conditions (final)
(A) A reflexive must be bound within the smallest category containing it, its case assigner, and a Subject (=its GC).
(B) A non-reflexive must be free in its coargument domain.

Given these definitions, we account for the non-complementarity in (3.25), but we lose our account of the complementary distribution in (3.24), because there, too, the non-reflexives would now be free within the PP. Now, there is arguably a difference between the PPs in (3.24) and those in (3.25). While they are all selected by the verb (or so we will assume), the prepositions in (3.25) make a clear semantic contribution to the sentences, while those in (3.24) seem semantically empty.

To be sure, it is not easy to motivate this distinction in every specific case (surely *on* and perhaps even *to* can have semantic content in other sentences). It is, however, suggestive that the Ps in the complements to *look, pull*, and *hide* can be exchanged for others as in *look behind NP, pull it around NP*, or *hide it next to NP*, and the whole PPs for proforms as in *look there, pull it up*, and *hide it away*. No such variation is possible in the case of *sent* or *rely* (**sent it on him,* **rely there*, etc.). Let us assume for concreteness, then, that *look, pull*, and *hide* truly select a PP that denotes a path (or location), while *sent to* and *rely on* are really complex verbs that semantically combine with the NP denotation, so that the P has no semantic function on its own.

Assuming this difference, how can it help us to explain the difference in the acceptability of non-reflexive pronouns? The intuition we are after is that the PPs in *rely [$_{PP}$ on NP], sent X [$_{PP}$ to NP]* and their likes are not the coargument domain for NP, because NP is 'really' an argument to V, whose arguments include the subject. Now, this would follow from our tentative definition of coargument domain above, if we could plausibly argue that V, not P, is the case assigner for NP in these cases. That, however, isn't obvious, and we will not pursue this option further here. Failing that, the non-domain status of these PPs would also follow if we replaced the notion of case assigner in the definition of coargument domain by the notion of thematic role (Θ-role) assigner, or semantic predicate. Consider the schemata in (3.27):

(3.27) (a) John$_1$ relies [$_{PP}$ on [$_{NP}$ himself$_1$/*him$_1$]] coargument domain

(b) John$_1$ looked [$_{PP}$ around [$_{NP}$ himself$_1$/him$_1$]] coargument domain

Since *him* receives a Θ-role from *relies* in (3.27a), all arguments to *relies* are within its coargument domain, in which it needs to be free. In (3.27b), on the other hand, *him* receives its Θ-role from *around* (while the entire PP in turn gets a Θ-role from *looked*), which doesn't have any further arguments and therefore makes PP the coargument domain, in which *him* must be, and is, free. This will be our account of the different behavior in these cases.

The alert reader will recall that a non-reflexive pronoun cannot be bound to a local subject, even if the subject and the pronoun aren't thematic coarguments. The crucial example involving exceptional case marking is repeated here:

(3.28) O'Leary$_6$ believes himself$_6$/him$_{*6}$ to deserve the crown of England.

Note that by a purely semantic definition of coargument domain in terms of Θ-role assignment, *him*'s coargument domain in (3.28) should be the embedded clause, in which it is free, which wrongly predicts possible binding of *him* by *O'Leary* in (3.28). If, however, we defined the coargument domain to include the *case* assigner (rather than the Θ assigner) and its other arguments, we're back at predicting *him* to be possible in (3.24) (assuming that the case assigner is the semantically vacuous P). It seems we have to bite the bullet and adopt a disjunctive formulation of coargument domain as in (3.29):

(3.29) NP's *coargument domain* is the smallest constituent X which contains (i) NP, (ii) NP's case assigner C, (iii) NP's Θ-role assigner T, and (iv) every XP whose case or Θ-role is assigned by C or T.

This, I grant, is an ugly definition, but it copes with both kinds of PP cases, as well as the ECM cases (and raising cases, to be discussed in chapters 11 and 12). We will return to some of the issues involved here in chapter 11, especially section (11.4.3), but the basic disjunction will stay with us.

In the discussion of binding domains in section 3.3 below, we will abstract away from these complications and use a simpler definition of coargument domain, ignoring the fact that similar data to the ones discussed in the present section are found in other languages as well (see e.g. Hestvik [1991] and the references there).

Before leaving the issue of PPs as binding domains, let me briefly mention cases of true and uncontroversial adverbial PPs, as opposed to selected PPs, as in (3.30):

(3.30) (a) Max$_1$ saw a ghost next to him$_1$/himself$_1$.
(b) John$_1$ found a dollar bill in front of him$_1$/$^{??}$himself$_1$.
(c) John heard a strange noise behind him/*himself.

The grammaticality marks in (3.30a), (3.30b), and (3.30c) are as given by Reinhart and Reuland (1993):686; Hestvik (1991); and Kuno (1987):65,

respectively;[8] this variability of judgments is surprising, since the examples are parallel in all relevant respects, and no single author offers reasons to distinguish them.

How does our current approach fare here? The acceptability of the non-reflexive is expected and predicted, since clearly the P heading the adverbial PP is semantically potent and thus defines its own coargument domain. However, the impossibility or degraded acceptability of a reflexive – as reported in (3.30) – is unexpected, since the smallest category containing a Subject is obviously the clause. We could introduce yet another quirk in the definition of binding domain for reflexives, if we wanted to rule out reflexives in adjunct PPs systematically; examples of such a strategy can be found in Hestvik (1991) and Kuno (1987):ch. 2. Generally, though, reflexive pronouns in adjunct PPs are found rather frequently in actual text:[9]

(3.31)
(a) Mrs B. who hears the steps behind herself feels rottenly and goes faster.
(b) He supplied the end of the cord near himself with a conductor in the shape of an iron key.
(c) Mr. Meynell, finding him in better health but suffering from the loneliness of his life, brought him to London and established him near himself.
(d) The seductress must be careful not to cast this spell near herself.

Given this, we will assume that reflexives in adjunct PPs can generally be bound from outside, leaving open the question of what accounts for the degraded status of (some of) these examples; see chapter 11, especially section 11.3.2, for further discussion.

This concludes our discussion of the distributional data from English. In the face of considerable unclarity in the data, we tentatively adopt the Binding Conditions A and B in (3.26), which fix the binding domains for reflexives and non-reflexives as the subject domain and coargument domain in the sense to be introduced in (3.3) below, respectively.

In the following sections we will extend our perspective to a wider variety of languages. We will see that neither the tripartition into reflexives (and reciprocals), non-reflexive pronouns, and full NPs, nor the notion of GC developed in this section seem directly applicable cross-linguistically. Nonetheless I will attempt to make visible the outlines of a general format of Binding Conditions.

[8] Reinhart and Reuland (1993):687 remark, though, that the reflexive in (3.30a) is 'much more marked' than in sentences like (3.24), which they attribute to the competing possibility of using a non-reflexive (though they don't say why that should be preferred); Kuno on the other hand uses stars but describes the contrast by merely stating that '... for many speakers the following sentences are better with a [non-reflexive; DB] pronoun than with a reflexive' (p. 65).

[9] From http://www.plew.de/english/Note.htm, a description of Benjamin Franklin's famous experiment at http://www.home.zonnet.nl/kitedude/VEK2.htm, the *Catholic Source Book* on Francis Thompson (http://www.newadvent.org/cathen/14703b.htm), and the *Realm of the Dragons* AD&D community and info. site, respectively.

A word on the terminology to be used. I will use the term *pronoun* in the usual sense, including forms that are not inherently anaphoric, such as first and second person pronouns. The term *reflexive* will be used in the text and the glosses for pronouns that need to be bound, including forms that can be bound across clause boundaries, where this is done by the authors whose presentation I allude to (a more apt term for these pronouns might be *command anaphor*, which would leave the term *reflexive* for a morphologically defined class).

I will use the general term *binding domain* instead of *governing category*. For one thing, it is unclear whether binding domains relevant in other languages are usefully described in those terms used in the definition of GC. For another, it is a particular property of English that the two relevant pronoun-classes, reflexives and non-reflexives, make reference to the same domain (if indeed they do) for their respective Binding Conditions. As we will see, different pronouns within the same language, and even different conditions pertaining to the same pronoun, can make reference to different domains. These domains will be designated by more informative names. The GC as defined in (3.18)/(3.26), for example, which is the binding domain for reflexives in English, will be called the *subject domain*.

3.2 Orientation

In this section we will extend our blueprint for binding conditions by one parameter, *orientation*. Orientation is not relevant in English, but has an important role to play in many other languages. To a first approximation, orientation means that a certain anaphoric element must be bound to or free from NPs with a particular grammatical function, in most cases a subject.

3.2.1 Subject orientation and anti-subject orientation

An illustrative example of subject orientation is the Chinese reflexive *ziji*. As widely discussed in the literature, *ziji* must be bound within its root clause, but not necessarily within any local domain. However, the antecedent has to be a subject:[10]

[10] Dalrymple (1993) (2.41); Tang (1989) (45); according to Huang (1982), *ziji* can only find a nonlocal antecedent if it is itself in subject position. But as Tang (1989) argues using (i), this doesn't seem correct:

(i) Zhangsan$_3$ juede Lisi$_8$ dui **ziji**$_{3/8/*1}$ mei xinxin.
Zhangsan think Lisi to self no confidence
'Zhangsan thought that Lisi had no confidence in himself.'

Also, *ziji*'s antecedent can sometimes be an experiencer or a phrase contained in a subject (cf. Huang and Tang's [1992] notion of *subcommand*).

(3.32) (a) Zhangsan$_1$ shuo **ziji**$_1$ hui lai. (Chinese)
Zhangsan say self will come
'Zhangsan says he will come.'
(b) Zhangsan$_1$ renwei Lisi$_2$ zhidao **ziji**$_{1/2/*3}$ de taitai shi yige
Zhangsan think Lisi know self DE *wife is one*-CL
da hao ren.
big good person
'Zhangsan thought that Lisi knew that his wife was a very good person.'

Subject orientation and domain restriction are not mutually exclusive. The Finnish reflexive *itse*, 'self,' must be bound by a subject within its minimal finite clause:[11]

(3.33) (a) Pekka$_1$ näki että Matti$_2$ katsoi **itseään**$_{2/*1}$ (Finnish)
Pekka saw that Matti watched self-POSS
'Pekka saw that Matti watched himself.'
(b) ∗ Puhuin Pekalle$_1$ **itseään**$_{1/2}$
spoke-1SG *Pekka self*-POSS
'I spoke to Pekka about himself.'

Subject orientation of pronouns is cross-linguistically very pervasive; apart from the cases mentioned here, it is reported to hold for Chinese *taziji*,[12] Czech *si*, Danish *(sig) selv*, Dutch *zich*, French *soi*, Icelandic *(sjálfur) sig*, Italian *sè*, Japanese *zibun*, Kannada *hon* and *ta-nu*, Latin *se*, Malayalam *swa*,[13] Marathi *swataah*, Norwegian *seg self*, Russian *sebja* and *svoj(u)*,[14] and Spanish *se*.

Anti-subject orientation is exemplified by the weak pronoun form *ó* in Yoruba (Kwa, Niger-Congo). This pronoun can be unbound, or bound to a non-subject; it must not, however, be anaphorically related to any commanding subject:[15]

[11] Unless noted otherwise, the data and generalizations in this section and the next are taken from Tang (1989) (Chinese); Toman (1992) (Czech); Vikner (1985) (Danish); Enç (1991) (Dogrib, Turkish); Koster (1984) (Dutch); van Steenbergen (1992) (Finnish); Pica (1983) (French); Everaert (1992) (Frisian); Dalrymple (1993):157 (Fula); Iatridou (1986) (Greek); Mohanan (1990) and Yang (1983) (Hindi); É. Kiss (1992) (Hungarian); Koster and Reuland (1992b) (Icelandic, Italian); Katada (1991) (Japanese); Bhat (1978) (Kannada); Benedicto (1992) (Latin); Yang (1983) and Mohanan (1982) (Malayalam); Dalrymple (1993) (Marathi, Norwegian); Reinders-Machowska (1992) (Polish); Avrutin (1994) (Russian); Fontana and Moore (1992) (Spanish); and Pulleyblank (1986) (Yoruba).

[12] See also Koster and Reuland (1992b).

[13] Latin and Malayalam according to Yang (1983).

[14] Rappaport (1986).

[15] Cf. Pulleyblank (1986) and Pulleyblank (1990); data from Dalrymple (1993), (1.107) and (1.108). Pulleyblank (1990):987 actually provides an example in which *ó*, contained in an adverbial clause, is bound to the matrix subject. The purported generalization is that the anti-subject orientation only holds for pronouns embedded within complement clauses. Further investigation of this is beyond the scope of this book.

(3.34) (a) Ṣẹ́gun$_1$ sọ pé Túndé$_2$ rò pé **ó**$_{3/*1/*2}$ sanra. (Yoruba)
Segun say that Tunde think that he fat
'Segun$_1$ said that Tunde$_2$ thought that he$_{3/*1/*2}$ was fat.'

(b) Tolú$_1$ sọ fún Ṣẹ́gun$_2$ pé Dúpẹ́$_3$ rò pé **ó**$_{2/*1/*3/4}$ sanra
Tolu say to Segun that Dupe think that he fat
'Tolu$_1$ told Segun$_2$ that Dupe$_3$ thought that he$_{2/*1/*3/4}$ was fat.'

Anti-subject orientation – sometimes within a delimited domain – is also found in Danish *ham (selv)*, Dogrib *ye*, Kannada *ava*, Norwegian *ham (self)*, and Russian *ego*.

In the remainder, we will assume that the notion of 'subject' is available to us in formulating Binding Conditions. In theories in which this notion is not a primitive, it will have to be considered as a shorthand for some complex derived notion. Binding Conditions will then take the following form, with an extra parameter added:

(3.35) An NP of **class** must (**not**) be coindexed with a **commanding** $\left\{\begin{array}{l}\textbf{NP}\\ \textbf{Subject}\end{array}\right\}$ within its **domain**.

3.2.2 Logophoricity

In the previous subsection, we have seen pronouns which restrict the class of their possible antecedents by reference to grammatical function (subject vs. non-subject). In this subsection we turn to pronouns, so-called *logophors*, which are oriented towards a semantically or pragmatically determined class of antecedents.

Various languages have a special set of pronouns used to refer to the 'source' of an embedded statement. Consider the following examples from Ewe (Kwa, Niger-Kordofanian):[16]

(3.36) (a) kofi$_1$ be **yè**$_{1/*2/*s}$ -dzo (Ewe)
Kofi say LOG *left*

(b) kofi$_1$ be **e**$_{*1/2/*s}$ -dzo
Kofi say he left

(c) kofi$_1$ be **me**$_{*1/*2/s}$ -dzo
Kofi say I left

'Kofi said that he/I left.'

The distribution of *e*, 'he,' and *me*, 'I,' is more or less as expected: *e* refers to a non-speaker, non-addressee person (hence the subscript $*s$, for 'not the speaker'), while *me* must be the speaker; moreover, *e*, for reasons of no concern here, cannot anaphorically relate to the matrix subject. What is of interest here is the additional pronoun *yè*, which can only refer to the subject of *be*, 'say,' not

[16] From Clements (1975) by way of Sells (1987):448 and Kuno (1987):146.

to any other person, speaker or not. Pronouns like that are called *logophoric*, glossed LOG, following the terminology proposed in Hagège (1974).

Logophoric pronouns are not restricted to verbs of saying. As the following Ewe examples show, the subject of *be happy, know*, or *see* in Ewe can antecede logophoric pronouns; similarly, in Tuburi, the experiencer of a psychological verb can bind a logophor:[17]

(3.37) (a) ana$_1$ kpɔ dyidzo be **yè**$_{1/*2}$ -dyi vi (Ewe)
Ana see happiness COMP LOG *-bear child*
'Ana$_1$ was happy that she$_{1/*2}$ bore a child.'

(b) kofi$_7$ (me-) nya be me -kpɔ **yè**$_{7/*2}$ (o)
Kofi (not) know COMP *I see* LOG
'Kofi$_7$ knew/didn't know that I had seen him$_{7/*2}$.'

(c) kofi$_7$ kpɔ be **yè**wo$_{7+2/*2}$ -do go
Kofi see COMP LOG-PL *-come out*
'Kofi saw that they (including Kofi) had come out.'

(3.38) hí:ní dʒō nē$_1$ gā **sɛ̄**$_{1/*2}$ lĪʔ tʃĪgì (Tuburi)
fear make him COMP LOG *fall illness*
'He was afraid that he would fall ill.'

Before going on, let us briefly consider alternative characterizations of these pronouns. Couldn't they just be reflexives that need to be bound by a higher subject? After all, the 'source of an embedded proposition' often *is* (the referent of the) subject of the verb embedding (the sentence expressing) that proposition. Indeed, logophoricity and long-distance subject orientation aren't always easy to tell apart. There are two straightforward ways to distinguish them, though; consider the following examples involving the Japanese logophoric pronoun *zibun*:[18]

(3.39) (a) Takasi$_1$ wa Taroo$_2$ ni [Yosiko ga **zibun**$_{1/*2}$ o nikundeiru
Takasi TOP *Taroo* DAT *Yosiko* NOM *self* ACC *be-hating*
koto] o hanasita (Japanese)
COMP ACC *told*
'Takasi told Taroo that Yosiko hated him (Takasi).'

(b) Taroo$_2$ wa Takasi$_1$ kara [Yosiko ga **zibun**$_{1/*2}$ to nikundeiru
Taroo TOP *Takasi from Yosiko* NOM *self* ACC *be-hating*
to] kiita
COMP *heard*
'Taroo heard from Takasi that Yosiko hated him (Takasi).'

The only acceptable antecedent for *zibun* in both sentences in (3.39) is *Takasi*. *Takasi* is the subject and topic in (3.39a), but an oblique in (3.39b). If *zibun* were subject oriented, we would expect it to refer to *Taroo* in (3.39b), which has the same grammatical function and morphological marking as *Takasi* in (3.39a). It is clear from these examples that *zibun* is not subject oriented. Rather, it takes the

[17] Ewe data from Sells (1987):449; Tuburi from Hagège (1974).
[18] From Sells (1987):453f.

source of the embedded proposition as its antecedent, which in both sentences in (3.39) is *Takasi*.

A second hallmark of logophoric pronouns is that they can sometimes occur without a sentence internal antecedent at all. (3.40) illustrates this with an example from Icelandic, involving *sér*, the dative of the logophoric pronoun *síg*:[19]

(3.40) Formaðurinn$_1$ varð óskaplega reiður. Tillagan væri avívirðileg.
the chairman became furiously angry the proposal was-SUBJ *outrageous*
Væri henni beint gegn **sér**$_1$ persónulega? (Icelandic)
was-SUBJ *it aimed against self personally*
'The chairman became furiously angry. The proposal was outrageous. Was it aimed at him personally?'

Despite the sentence boundary, the second and third sentence are clearly reporting the chairman's thoughts (note also the subjunctive marking on the verbs in these sentences). As a consequence, the logophoric *sér* can be used here without any sentence internal binder. This again shows that we are not dealing with a reflexive that needs to be syntactically bound, however involved its syntactic binding domain.

Having demonstrated the difference between logophoric and (subject-oriented) reflexive binding, note secondly that logophoric pronouns are different from English non-reflexives, too. They do require an antecedent, and moreover one with a special pragmatic property. While one could argue that pronouns need (discourse) antecedents as well, the difference is clear. Contrast the Ewe example (3.36) above and a random English sentence like *Mary said that she left*. Out of the blue, the latter might tempt speakers strongly towards an interpretation in which *she* is bound by *Mary*, but a disjoint reading, in which *she* is discourse related, is clearly available. This is very different from the case of a logophor, which absolutely needs to take the source of information as its antecedent.

To get a better intuition about what counts as the 'source of information,' note the following paraphrases for the examples above:

(3.41)
(a) Kofi said: "*I* left."
(b) Ana was happy thinking: "*I* am bearing a child."
(c) Kofi knew/didn't know: "X has seen *me*." (where X is the speaker of the sentence)
(d) Kofi saw (something that triggered the mental representation): "*We* have come out."
(e) He was afraid (of being in a state reflected by the mental representation): "*I* am falling ill."
(f) Takasi told Taroo: "Yosiko hates *me*!"
(g) Taroo heard from Takasi: "Yosiko hates *me*!"
(h) The chairman, furiously angry, thinks: "The proposal is outrageous. Is it aimed at *me* personally?!"

[19] From Sigurðsson (1986), via Sells (1987):453.

In all these cases, I have replaced the embedded clause by a direct quotation. Accordingly, the logophoric pronouns of the original sentences are replaced by first person pronouns (in italics). We formulate this as our rule of thumb for logophoric pronouns:

(3.42) A logophoric pronoun can be used if it is embedded in a constituent c such that (i) c is embedded, (ii) c denotes a proposition p, which (iii) can be paraphrased as a mental state or reported utterance of the pronoun's antecedent such that the paraphrase contains a first person pronoun in place of the pronoun.

What exactly qualifies as a logophoric antecedent, however, varies from language to language. Usually some lexical element indicates the presence of a 'logophoric environment,' e.g. a verb of saying, thinking, etc., or a special embedding complementizer. Further conditions may obtain. For example, in Ewe, only sentence embedding verbs license logophors, while verbs like *remember*, which selects an NP complement, or *hear*, with an ECM-type complement, do not (cf. [3.42ii]):[20]

(3.43) (a) ∗ ama$_1$ do ŋku nyɔnuvi hi dze yè$_1$ gbɔ dyi. (Ewe)
Ama set eye girl REL *stay* LOG *side on*
'Ama remembered the girl who stayed with her.'
(b) ama$_1$ gblɔ be yè -do ŋku nyɔnuvi hi dze yè$_1$ gbɔ dyi.
Ama say COMP LOG -*set eye girl* REL *stay* LOG *side on*
'Ama said that she remembered the girl who stayed with her.'

(3.44) (a) ∗Ama remembered: "The girl who stayed with me." (not propositional)
(b) Ama said: "I remember the girl who stayed with me."

Logophoricity is attested in many languages of the world. It is important to keep the option of logophoricity in mind when attempting to describe Binding Conditions in a given language, precisely because it can so easily be mistaken for something else, e.g. long-distance subject-oriented anaphors.

3.2.3 Towards a formal treatment of logophoricity

Sells (1987) provides a formal implementation of logophoricity within the framework of Discourse Representation Theory (Kamp and Reyle, 1993). We will follow the gist, though not the letter, of his proposal in what follows; many details, however, will be omitted.

We take as our role model the treatment of first and second person pronouns from chapter 2, section 2.2.2:

(3.45) (a) $[\![\text{I/me/my/myself}_n]\!]^{g,s,u} = g(n)$ if $g(n) = s$, undefined otherwise
(b) $[\![\text{you/your}_n]\!]^{g,s,u} = g(n)$, if $g(n)$ is the person s addresses in u

[20] Cf. Sells (1987):449f.

Extending this treatment, we introduce another contextual parameter, which we call o for 'origo' (Latin for 'source', given that s is already in use). A logophoric pronoun will always refer to the individual o:

(3.46) $[\![\text{pronoun}_n^{log}]\!]^{g,s,u,o} = o$, if $o = g(n)$

Note that we assume here that logophors, just like first and second person pronouns, are indexed and that their lexical content is just a presupposition.

The origo-parameter can be shifted by verbs of saying, thinking, etc. to the sayer, thinker, etc. (all other aspects of these verbs' meanings are simplified as far as possible):

(3.47)
(a) $[\![\text{say (that) S}]\!]^{g,s,u,o} = \lambda x.x$ says something which entails $[\![\text{S}]\!]^{g,s,u,x}$
(b) $[\![\text{hear from NP (that) S}]\!]^{g,s,u,o} = \lambda x.x$ hears y, $y = [\![\text{NP}]\!]^{g,s,u,o}$, says something which entails $[\![\text{S}]\!]^{g,s,u,y}$
(c) $[\![\text{believe (that) S}]\!]^{g,s,u,o} = \lambda x.$ what x believes entails $[\![\text{S}]\!]^{g,s,u,x}$
(d) $[\![\text{S frightens NP}]\!]^{g,s,u,o} = 1$ iff x, $x = [\![\text{NP}]\!]^{g,s,u,o}$, prefers a state of affairs in which $[\![\text{S}]\!]^{g,s,u,x}$ is false to one in which it is true

Note, in contrast, that the speaker parameter (just like the utterance-situation parameter) cannot be shifted, except for direct quotation. This can be seen, for example, from the fact that, unlike in a direct quotation, English first person pronouns cannot be used to refer to the speaker of an embedded sentence:[21]

(3.48)
(a) $*\text{Gil}_6$ said that I_6 was happy.
(b) Gil_6 said: "I_6 am happy."
(c) Gil_6 said that she_6 was happy.

Rather, we appear to see ordinary coreference in these cases: *she* corefers with *Gil*, who happens to be the source or speaker of the embedded sentence. This is no different from the cases in (3.49), which don't involve reported speech or attitudes in the embedded clauses:

(3.49)
(a) Gil_6 met a pilot who liked her_6.
(b) Gil_6 arrived even though Roger had said that she_6 wouldn't.

According to the treatment in (3.47), re-setting the logophoric center is a lexical property of verbs like *say, believe*, etc. Other sentence-embedding verbs do not have this option, just as little as, say, sentence-embedding prepositions:

(3.50)
(a) $[\![\text{look as if S}]\!]^{g,s,u,o} = \lambda x.$ the visual appearance of x makes it likely that $[\![\text{S}]\!]^{g,s,u,o}$ is true (**not:** ... that $[\![\text{S}]\!]^{g,s,u,\mathbf{x}}$ is true!)
(b) $[\![\text{S unless S'}]\!]^{g,s,u,o} = 1$ iff $[\![\text{S}]\!]^{g,s,u,o}$ is true and will be as long as $[\![\text{S'}]\!]^{g,s,u,o}$ is

It is rather clear why *unless* cannot, in principle, establish a new logophoric center: there is no individual argument to the function denoted by *unless*, therefore

[21] The classical philosophical reference here is Kaplan (1977); see Schlenker (1999) for recent discussion.

a fortiori none that could serve as the new origo. This is different for *look like*, which does have an individual argument. We could thus give it a meaning as indicated in the parentheses in (3.50a), which would wrongly shift the origo to the person whose looks are described.

The intuition here is, of course, that *say* and *think*, but not *look like*, involve the report of an utterance or thought, and thus only they have a source to come along with it. This, however, is expressed nowhere in our formal treatment, and it is not easy to see how it could be.[22]

Exploring this issue further would be beyond the scope of this book. We will leave our formalization at this. Sells (1987) argues that we need in fact not just one origo parameter but three different ones, which he calls 'source,' 'self,' and 'pivot' (plus, of course, the familiar 'external speaker'). As these present nothing fundamentally new, I refer to Sells' work for further details.

There are a number of essential points we leave unaddressed or unresolved: first, the formal counterpart of the 'origo' intuition (i.e. the point just mentioned above); second, cases such as (3.40) above, in which the re-set origo parameter extends beyond the scope of the embedding element (these should follow from a general treatment of *modal subordination*, cf. Roberts (1987, 1989, 1996)); and third, cases in which the origo is not introduced by a specific lexical element, but nonetheless seems to represent a 'source' in the intuitive sense (Sells' *third person point of view* cases; see Sells [1987] for details).

3.3 Binding domains cross-linguistically

In section 3.1 we defined the notion of *governing category* and *coargument domain*, which could serve to identify the binding domain for reflexive (must be bound within) and non-reflexive (mustn't be bound within) pronouns, respectively. Other languages make reference to different domains. It seems, however, that the number of binding domains cross-linguistically might still be rather limited. For the purpose of the presentation to follow, I will focus on four different domains, listed in table 3.1.[23]

It is important to realize that these domains are collapsed in a great number of structures. In a simple transitive clause like (3.51), all four domains for the object NP are the same, namely the root clause:

[22] It may be tempting to try to give the logophoric pronouns a presupposition similar to the first and second person pronouns, e.g. $[\![\text{pronoun}_n^{log}]\!]^{g,s,u,o} = o$, presupposition: o is the source of P. But the question here is: what is P? Unlike u in the definitions above, P is not a contextual parameter. We want it to refer to the reported utterance or thought, but where should this come from?

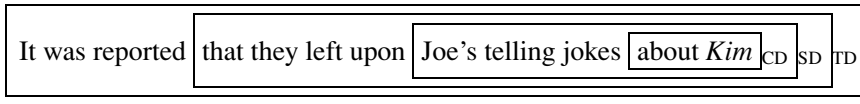
[23] These domains correspond, from bottom to top, to the domains Root S, Minimal Finite Domain, Minimal Complete Nucleus, and Coargument Domain in Dalrymple (1993); and to Domains 3, 2, and 1 in Koster and Reuland (1992b) (who deny the relevance of the C-Domain, a point we will return to later). Interestingly, it is noted in both works that these domains seem to be the only ones required for their quite comprehensive, though not typologically representative, sample of languages.

Table 3.1 *Cross-linguistic binding domains*

Domain	Definition
the . . . for A is	the smallest constituent containing A, A's case assigner C and . . .
Coargument domain, CD	all arguments of C
Subject domain, SD	a Subject (within NP or S)
Tense domain, TD	a finite clause
Root domain, RD	the entire sentence

(3.51) The Schindlers provoked the Flaunders.

Below, on the other hand is a maximally complex example in which the four domains, relative to the NP *Kim*, are different:



To find out which domain is crucial for a given pronoun type thus requires constructing intricate and often somewhat involved sentences. A lot of studies, especially older ones, on Binding Theory in particular languages do not distinguish between the different categories, which is why our cross-linguistic knowledge about these distinctions is very incomplete and likely to be in need of revision.

The root domain

The root domain is significant in English only as a negative domain for full NPs. There are languages, however, in which certain pronouns must be bound in the root domain, though not necessarily locally. Examples of such pronouns mentioned in the literature are Chinese *ziji*, Fula *Dum*, Greek *o idhios* and *ton idhios*, Icelandic *sig*, Italian *sè* and *proprio*, Japanese *zibun*,[24] Kannada *ta-nu*, Latin *se*, Malayalam *swa* and *tanne*, Marathi *aapan*, and Yoruba *ó*. It should be noted that many of these are also subject oriented, and also that the data provided are often not sufficient to rule out the possibility that some of these are actually logophoric.[25] For the remaining core cases, however, the pertinent Binding Condition will take the general form in (3.52), with illustrative examples given in (3.53) and (3.54):[26]

[24] Also Sportiche (1986).

[25] In the same vein, it is not clear in all of these cases that the antecedent actually needs to c-command, or command at all for that matter, the pronoun (this is, for example, not the case in Chinese and Fula).

[26] From Iatridou (1986):769 and Yang (1983):183, respectively; *o idhios* must also be locally free, hence it cannot be bound by *Maria* in (3.53).

(3.52) *o idhios*/*caki* (*ziji*/*zibun*/*swa* . . .) **must** be coindexed with a commanding NP within its **root domain**.

(3.53) O Yanis$_1$ ipe ston Costa$_2$ [oti i Maria$_3$ aghapa **ton idhio**$_{1/2/*3/*4}$]
the Y. *said to-the C.* *that the M.* *loves* *himself*
(Greek)
'Yanis told Costa that Maria loves him.'

(3.54) John$_1$ -in [Bill$_2$ -i [Mary$_3$ -ka [Tom$_4$ -iy **caki**$_{1/2/3/4/*5}$**-e**
J. -TOP *B.* -NOM *M.* -NOM *T.* *-'s* *self*
tæhan thæto] -lɨl silhɔha -n -ta -ko] sængkakha -n
toward attitude -ACC *hate* -ASP -DEC -COMP *think* -ASP
-ta -ko] mit -nin -ta. (Korean)
-DEC -COMP *believe* -ASP -DEC
'John$_1$ believes that Bill$_2$ thinks that Mary$_3$ hates Tom$_4$'s attitude toward self$_{1/2/3/4/*5}$.'

The Korean example aptly illustrates how a Binding Condition like (3.52) is to be read: *caki* can be bound within *any* domain up to the root domain, but not from outside of the root domain (not: in the root domain, but not in any smaller domain). So, generally, if a Binding Condition says that an element needs to be bound in domain D, this means that its antecedent has to be somewhere within D, not necessarily at the 'top' of D. Note, on the other hand, that certain elements may be subject to more than one Binding Condition. On top of (3.52), Greek *o idhios*, for example, also needs to be *free* in its tense domain (which is why it can't be coindexed with *Maria* in [3.53]), so that effectively its binding domain 'starts' at the next higher clause and 'ends' at the root node.

The tense domain

The tense domain functions as the binding domain, for example, for Czech *se*/*sebe* and *svůj*,[27] Danish *sig*, Finnish *hän*, *itse*, and *hän itse*, French *soi*, Marathi *swataah* (in the less restrictive version), Norwegian *seg* and *sin*, Polish *siebie* and *swój*, and Russian *sebja* and *svoj(u)*.[28] Tense-domain-bound reflexives might look like subject-domain-bound reflexives (i.e. like English) at first glance, cf. (3.56a) and (3.57a). What is distinctive, however, is that the pronoun can be bound across subjects in NPs and non-finite complements such as gerunds or infinitival clauses (i.e. the subject domain, marked by parentheses):[29]

(3.55) *swataah*/*seg (sig*/*soi*/*siebie*/*uskii . . .)* **must** be coindexed with a commanding NP within its **tense domain**.

[27] According to Toman (1992):154f., the exact domain seems to be 'inflected clause' (including infinitival clauses), as Czech reflexives cannot be bound across infinite sentence boundaries; they can be bound across small clause and NP subjects though, which is why I group them here.
[28] Also Rappaport (1986).
[29] Examples are Dalrymple's (1993) (1.49), (1.42), (1.43), (1.87), and (1.86).

(3.56) (a) ∗ Tom$_1$ mhanat hota [$_{TD=SD}$ ki Sue ni **swataahlaa**$_1$ maarle]
Tom said that Sue ERG *self hit*
(Marathi)
'Tom said that Sue hit himself.'

(b) [$_{TD}$ Jane$_2$ laa [$_{SD}$ Tom ne **swataaci**$_2$ pustake phekun dilyaace]
Jane DAT *Tom* ERG *self books throw*
kaḷḷe]
learned
'Jane learned about Tom throwing away (her)self's books.'

(c) [$_{TD}$ Jane$_2$ [$_{SD}$ John ne **swataahlaa**$_2$ maarlyaavar] rusun]
Jane John ERG *self hitting angry*
'Jane remained angry upon John hitting (her)self.'

(3.57) (a) ∗ Jon$_1$ var ikke klar over [$_{TD=SD}$ at vi hadde snakket om
Jon was not aware of that we had talked about
seg$_{1/2}$] (Norwegian)
self
'Jon was not aware that we had talked about him.'

(b) Jon$_1$ likte [$_{SD}$ din artikkel om **seg$_1$**]
Jon liked your article about self
'Jon likes your article about him.'

The subject domain

Ignoring the non-complementary cases (i.e. those where a reflexive or reciprocal is contained in a subject), the subject domain is the relevant one for English reflexives. It furthermore appears to be the binding domain for reciprocals in all languages I could find information about on this issue, including Danish *hinanden*,[30] Dutch *elkaar*, English *each other*,[31] Finnish *toiset*, Hungarian *egymas*, Icelandic *hvor annar*, Italian *l'uno, l'altro*, Norwegian *hverandre*, the reciprocal use of Polish *siebie*, and, shown here, Russian *drug druga*:[32]

(3.58) (a) [$_{SD}$ Pisateli$_1$ čitali [vospominanija **drug** o **drug**
writers-NOM *read reminiscences*-ACC *each about other*
-e$_1$]] (Russian)
-LOC
'The writers read reminiscences about each other.'

(b) ∗ Pisateli$_1$ čitali [$_{SD}$ vospominanija Tolstoja **drug**
writers-NOM *read reminiscences*-ACC *Tolstoi*-GEN *each*
o **drug** -e$_1$]
about other -LOC
'The writers read the reminiscences of Tolstoj about each other.'

[30] Pica (1983).
[31] But see Lebeaux (1983).
[32] Rappaport (1986) (18). Yang (1983) actually speculates the SD is the binding domain for reciprocals universally, but see Huang (2000):n. 46, pp. 101f. for possible exceptions.

Cases of (non-reciprocal) pronouns that need to be bound in the subject domain alongside English *herself*/*himself* are Chinese *taziji*,[33] all Danish *-self* pronouns, Dutch *-zelf* pronouns, French *se*, Frisian *'m(sels)*, Greek *ton*,[34] Hungarian *maga*, Icelandic *sjálfur* forms, Italian *si* and *se stesso*, Japanese *-zibun* forms, Marathi *swataah* (in the restricted dialect), Norwegian *-self* forms and possessive *sin*, and Turkish *kendi*:[35]

(3.59) *zichzelf (hvor hanna/toiset/each other/ton/si/swataah . . .)* **must** be coindexed with a commanding NP within its **subject domain**.

(3.60) (a) Peter$_1$ zag [$_{SD}$ Mary$_2$'s foto van **zichzelf**$_{2/*1/*3}$]. (Dutch)
P. saw M.'s pictures of self
'Peter saw Mary's pictures of him/herself.'
(b) Mary$_1$ liet [$_{SD}$ Peter$_2$ op **zichzelf**$_{2/*\ 1/*\ 3}$ schieten]
M. let P. at self shoot
'Mary let Peter shoot at himself.'

The subject domain is also significant as a *negative* domain. Dogrib *ye*, Finnish *hän*, Greek *ton (idhios)*, Italian *lui*, and Turkish *o* need to be free in their subject domain.

Moreover, Danish *ham*/*hende* and *ham selv*, Finnish *hän (itse)*, Icelandic *hann sjálfur*, Malayalam *taan*, Norwegian *ham (self)*, Polish *nim* and possessive *jej*, Russian *ego* must be *subject*-free within the subject domain:

(3.61) *nim (jej/ham (selv)/hän (itse)/ taan . . .)* **must not** be coindexed with a **commanding subject** within its **subject domain**.

(3.62) Piotr$_1$ czyta~~ł~~ [$_{SD}$ Janka$_2$ artyku~~ł~~ o **nim**$_{?1/*\ 2}$] (Polish)
P. read J.'s article about him
'Piotr read Janek's article about him.'

The coargument domain

Turning finally to the coargument domain, it appears that this domain is only relevant as a negative domain. For example, Marathi *to* at first glance behaves like a non-reflexive pronoun in that it cannot be bound in its local clause, but across a finite sentence boundary, cf. (3.63a/3.63b); however, unlike a non-reflexive pronoun in English, it can occur with an antecedent within its subject domain, as long as the antecedent is not a coargument, cf. (3.63c/3.63d):[36]

(3.63) (a) Jane$_1$ ne **tilaa**$_{2/*1}$ bockaarle. (Marathi)
Jane ERG *her*-ACC *scratched*
'Jane scratched her.'
(b) Mary$_2$ dukhi hoti. **tilaa**$_2$ jaataa aale naahi.
Mary sad was she-DAT *go could not*
'Mary was sad. She could not go.'

[33] Koster and Reuland (1992b); Tang (1989).
[34] Also Enç (1989).
[35] Examples are Koster's (1984) (45) and (23).
[36] Examples are Dalrymple's (1993) (1.63), (1.64), (1.60), and (1.61).

(c) Jane$_1$ ne **ticyaakartaa**$_1$ saaḍi gheṭ li.
Jane ERG *her for* *sari* *bought*
'Jane bought a sari for her (Jane).'

(d) Jane$_1$ ne John laa **ticyaabaddal**$_1$ maathiti dili.
Jane ERG *John* DAT *her about* *information gave*
'Jane gave John information about her (Jane).'

Danish *ham* and *hende*, Dutch *zich*, Frisian *'m*, Kannada *ava-*, Marathi *to* and *aapan*, and Norwegian *seg* all need to be free in their coargument domain:

(3.64) *to (hende/zich/ava-...)* **must not** be coindexed with a commanding NP within its **coargument domain**.

It is worth emphasizing that all these elements can be 'locally' bound (e.g. by a non-coargument subject within a finite clause), as in (3.63c/3.63d), and indeed some of them actually *must* be bound within the next higher tense domain (Norwegian), or subject domain (Frisian, Dutch, Danish).

These pronouns have thus a very limited distribution, essentially as arguments to prepositions and inherent-reflexives (cf. chapter 1, section 1.4). In chapter 11 we will explore a possible rationale behind the peculiarities of this domain, namely that all elements that must be free in the coargument domain are simply incapable of marking a predicate as reflexive.

This concludes our cross-linguistic discussion of domains. It is worth stressing that the choice of binding domain is not made once per language (say, tense domain for Marathi and Norwegian, subject domain for English), or even twice (say subject domain as negative domain for non-reflexives, and first clause with an *accessible* subject as positive domain for reflexives, as is sometime suggested for English). Rather, the choice of domain appears to be morpheme-specific, where within the same language different forms can have different binding domains (for example, tense domain as positive domain for Norwegian *seg*, but subject domain for *ham selv* and the reciprocal *hveandre* in the same language). This challenges attractive ideas about setting a single 'domain parameter,' and leaves us with what appears to be a less restrictive approach to the acquisition of binding domains (cf. Manzini and Wexler [1987]; and Wexler and Manzini [1987]).[37]

Exercise 3.3

Indicate all four domains relative to *Y* in the following schematized structures. Provide labeled bracketings for clarity:

[37] There is, of course, the possibility that something about the *form* of a given pronoun, coupled with a universal theory of binding domains, yields the different domains (Déchaine and Wilschko [2000]; Reuland [2001, a.o.]). There are, however, to the best of my knowledge no cross-linguistically valid generalizations about form–domain correspondences that could serve as the basis for such a theory, and, accordingly, such approaches often remain stipulative.

(3.65) (a) X loves Y
(b) Z says that X loves Y
(c) Z left upon X's hitting Y
(d) X has information about Y

Exercise 3.4

(i) Formulate, as far as the following data warrant, the positive and negative binding condition for Greek *ton eafton tou* and *o idhios* (cf. also [3.53]):[38]

(3.66) (a) O Yanis$_1$ aghapa ton eafton tou$_{1/*2}$/ *ton idhio$_{1/2}$.
the Y. loves PRON
(Greek)
'Yanis loves himself.'
(b) O Yanis$_1$ theori oti o idhios$_{1/*2}$ ine o kaliteros
the Y. thinks COMP PRON is the best
ipopsifios.
candidate
'Yanis thinks that he is the best candidate.'
(c) O Yanis$_1$ theli o Costas$_2$ na voithisi ton idhio$_{1/*2}$/ton
the Y. wants the C. helps PRON
eafton tou$_{2/*1}$.
'Yanis wants Costas to help him(self).'
(d) O Yanis$_1$ theori ton eafton tou$_{1/*2}$ ton kaliero
the Y. considers PRON the best
ipopsifio.
candidate
'Yanis considers himself the best candidate.'

(ii) Would you expect that *ton eafton tou* in (3.66d) can be replaced by *o idhios*? Explain!

Exercise 3.5

Complete, as precisely as the data allow, the following Binding Condition for the Hindi pronoun *uskii*:[39]

(3.67) *uskii* must_____be coindexed with a commanding_____in its _____

[38] Data from Iatridou (1986).
[39] Data from Mohanan (1990), via Dalrymple (1993):37f.

(3.68) (a) ravii$_1$ **uskii**$_{*1}$ saikil -par baiṭ^{h}aa (Hindi)
Ravi his bicycle -LOC *sit*-PERF
'Ravi sat on his bike.'

(b) vijay -ne$_1$ ravii -ko$_2$ **uskii**$_{2/*1}$ saikil -par biṭ^{h}aayaa
Vijay -ERG *Ravi* -ACC *his bicycle* -LOC *sit*-CAUSE-PERF
'Vijay seated Ravi on his (Ravi's) bike.'

(c) raajaa -ne$_1$ kahaa ki mantrii$_2$ **uske**$_{1/*2}$ g^{h}ar gayaa
king -ERG *say*-PERF *that minister his house*-LOC *go*-PERF
'The king said that the minister went to his (the king's) house.'

(d) raajaa -ne$_1$ mantrii -ko$_2$ **uske**$_{*1/*2}$ g^{h}ar jaanee -kii
king -ERG *minister* -ACC *his hous*-LOC *go* NONFIN-GEN
aagyaa dii
order give
'The king ordered the minister to go to his (someone else's) house.'

3.4 Long-distance reflexives (LDRs)

There are many languages in which reflexives lead a 'double life': they can be locally bound, similar to English *herself*; or they can find an antecedent outside their minimal clause. In the latter case, that antecedent often needs to be a subject and/or a logophoric center. Such languages include Latin, Icelandic, and Japanese. We will follow the custom in the literature in this section and refer to such pronouns as *long-distance reflexives* (LDRs). For example, the Latin reflexive *se* (acc.)/*sibi* (dat.) can be bound to a non-local antecedent if that antecedent is a verb of saying, cf. (3.69a); we do not of course have ungrammatical examples, but note that in embedded clauses that are not embedded under verbs of saying, only plain pronominal forms are found, cf. (3.69b):[40]

(3.69) (a) Iccius nūntium mittit, nisi subsidium **sibi**$_7$ submittātur ...
Iccius message sends if-not relief REFL *furnished*-PASSIVE (Latin)
'Iccius sends a message that unless relief be given to himself, ...'

(b) Ibi in proximiīs villīs ita bipartītō fuērunt$_{(1)}$, ut
there in nearest farmhouse so in two parts made-they that
Tiberis inter **eōs**$_1$ et pōns interesset.
Tiber between them and bridge lay between
'They set (themselves) up in farmhouses very nearby, divided in two, so that the Tiber and the bridge were in between them.'

There are two general lines of analysis for LDRs: movement analyses, according to which LDRs are locally bound after covert movement into the local clause of their antecedent; and non-movement analyses. Among the latter some analyze

[40] Examples from Kuno (1987):137.

LDRs as command anaphors bound within a certain (rather large) syntactic domain (e.g. Progovac (1992); Manzini and Wexler (1987)), while most claim that LDRs fall outside the domain of sentence grammar and are purely a matter of logophoricity (e.g. Hellan (1992); Kameyama (1984); Kuno (1987); Maling (1984); Thráinsson (1992)).[41]

We do not intend to review the rather extensive literature on the issue here (see e.g. Huang (2000):ch. 2.3 for a critical overview). Generally, proponents of movement accounts claim to offer a unified analysis of the short-distance and long-distance binding cases of LDRs, and stress the parallelism to languages with clitic climbing (it is noteworthy that LDRs are overwhelmingly prosodically weak, monosyllabic forms; see below). Furthermore, these analyses often capture specific restrictions on LDRs connected to things like intervening subjects or mood.

Proponents of non-movement accounts, on the other hand, often point out counter-examples to the restrictions dealt with on the movement analyses. Also, they argue that the existence of logophoric pronouns needs to be acknowledged anyway, and that LDRs can sometimes be found without any sentence-internal antecedent, which is, of course, typical for logophoric pronouns, but finds no natural account under movement approaches. Furthermore they point out that the purported movement of LDRs often needs to violate established restrictions on movement.

It is probably fair to say that the problems for movement accounts are considerable and severe, but that logophoric approaches are only as restrictive as their underlying theory of logophoricity, an area where more work is required.

Maling (1984), Sells (1987), and Thráinsson (1992), among others, make a convincing case that Icelandic LDRs are logophoric in nature. Their data also provide some nice examples of what does and what doesn't count as a logophoric antecedent. To emphasize this latter point, I have provided the direct speech paraphrases below the examples:[42]

(3.70) (a) Hann$_2$ sagði [að **sig**$_2$ vantaði hæfileika]. (Icelandic)
he said that self lacked ability
'He said that he lacked ability.' – 'He said: "I lack ability."'

(b) ∗ Honum$_2$ var sagt [að **sig**$_2$ vantaði hæfileika].
he was said that self lacked ability
'He was told that he lacked ability.' – 'He was told: "You/# I lack ability."'

[41] Among the movement analyses, we can furthermore distinguish between analyses that assume LDRs to move as heads (essentially parallel to clitics in clitic-climbing languages; cf. Cole *et al.* [1990]; Cole and Sung (1994); Cole and Wang (1996); Cole *et al.* [2000]; Hestvik (1992); Pica (1983, 1987)), and those that assume them to move as phrases (Huang and Tang (1992)), as well as mixed ones (e.g. Katada [1991]).

[42] (3.70) is Maling (1984)'s (37a/b), (3.71), and (3.72) are from Sigurðsson (1986), both via Sells (1987):450ff.

(3.71) (a) Barnið$_1$ lét ekki í ljós [að það hefði verið hugsað vel
child-the$_1$ let not in light that there had been though well
um **sig**$_1$]
about self
'The child didn't reveal that it had been taken good care of.'
'The child didn't say: "I've been taken good care of."'

(b) ∗ Barnið$_1$ bar þess ekki merki [að það hefði verið hugsað
child-the$_1$ bore it not signs that there had been though
vel um **sig**$_1$]
well about self
'The child didn't look as if it had been taken good care of.'
#'The child didn't look: "I've been taken good care of."'

As mentioned before, logophoric pronouns can, under the right circumstances, appear without a sentence-internal antecedent at all. This is the case, too, for Icelandic LDRs:

(3.72) Formaðurinn$_4$ varð óskaplega reiður. Tillagan væri avívirðileg.
The chairman became furiously angry. The proposal was outrageous.
Væri henni beint gegn **sér**$_4$ persónulega. (Icelandic)
Was it aimed at self personally?

Because LDRs usually occur in subjunctive, rather than indicative, subordinated sentences (where such a distinction exists), it has been suggested that their binding domain could be syntactically described as 'first dominating clause with independent (=indicative) inflection' (plus, possibly, subject-orientation). While this might be the case for some languages, Thráinsson (1992) shows that subjunctive mood is neither a necessary nor a sufficient condition for LDRs in Icelandic. Rather, the connection appears to be an indirect one: subjunctive mood marks certain embedded contexts (e.g. non-factual ones), which, in turn, coincide to a large degree with those created by verbs of saying and thinking.

A striking generalization about LDRs is that they are morphologically simple. Languages that have only complex reflexives (like English) systematically lack LDRs, and in those that have simple and complex forms (e.g. Icelandic) only the simple ones are found to be LDRs. In some movement approaches to LDRs, this remarkable property has been taken to correlate with the head/phrase distinction: complex forms like *himself* are syntactically branching and thus inherently phrasal, confined to phrasal movement, while simple forms like Icelandic *sig* can act as heads and undergo head-movement (Pica, 1983, 1984). The success of such a story partly relies on how plausible it is to assume that head-movement is less local than phrasal movement (the natural expectation might be that it is the other way around), an issue we won't go into here. Non-movement accounts, on the other hand, have little to offer in the way of explaining the general morphological simplicity of LDRs (either) (see Hellan [1992]:29 for a few speculative remarks).

Exercise 3.6

Explain the grammaticality contrast between the two Icelandic examples in (3.73) below (Maling [1984]'s [29a/b]) by paraphrasing them as direct speech:

(3.73) (a) Jón$_4$ trúir [að hann$_4$ verði alltaf froskur [nema
John believes [that he will be forever a frog [unless
konungsdóttir kysse sig$_4$]]
a king's daughter kisses self]]
'John believes that he will remain a frog forever unless a princess kisses him.'

(b) ∗ því er trúað [að hann$_4$ verði alltaf froskur [nema
it is believed [that he will be forever a frog [unless
konungsdóttir kyssi sig$_4$]]
a king's daughter kisses self]]
'It is believed that he will remain a frog forever unless a princess kisses him.'

3.5 Some pronominal systems

Now that we've seen a broader array of possibilities for domains and orientation, we can have a glance at some complete pronominal systems different from English.

3.5.1 Danish and Norwegian

Many Germanic languages other than English show a four-way split in their pronominal system, cf. table 3.2 (the labels SE- and P-form are taken over from the literature, reminiscent of the French reflexive clitic *se* and the generative term 'pronominal').

Table 3.2 *Germanic pronoun systems*

		SE-form:	P-form:
bare	Danish	sig	ham
	Dutch	zich	hem
	German	sich	ihn
	Norwegian	seg	ham
+'self'	Danish	sig selv	ham selv
	Dutch	zich self	hem self
	German	sich selbst	ihn selbst
	Norwegian	seg selv	ham selv

Table 3.3 *Danish and Norwegian pronoun system*

	SE-form bound to subject in tense domain	P-form free from subject in coargument domain
bare free in c-dom.	D: *sig* N: *seg*	D: *ham, hende* N: *ham*
+'self' bound in s-dom.	D: *sig selv, hende selv* N: *seg selv*	D: *ham selv* N: *ham selv*

Typically, the plain SE-form is found in all non-referring constructions, e.g. inherent reflexives and detransitivized forms such as middles. When referring, SE-forms tend to be locally subject oriented, while P-forms are often locally anti-subject oriented. Bare forms (SE- or P-) are usually locally free, whereas 'self' forms must be locally bound.

I will illustrate this using Danish and Norwegian, which provide particularly neat illustrations of such a system. As suggested in table 3.2, their pronominal systems can be thought of as arranged by two binary choices: SE-form vs. non-reflexive pronoun; and bare form vs. 'self' form. Each value for these choices is associated with one binding requirement. As the choices cross-classify, so do the conditions, as shown in table 3.3.[43]

The examples given below illustrate these systems.[44] (3.74) illustrates the simple case in which tense, subject, and coargument domain coincide. No bare forms can be used, given that these need to be free in the coargument domain, which includes all NPs here. Among the *selv*-forms, *sig*/*seg selv* must be chosen if the antecedent is a subject, but *hende*/*ham selv*, if it is a non-subject:

(3.74) (a) Susan$_1$ fortalte Anne$_2$ om $\left\{\begin{array}{l}*\text{hende}_1/ *\text{hende selv}_1/*\text{sig}_1/ \textbf{sig selv}_1\\ *\text{hende}_2/ \textbf{hende selv}_2/*\text{sig}_2/*\text{sig selv}_2\end{array}\right\}$
Susan told Anne about
(Danish)

(b) Harald$_1$ fortalde Jon$_2$ om $\left\{\begin{array}{l}*\text{ham}_1/ *\text{ham selv}_1/*\text{seg}_1/ \textbf{seg selv}_1\\ *\text{ham}_2/ \textbf{ham selv}_2/*\text{seg}_2/*\text{seg selv}_2\end{array}\right\}$
Harald told John about
(Norwegian)

[43] Note that both 'self' and SE correspond to 'bound in domain D,' while both 'bare' and P correspond to 'free in domain D.' It is tempting to think that +/−SE corresponds to '+/− bound to subject in domain D_1,' and that +/−self corresponds to '+/− bound (at all) in domain D_2.' But note that neither D_1 nor D_2 are the same across positive and negative conditions. Therefore, the formal similarity between the conditions associated with the opposing values cannot be captured in any straightforward way.

[44] I did not differentiate between tense domain and coargument domain for the sake of simplicity; cf. Dalrymple (1993) for evidence for the choice of the latter. Danish examples from Vikner (1985).

Let us now turn to a case where coargument domain and tense domain are distinct. Since *bad*, 'asked,' is an object control verb, the understood subject of the embedded clause in (3.75) is Anne, not Susan. Therefore, coreference with *Susan* is a case of binding outside the coargument domain (and subject domain), but within the tense domain. In such a case, the *selv*-forms are excluded because they require an antecedent within the subject domain. The bare SE-form *sig*/*seg* is possible, because it is free in its coargument domain and at the same time bound to a subject within the tense domain; the bare P-form *hende*/*ham* is possible, too, because it is (subject-)free in the smaller coargument domain:

(3.75) (a) Susan$_1$ bad Anne om at ringe til **hende**$_1$/ ∗hende selv$_1$/
Susan asked Anne for to ring to
sig$_1$/ ∗sig selv$_1$ (Danish)
'Susan asked Anne to call her.'

(b) Jon$_1$ bad oss snakke om **ham**$_1$/ ∗ham selv$_1$/ **seg**$_1$/ ∗seg
Jon asked us to talk about
selv$_1$ (Norwegian)

As expected, *sig*/*seg* will no longer be available if the binding is to a non-subject (note that the understood subject of *ringe* is *Susan*, i.e. the pronoun is indeed bound by the matrix object):

(3.76) Susan$_1$ lovede Anne$_2$ at ringe til **hende**$_2$/ ∗hende selv$_2$/∗sig$_2$/
Susan promised Anne to ring to
∗sig selv$_2$ (Danish)

Note finally that the bare SE-forms (and only those) are used in non-thematic positions such as with inherently reflexive verbs:[45]

(3.77) Peter sov over ∗ ham/ ∗ham selv/ **sig**/ ∗sig selv (Danish)
Peter slept over
'Peter overslept.'

This contrasts with the reflexive object in true transitive constructions, which, as seen in (3.74) above, cannot be bare SE. It looks as if the subject qualifies as a binder in the sense of the SE-vs.-P distinction (otherwise we would expect to see *ham*), but not in the sense of the bare-vs.-'self' distinction (otherwise we would expect *sig selv*). We will return to some of these issues in chapter 11.

[45] As Vikner (1985):8f. notes, certain verbs that allow either thematic or non-thematic objects give the misleading impression that *sig* and *sig selv* have a similar distribution, e.g. in (i):

(i) Peter vaskede sig / sig selv. (Danish)
Peter washed

It turns out, however, that *vaskede* simply has two argument frames, one transitive and one inherently reflexive. Thus *Peter vaskede sig selv* corresponds to English *Peter washed himself*, while *Peter vaskede sig* corresponds to English *Peter washed*. As Vikner shows convincingly, the non-complementarity disappears as soon as one uses verbs that are either always transitive (*sig selv*), or always inherently reflexive (*sig*).

Table 3.4 *Marathi pronoun system*

	bound to	free in
swataah	subject in subject-domain	–
aapaṇ	(logical) subject in root-domain	coargument-domain
to	–	coargument-domain

3.5.2 Marathi

Marathi, as discussed in Dalrymple (1993), has three different pronoun forms, two of which require a binder; cf. table 3.4.[46]

As discussed in section 3.3 above, *to*, like English *her*/*him*, must be locally free, but unlike *her*/*him* (and like bare pronouns in Danish and Norwegian) only in its coargument domain (as opposed to subject domain). *Swataah* plays the role of a reflexive pronoun, except that its domain is slightly larger than that of English, Danish, or Norwegian 'self'-forms (tense domain in Marathi vs. subject domain in the Germanic languages), and it is subject oriented. In addition there is the long-distance reflexive *aapaṇ*, which must be locally free, but bound within the root S;[47] it can thus occur in embedded clauses, within NPs, and as the object of prepositions that assign their own θ-role:[48]

(3.78) (a) Tom$_1$ mhanat hota ki Sue$_2$ ni **aaplyaalaa**$_{1/*2/*3}$ maarle (Marathi)
Tom said that Sue ERG *self*-ACC *hit*
'Tom said that Sue hit him.'

(b) Jane$_1$ ni **aaplye**$_1$ pustak phekun dile
Jane ERG *self*-GEN *book threw give*
'Jane threw away self's book.'

(c) Jane$_1$ ne **aaplyaakartaa**$_1$ saaḍi gheṭ li
Jane ERG *self-for sari bought*
'Jane bought a sari for herself.'

Note once again that the distribution of *aapaṇ* significantly overlaps with that of *swataah* (for all positions whose tense domain is bigger than the coargument domain), and is even a proper subset of that of *to*. This kind of non-complementarity is common among languages: the Turkish pronoun *kendi* is a domain reflexive – it needs to be bound within the root-domain. But Turkish also has a second form *kendisi*, which obeys no Binding Conditions at all. The distribution of *kendi* is thus a proper subset of that of *kendisi*:[49]

[46] The line for *swataah* describes what Dalrymple calls the 'restrictive' dialect of Marathi. For speakers of the less-restricted dialect, the relevant domain is the tense domain.
[47] Dalrymple (1993):21–24 convincingly shows that *aapaṇ* is not a logophor.
[48] Dalrymple (1993)'s (1.49), (1.35), (1.31).
[49] Keenan (1988):134, following Enç (1989).

(3.79) Herkes$_1$ ayna-da $\left\{ \begin{array}{l} \textbf{kendisi-(n)i}_{1/2} \\ \textbf{kendi-(n)i}_{1/*2} \\ \textit{(him)self}\text{-ACC} \end{array} \right\}$ gördu (Turkish)
everyone mirror-LOC *saw*
'Everyone saw themselves in the mirror.'

Even more extreme, Fijian (Oceanic, Austronesian) has no reflexives at all. The Fijian pronoun *koya* occurs in all kinds of configurations, from locally bound, to non-locally bound, to free:[50]

(3.80) a mokuti **koya**$_{1/2}$ o ira kece$_1$ (Fijian)
PAST *hit* *him(self)* PL *all*
'They all hit themselves.'

These are but a few illustrative examples of pronominal systems found in the languages of the world. Our general schema for Binding Conditions at this point looks as in (3.81):

(3.81) An NP of **class** must (**not**) be coindexed with a **commanding** $\left\{ \begin{array}{l} \textbf{NP} \\ \textbf{subject} \end{array} \right\}$ within its $\left\{ \begin{array}{l} \textbf{coargument} \\ \textbf{subject} \\ \textbf{tense} \\ \textbf{root} \end{array} \right\}$ domain.

A **class** can be just one lexical item (with all its different case, person, and number forms, e.g. Marathi *aapan*, Turkish *kendisi*), or a set of (all forms of) stems that have a certain morphological shape in common ('self' vs. bare, SE-based vs. P-based in Germanic), or, of course, their complement (all non-pronominal NPs). Any one of the parameters in (3.81) can serve to define natural super-classes of forms (e.g. all that need to be bound vs. all that need to be free, only the subject-oriented, etc.).

Different languages distinguish different numbers of classes, and whether there is a minimal or a maximal number of classes, whether there are any implicational relations between the different classes cross-linguistically, and whether any super-classes (e.g. Chomsky [1981]'s anaphors versus pronominals) play a cross-linguistically privileged role, remains yet to be found out. What should have become clear, however, is that it is not useful to ask which elements in a given language are the counterparts of, say, reflexive and non-reflexive pronouns in English, or what *the* binding domain (or governing category) in a given language is.

Exercise 3.7

Define the following classes using the schema in (3.81):

1. the class of *domain reflexives*, i.e. all pronouns that must be syntactically bound at all
2. the class of logophors
3. the class of bare SE anaphors in Danish and Norwegian

[50] Keenan (1988):132.

Exercise 3.8

Using table 3.4, predict which of the three pronouns *swataah, aapan*, and *to* can occur in the position of PRONOUN in the Marathi sentence frames in (3.82a) and (3.82b) (Dalrymple [1993] [1.60/61]):

(3.82) (a) Jane$_1$ ne PRONOUN$_1$-FOR saaḍi gheṭ li (Marathi)
Jane ERG *self-for* *sari* *bought*
'Jane bought a sari for herself.'

(b) Jane$_1$ ne John laa PRONOUN$_1$-ABOUT maahiti
Jane ERG *John* DAT *self-about* *information*
dili
gave
'Jane gave John information about herself.'