A reductionist treatment of control and anaphora
Abstract: The main proposal made here is that control and anaphora are not separate grammatical phenomena but involve different interface manifestations of a single underlying anaphoric DP. Specifically, it is hypothesized that obligatory-controlled PRO and overt, “SELF”-anaphors are syntactically-conditioned allomorphs. Given a Y-modular grammar and radical Late Insertion, this entails that obligatorily-controlled PRO and SELF-anaphors are born with identical features and categorial structure and are thus indistinguishable in the narrow syntactic module. Nevertheless, I propose, these elements end up in systematically distinct syntactic positions pre-SpellOut – they are thus distinguishable purely by means of their immediate syntactic context and, in this manner, are contextual allomorphs in the narrow Syntax. The DP that ends up being “PRO” is in canonical subject ([Spec, TP]) position pre-SpellOut; furthermore, the cartographic C-domain that directly embeds this TP is featurally-deficient in a manner that, I propose, characterizes C-layers that embed control complements more generally. The anaphoric DP that becomes a “SELF”-anaphor is in canonical object position and its antecedent is in the same minimal domain as itself. All LF and PF distinctions between OC PRO and SELF-anaphors, I argue, are directly motivated by these contextual distinctions in the narrow syntax.

Keywords: Control, anaphora, allomorphy, features, phases.
1 Introduction

The central thesis of this paper is that control and anaphora are different manifestations of the same underlying grammatical phenomenon. Specifically, I claim that obligatorily-controlled (henceforth, OC) PRO and overt anaphors are syntactically-conditioned allomorphs of one another. In the most abstract sense, this entails that these two elements are completely identical at some deeper, underlying level and that all observed surface (i.e. LF and PF) distinctions develop only at some later stage in the grammatical derivation. Controlled and anaphoric elements are born identical, I propose – specifically as a referentially anaphoric DP whose anaphoricity is characterized by an unvalued R feature. However, they end up in systematically distinct syntactic positions pre-SpellOut, making them contextual allomorphs in the syntax. The sole difference between them at this point in the derivation is their distinct positions in the syntax. I assume a Y-modular grammatical architecture, Late Insertion (Halle and Marantz, 1993) and phase-cyclic SpellOut (Chomsky, 2001) – within such a model, the systematically distinct positions that these elements occupy alone trigger all other differences between these items at the LF and PF interfaces (the so-called “surface” distinctions).

In the course of this paper, I develop the details of the allomorphy thesis for a particular type of PRO – obligatorily-controlled PRO– and for a particular type of overt anaphor, namely the morphologically-complex or “SELF” (to adopt the term from Reinhart and Reuland (1993)) anaphor. The ultimate goal, however, is to extend the allomorphy hypothesis to other elements within the control-anaphora taxonomy such as, on the control side, arbitrary PRO and non-obligatorily controlled PRO and, on the anaphora end, long-distance, morphologically simplex “SE” anaphors.

2 Clarifying the terminology

The term “allomorphy” is a fairly loaded one in the literature, as there are differing views on what types of grammatical phenomena it should cover. So before outlining the details of my proposal, I first clarify what I mean by the term “syntactically-conditioned allomorphy” as applied to the elements OC PRO and SELF-anaphors.
2.1 Different definitions of allomorphy in the literature

For instance, Harley (2006) distinguishes between allomorphy and homosemy. Whereas a particular type of allomorphy, like phonologically-conditioned allomorphy, is “quite general, applies indiscriminately to every phonological word produced by affixation of a particular listeme, and is motivated by the phonotactic rules of the language” what she terms (stem-conditioned) homosemy involves an arbitrary process that cannot be learned from any independent (phonological or syntactico-semantic) rules of grammar; rather, the rules involved in homosemy must be listed and memorized. Thus, the variation with respect to the regular English past-tense – namely, the multiple pronunciations of -ed as /d/ (/k'ænd/), /t/ (/hɪst/), and /@d/ (/wənt@d/) – is an instantiation of (phonologically-conditioned) allomorphy, not homosemy, because it is independently predictable from phonotactic rules of English. In contrast, the irregular past-tense forms created by vowel-ablaut of the verb-root, such as in sink - sank, count as instances of (root-)homosemy because they are unpredictable.

In contrast, Embick (2010) and many others treat both regular and irregular grammatical processes like these as allomorphy: whereas allomorphy with respect to past-tense -ed is phonologically-conditioned and follows from broader principles of English phonology, the allomorphy in the verbal-ablaut case for the English past-tense is root-conditioned and follows from the nature of the root itself. At the other end of the classificatory spectrum, Bonet and Harbour (To appear) argue that only idiosyncratic morphological variation involving irregular multiple exponents of some underlying form constitutes allomorphy. Regular and predictable morphophonological processes such as those involved in the /t/-/d/-/@d/ exponents of English past-tense -ed, they claim, should be not be treated as allomorphic “precisely because regular operations of the phonology are capable of yielding the surface variation from a unique underlying exponent” [p. 7].

2.2 What I mean by the term “allomorphy”

The purpose of this paper is not to contribute in any substantive way to this debate. For the rest of this paper, I will assume in line with Embick (2010), and without further debate, that both
regular and irregular grammatical variation of the kind discussed above constitute allomorphy. In both cases there is a meta-theoretic rule involving a one-many mapping relation between an abstract underlying feature (or feature-bundle) and multiple morphophonological variants, as shown below (this diagram is taken from Bonet and Harbour, To appear):

\[
[F] \Leftrightarrow \begin{cases} 
\phi_1 & \text{Context}_1 \\
\phi_2 & \text{Context}_2 \\
\vdots & \vdots 
\end{cases}
\]

This is the broad sense in which I use the term “allomorphy” – nothing substantive hinges on this and this should therefore be treated purely as a terminological choice.

2.2.1 Syntactically-conditioned allomorphy

My use of the term “syntactically conditioned allomorphy” should be understood in a parallel manner to phonologically-conditioned allomorphy in such a system. As such, my underlying feature/feature-structure [F] is a syntactic one which the narrow syntactic module is equipped to interpret. The conditioning environments are syntactic: they represent structural positions, such as head, specifier, and complement, in the syntactic tree. The rules that affect the underlying abstract feature(-bundle) [F] in these positions are rules that are held to independently apply in the Narrow Syntax in a standard Minimalist (Chomsky, 2001, and subsequent) framework: phase-locality and boundedness/impenetrability, Agree and (phasally) cyclic Spell-Out are the syntactic operations that I make crucial use of in my proposal.

The (putative, as per my claim) syntactically-conditioned allomorphy between OC PRO and SELF-anaphors is intended to capture, at the most general level, the following properties. First, it indicates that there is a one-many mapping relation between a single underlying DP with abstract feature structure and two distinct exponents of this DP, namely a silent “PRO” and a morphologically complex “SELF”-anaphor. I will propose that the underlying DP that yields both these surface forms has an unvalued [R] feature which marks referential deficiency/anaphoricity in the narrow syntax. This is shown diagrammatically below (instantiating the more general meta-theoretic rule above):
2.2.2 Other potential issues

There are other ways in which such a model might be held to deviate from standard notions of allomorphy. Standard allomorphy considers conditioning environments to be those that are immediately local, typically defined in terms of (hierarchical or linear) adjacency, to the abstract morpheme in question. The syntactic conditioning environments that I assume for these exponents above, however, might not all be “local” enough to the \( DP_{[R:]} \) feature-structure in this sense. It seems like sisterhood should be an instantiation of an immediately local relationship in the syntactic sense and sisterhood does play a role as a conditioning environment in my proposal – the syntactic conditioning context for exponence as a SELF-anaphor is sisterhood to \( v \) (i.e. a canonical direct-object position). At the same time, and as will become clearer with explication, broader notions of locality like phase-locality also (though perhaps only indirectly, due to Agree) influence exponence in my proposal. It is nevertheless possible that a phasal domain instantiates a conditioning environment that is not local enough in the allomorphic contextual sense, in which case a different term should be used to describe the model at hand. Since it is not a priori entirely clear what should count as local in the syntactic sense, I leave aside this issue for the time being.

There is one final clarificational point to be made and this is perhaps the most important one in that it is substantive and not just terminological in nature. As I describe in detail later, what ultimately feeds distinct SpellOut rules (as a silent PRO on the one hand, and an overt SELF-form, on the other) for these items, is that in the “SELF-anaphor scenario” \( DP_{[R:]} \) is able to Agree with its antecedent DP prior to SpellOut whereas in the PRO-scenario, it hasn’t checked its unvalued R feature against its antecedent DP at SpellOut. In other words, the abstract syntactic feature-bundle that is held to represent both these exponents underlingly – \( DP_{[R:]} \) – arguably gets differentiated even prior to SpellOut/exponence, which is potentially at odds with what is diagrammatically depicted above. Again, I do not say much more about this
but it is, nevertheless, important to keep in mind while considering the rest of this allomorphic hypothesis.

3 Returning to the proposal: theoretical motivations

As has been discussed in detail elsewhere (Landau, 2004; Hornstein, 1999; Darzi, 2008), there are compelling theoretical reasons to get rid of PRO and overt anaphors as separate grammatical primitives. Here, I address the theoretical redundancy and problems involved in the treatment of PRO and overt anaphors in turn.

3.1 The problem with a primitive element “PRO”

In the GB era, it was seen that OC PRO seemed to behave like an anaphor in some ways and like a pronominal in others. Thus, in the sentences in (1) and below, PRO may not be bound by a local antecedent (2); its referential antecedent must be strictly anti-local (1):

(1) John$_i$ wanted [PRO$_i$] to eat

(2) * John$_i$ wanted PRO$_i$

On the other hand, OC PRO behaved like anaphors in that it was obligatorily coreferential with its antecedent and could not refer to a salient entity in the conversational discourse (the way that a deictic pronoun could, for instance). To resolve this dilemma, Chomsky (1981, 1986) and others proposed that OC PRO was lexically specified as [+pronominal] and [+anaphoric] – a mutually inconsistent feature-combination that could only be vacuously satisfied by its being ungoverned (in [Spec, T$_{nonfin}$]). The traditional version of the GB-style Case filter (Chomsky, 1980) claimed that overt N/DPs must be Case-marked. Since a N/DP in ungoverned position could, by definition, not be Case-marked, it thus could not also be overt. Thus, as Hornstein (1999) points out therefore, the conception of PRO as a pronominal anaphor that could not be governed also served to explain its phonetic silence.

Although descriptively elegant, this theory was not explanatory in any way; it didn’t explain how a nominal element could come to simultaneously possess two apparently mutually incompatible features. It was also redundant in that it required an entire grammatical apparatus,
namely the control module, to account for the distribution of a single element, PRO. With the advent of Minimalism, the problems got even more complicated. First of all, the notion of government was dispensed with, so the distribution of PRO could no longer be defined in terms of it (Martin, 2001). And the Minimalist equivalent of a syntactic workspace, namely the phase, was not equivalent to a governing category and could not trivially be used to replace it. Second, and perhaps even more problematic, features such as [+anaphoric] and [+pronominal] could no longer be treated as primitives (see Reuland, 2001, and others) but had to be derived from even more primitive notions such as φ-feature composition and the nature of the syntactic environment. In Minimalism (Martin, 2001; Chomsky, 2001), the distribution of PRO is governed by the following one-one mapping relation:

\[
\text{NULL CASE} \leftrightarrow \text{PRO}
\]

In addition to being a theoretically costly move (see also works in the Hornstein, 1999, traditions), it also faced serious empirical challenges. Landau (2003); Szabolcsi (2009); Sundaresan and McFadden (2009) and many others have shown that elements other than OC PRO may occupy the subject position of untensed/nonfinite clauses in languages like Hungarian, Hebrew, Tamil, Malayalam, Sinhala and perhaps even in gerundival constructions in Modern English. Sigurðsson (2004) also presents evidence from floating quantifiers in Icelandic which strongly suggest that OC PRO is capable of bearing (non-null) Case. Finally, Darzi (2008) and others present evidence of “finite” control in Persian. All these types of evidence strongly undermine the bijectionary mapping between null Case and OC PRO, shown above.

A major counter-proposal against the Minimalist null Case hypothesis for PRO has been that of Hornstein (1999); Boeckx and Hornstein (2003) and others in the Hornsteinian tradition, which propose that control should be treated as identical to A-movement. While this is appealing in its theoretical parsimony, its empirical coverage has been called into question by a series of papers as in Landau (2004) and subsequent, which show that OC PRO and A-traces show systematically distinct empirical behavior. Bobaljik and Landau (2009), for instance, present strong counter-evidence from Icelandic showing that infinitival complements of control-predicates constitute a separate structural Case-assignment domain in contrast to clauses involving A-traces like raising complements which show some manner of Case-dependence with the
matrix clause. Also, as Landau (2003) shows, it is unclear how other properties like partial control can be accommodated under a movement analysis. In a slightly different vein, Culicover and Jackendoff (2001) present strong empirical evidence to demonstrate that the nuanced semantics seen in control environments cannot be easily explained under an A-movement-type analysis. The empirical facts and theoretical arguments against a movement analysis of control are robust and well-documented; thus, I don’t say anything more about this here but assume, in line with Bobaljik and Landau (2009) and others, that such a theory is on the wrong track.

To summarize then, the issue of control has always been a problem in generative linguistic traditions. There is thus ample theoretical motivation to get rid of PRO as a grammatical primitive and derive its empirical properties from other independent principles of the grammar. It is the hope that an allomorphy hypothesis between OC PRO and SELF-anaphors will be the first step towards such a goal.

### 3.2 The problem with a primitive “SELF” or “SE”

The GB Binding Conditions delimited the surface distribution of nominals in terms of the governing category (a formalized syntactic work-space) for each, as follows:

**THE BINDING CONDITIONS (CHOMSKY, 1981)**

(A) An anaphor [+anaphoric] is bound in its governing category.

(B) A pronominal [+pronominal] is free in its governing category.

(C) An R-expression is free.

This seemed to effectively capture the distribution of nominals in structures like the following:

(3) John<sub>i</sub> punched himself<sub>{i,∗j}</sub> (CONDITION A)

(4) John<sub>i</sub> punched him<sub>{∗i,j}</sub> (CONDITION B)

(5) John<sub>i</sub> told Bill<sub>j</sub> that [CP he<sub>{i,j,k}</sub> was stupid ] (CONDITION B)

---

1Chomsky (1981, p.211) defines a governing category as follows:

β is a governing category for α if and only if β is the minimal category containing α, a governor of α, and a SUBJECT accessible to α.
(6) John, told Bill, that \( [C_P \text{ Alfred}_{i,j,k} \text{ was stupid }] \) (CONDITION C)

Nevertheless, there were both theoretical and empirical problems with this treatment of binding. On the theoretical side, the problem was that the Binding Conditions were essentially a formalized description of the data and didn’t have much explanatory power – a point that has been noted elsewhere in the literature (see, for instance Reuland, 2001; Hicks, 2009, and others). On the empirical level, too, there were problems, as discussed at length in Reinhart and Reuland (1993). First, the Binding Conditions above don’t seem equipped to handle the distinction between SELF- and SE-anaphors on the one hand, and that between SE-anaphors and pronouns on the other. SE-anaphors seem to exhibit properties of both SELF-anaphors and deictic pronouns. On the one hand, they must be syntactico-semantically bound within a certain domain, just like SELF-anaphors (consider (7) from Dutch); on the other hand, they contrast with SELF-anaphors and resemble deictic pronouns in that they can be (in many languages, must be) long-distance bound – i.e. they show a Condition B effect (both (7) and (8)).

(7) Jan haat zichzelf/*zich
   Jan hated SELF/*SE/him
   “Jan, hated himself.”

(8) Jan zag jou achter zich/hem staan
   Jan saw you behind SE/him stand
   “Jan, saw you standing behind him.”

To deal with such problems, Reinhart and Reuland (1993) devise a tripartite (rather than a binary) feature-classification of nominals into SELF (-R(eferential), +Reflexive), SE (-R, -Reflexive), and deictic pronouns (+R, -Reflexive). They also propose the following revised set of Binding Conditions for these types of nominals:

CONDITION A: A reflexive-marked predicate is reflexive.

CONDITION B: A reflexive predicate is reflexive-marked.

Nevertheless, and despite the strong contribution Reinhart and Reuland (1993) made, it was still only a description (albeit a very good one) of the binding facts and did not serve to explain the reasons behind these facts (as Reuland, 2001, concedes).
With the advent of Minimalism (Chomsky, 1993, and subsequent), the status of binding became more precarious. First, the notion of a governing category was dispensed with, and the Minimalist phase could not readily replace it in all cases. Second, with the move toward a more derivational and less representational grammar model, it was unclear where in the derivation binding was supposed to take place – in D-Structure, S-Structure, or perhaps later. Yet it is clear that different types of binding (binding of pronouns, for instance) seem to occur at different stages of the derivation: Hicks (2009) provides evidence from heavy NP-shift and other factors to show that some Condition B effects are sensitive to morphophonological factors. On the other hand, other kinds of binding seem to be purely semantic/LF-based (see also Lebeaux, 1998, for an earlier in-depth analysis of such differences). I do not enter into a detailed discussion of the relevant facts here – the purpose of this brief discussion is merely to show that pronominal binding and the status of anaphors, bound- and free/deictic pronouns is still, in effect, an open question in the Minimalist Program – a point that has been noted recently in many research papers, among them: Reuland (2001); Heinat (2008); Hicks (2009) and Kratzer (2009).

### 3.3 Summary and previous reductionist attempts

The object of the preceding discussions in this section has been to show that the statuses of control and binding are both independently fraught with theoretical and empirical problems.

Interestingly, the idea that obligatory control and binding are equivalent is not an entirely new one. Thus, Manzini (1983) proposes a reductionist treatment wherein: “control theory is constructed on essentially the same notions on which binding theory is constructed in Chomsky (1981)” [p.1]. Similarly, Bouchard (1985), building on the more detailed dissertation Bouchard (1984), claims that: “In Chomsky (1981) and related work it is generally assumed that PRO is a pronominal anaphor. In Bouchard (1984) I depart from this position and assume that PRO is either a pronominal or an anaphor, but never both at the same time. I will show that these two PROs respectively behave like pronouns and anaphors in constructions where it has been independently motivated that pronouns and anaphors clearly differ in behavior.” Though differing in many details, similar reductionist theses attempting to reduce binding and control to a sin-

---

2I thank Idan Landau (p.c.) for bringing the significance of this point to my attention.
Single core phenomenon have been advanced by Koster (1984) (though this also attempts to bring movement and A-traces into the mix), Borer (1989), Clark (1990) and others. The reductionist proposal made here is thus not unique in a broader sense. At the same time, I believe that the current status and properties of Minimalist theory allows us to remodel this discussion within an allomorphic hypothesis such as that proposed here. Specifically, we are now in a position to make the strong hypothesis that the DPs corresponding to the surface forms of OC PRO and SELF-anaphors are featurally one and the same (despite their patent surface differences) – a claim that simply would not be possible without the fundamental grammatical premises of Late Insertion (creating a fundamental and systematized divide between underlying structure and surface exponence) and Y-modularity.

4 Empirical motivations for the proposal

Empirically too, there are reasons to believe that a syntactically-conditioned allomorphy hypothesis for controlled PRO and SELF-anaphors is on the right track. First of all, both PRO and anaphors are referentially deficient – a point that all proposals, however much they may vary in the details of implementation or other assumptions (for instance, the contrasting proposals in the Hornstein, 1999; Landau, 2004, traditions) unanimously agree on. What this means is that these elements have no inherent reference but must acquire this from some other antecedent in the syntax or logophoric context. An important point to keep in mind is that the type of referential defectiveness exhibited by these items is different from that seen with deictic pronouns. Although deictic pronouns are also referentially deficient, they are different from anaphors and PRO in that they can refer to entities in the salient discourse in a way that anaphors, PRO and even logophors cannot. This distinction between deictic pronominals and anaphoric ones is captured in the literature in many ways: Reinhart and Reuland (1993), for instance, endowed both SELF- and SE-anaphors with the feature [-R] but deictic pronouns with the feature [+R]

---

3A logophoric context is a special type of salient context where the SELF- or SE-anaphor seems to refer to an implicit author or narrator (see Reinhart and Reuland, 1993; Anand, 2006; Schlenker, 1999; Giorgi, 2006, and many others for discussion of this phenomenon).
to capture precisely this distinction. Heim and Kratzer (1998) distinguish between these elements by means of making them be semantically bound in different ways: deictic pronouns are accidentally coreferent with their antecedents. The reference resolution is implemented by means of an assignment function g which maps the pronominal variable to a salient entity in the discourse context or larger syntactic environment at LF; bound variables, on the other hand, are bound by means of λ-abstraction within the syntactico-semantic structure itself. In more recent papers (Heim, 2008; Kratzer, 2009), Heim and Kratzer have both argued that PRO and anaphors should be treated as bound variables and that deictic pronouns should be treated as accidentally coreferent. From now on, I will simply call the referential defectiveness of PRO and overt-anaphors “anaphoricity” to distinguish it from the referential defectiveness of deictic pronouns.

Having established this, I now proceed to describe in more detail the syntactic, semantic, and morphophonological properties of OC PRO and SELF-anaphors, which provide further empirical motivation for the allomorphy hypothesis proposed here.

4.1 Syntactic properties of PRO and SELF-anaphors

The other significant motivation for the allomorphy hypothesis between these nominals is their complementary distribution in the syntax. Simply put, SELF-anaphors must be in a syntactically local relationship with their antecedents (descriptively, Condition A); SE-anaphors can, often must, be in a syntactically non-local relationship with their antecedents (the Condition B effect of SE-anaphors, noted above). OC PRO, on the other hand, must be in a non-local relationship with its antecedent (descriptively, that PRO must be ungoverned). I discuss the details of these points, though for now only with respect to SELF-anaphors and OC PRO. Complementary distribution is one of the classic tests for allomorphy between two elements: the basic idea is that two underlyingly identical (in this case, identically anaphoric) elements are differentiated solely in terms of their syntactic positions/environment. Additionally, I show that control

---

4Interestingly, Reinhart and Reuland (1993) gave PRO the feature [+R] – though they don’t discuss PRO that much in the paper. However, Landau (2004) points out that there is good reason to treat PRO as [-R] as well – a point that I agree with him on.
(for the most part) requires minimality whereas binding requires locality. This property can be related to the more descriptive distributional pattern that OC PRO occupies a canonical subject position (in [Spec, TP]) whereas locally-bound, self-anaphors occur in canonical object position.

### 4.1.1 Near-complementary distribution

Consider the following examples in English:

1. **Simplex transitive**
   
   Bill, hit himself, */PRO*

2. **Control complement**
   
   John, tried \[CP PRO to eat rice.\]

The structures in (9) show that the self-anaphor must occur in a syntactically local relationship with its antecedent (descriptively, this is Binding Condition A). The grammaticality of PRO in (10) and its ungrammaticality in (9) displays conversely, that PRO must appear in a syntactically non-local relationship with its antecedent (descriptively, this is equivalent to the notion that PRO cannot be governed (Chomsky, 1981, and others)). The complementarity of OC PRO and self-anaphors falls out of these distributional constraints on each.\(^5\)

\(^5\)This is a non-trivial caveat, as Idan Landau (p.c.) points out. In other words, he claims that the “promise”-type putative exceptions to minimal OC are numerous enough to ultimately not count as mere exceptions. Thus, he points out, a more nuanced approach is called for which takes the individual syntactico-semantics of the embedding control predicate into account (see also Culicover and Jackendoff, 2001, for a similar view). Nevertheless, it is undisputed that the most restrictive type of OC, characterized by complements embedded under predicates like try and begin, does impose a strict minimality restriction on the controller DP. As far as I know, similar systematic restrictions are not found in binding constructions on the whole. If anything, the predominant pattern seen with long-distance anaphors is their requirement of a subject antecedent – which often has the syntactic effect of anti-minimality (Koster and Reuland, 1991). I leave this issue for further research.

\(^6\)There is, however, one class of clauses (complements of want-class verbs) where controlled pro and self-anaphors (as well as overt non-anaphoric DP-subjects) occur in apparent free-variation.

i. Susan, wants PRO \(_{i+j}\) to drink a Hefeweizen.

ii. Susan, wants herself \(_i\) to drink a Hefeweizen.

This seems to pose a challenge to the idea of complementarity between OC PRO and self-anaphors described here. But there are reasons to believe that the structures with self-anaphors are underlingly different from those with PRO (see Bošković, 1997; Moulton, 2005, for more).
4.1.2 Subjecthood vs. objecthood: minimality vs. locality

OC PRO and SELF-anaphors also differ from one another with respect to whether they are in canonical subject or object position in the syntactic structure. OC PRO is only found in canonical subject position (i.e. in [Spec, TP]), not in object position:

This is shown below:

(11) Susan\(_i\) tried \([\text{PRO}_{i/j} \text{ to drink a keg of beer}]\)
(12) * Susan\(_i\) slapped \text{PRO}_{i/j}

Overt anaphors can be in canonical subject position but this only happens in cases of long-distance binding which, typically furthermore, involve morphologically simplex (or “SE”) anaphors. Locally bound, morphologically complex SELF-anaphors of the kind under investigation here, in contrast to both SE-anaphors and OC PRO occur in non-subject positions, as shown below:

(13) John\(_i\) hit himself\(_i\).
(14) * Himself\(_{i/j}\) hit John\(_i\).

Standard control- and reflexive-binding structures also differ with respect to the properties of their controller and binder antecedents, respectively. Crosslinguistically, controllers are designated; binders are not Lasnik (1992). (15) below shows for Tamil that there is only one designated controller for OC PRO in a given structure – typically the one closest to PRO – yielding, in most cases, object- over subject-control.

(15) raman krishnan.ai \([CP \text{ PRO } \text{viit[w.ukku poog.a ] so.nn.aan} \text{ PRO } \text{home.DAT go.INF tell.PST.3MSG} \] \\
raman.NOM krishnan.ACC PRO home.DAT go.INF tell.PST.3MSG \\
“Raman\(_i\) told Krishnan\(_j\) [\text{PRO}_{j/4} \text{ to go home}]\)”

In contrast, (16) shows that more than one antecedent can potentially serve as a binder for the SELF-anaphor:

(16) Max\(_i\) told Josh\(_j\) about himself\(_{i/j}\)

The descriptive generalization here seems to be that controllers for PRO must be non-local but minimal (Rizzi, 1990). Binders for reflexives must be local but need not be minimal. Later
in this paper, I discuss a plausible way to explain these differences as well as the canonical-subjecthood vs. objecthood statuses of OC PRO and locally-bound SELF-anaphors, respectively, in terms of intervention effects and phase boundaries.

4.2 Semantic properties: anaphoricity

Here I discuss two different interpretive properties associated with OC PRO and SELF-anaphors. One is their sloppy-identity behavior under VP-ellipsis which suggests that they are both bound variables. The other is the phenomenon of *de se* effects – obligatory in the case of (subject-controlled) PRO but non-obligatory in the case of SELF-anaphors. However, I show that this systematic distinction is not actually problematic for a view that PRO and SELF-anaphors are underlingly/featurally identical.

4.2.1 VP ellipsis: sloppy and strict effects

There is evidence from sloppy ellipsis effects, among others, that PRO and SELF-anaphors are both bound variables and are not merely accidentally coreferent with their antecedents (Heim and Kratzer, 1998).\(^7\) Consider the examples below which present the sloppy-identity effects for both OC PRO and SELF-anaphors:\(^8\)

(17) Jason\(_i\) saw himself\(_i\) and Martin\(_j\) did too saw himself\(_j/\ast\)

(18) Mary\(_i\) tried \[CP PRO\(_i\) to drink beer\] and Susan\(_j\) did too tried \[CP PRO\(_j\) to drink beer\]

This interpretive parallel between OC PRO and SELF-anaphors is another motivation for the proposal that OC PRO and SELF-anaphors are born identical.

\(^7\)But see Giorgi (2006) for a dissenting view.

\(^8\)Ken Safir (p.c.) claims that structures like (i) show that SELF-anaphors can sometimes yield strict readings under VP-ellipsis:

i. Bush considers himself above the law, but we certainly don’t. (STRICT AND SLOPPY)

I agree with Ken’s judgments but don’t say anything more about this at the moment, especially in the view of some proposals (Heim, 1993) that certain bound variables can indeed be involved in strict readings and given that additional pragmatic factors seem to be able to affect interpretation.
4.2.2 *De se effects with PRO and SELF-anaphors*

Nevertheless, there is a serious empirical challenge to the allomorphy hypothesis in terms of the *de se* effects associated with each. Simply put, OC PRO is obligatorily *de se* when it is controlled by a matrix subject (or *de te* in the case of object control (Anand, 2006, for more)) whereas SELF-anaphors, while compatible with a *de se* interpretation, do not seem to be obligatorily *de se*. Thus we have distinctions like the following:

(19) John$_i$ wanted PRO$_i$ to eat pizza (obligatorily *de se*)
(20) John$_i$ wanted himself$_i$ to eat pizza (*de se*; *de re*)
(21) John$_i$ suspected himself$_i$ (*de se*; *de re*)

These systematic distinctions potentially challenge the idea that PRO and SELF-anaphors are underlyingly (i.e. featurally) identical elements. However, there is reason to believe that such obligatory *de se* effects as are associated with PRO are not due to anything inherent in PRO (let’s call this the “inherence hypothesis”) but are, rather, a direct function of their syntactic environment (the “contextualist hypothesis”)– namely as the subject of a TP that is embedded under a feature-defective (characterizing “control” environments) C-layer.

To see this, consider the following crosslinguistically robust generalizations:

- Only subject-controlled PRO is obligatorily *de se*; object-controlled PRO is not. The inherence hypothesis would be forced to deal with this by proposing that there are two different PRO elements involved in each. However, it is much more elegant and explanatory to propose that this distinction is directly because the same PRO element is controlled by different DPs in each case.

- SE-anaphors in subject-position are also obligatorily *de se*. The Italian sentence in (22) from Chierchia (1989) illustrates this:

(22) Pavarotti crede che i propri pantaloni siano in flamme

Pavarotti believes that the self pants are in flame

“Pavarotti$_i$ believes that his$_i$ pants are on fire.”

(Obligatorily *de se* = ‘Pavarotti believes, “My pants are on fire.”’)
Crucially, SE-anaphors pattern like subject-controlled PRO in that they take subject, not object, antecedents (Reinhart and Reuland, 1993; Reuland, 2001, and many others) – suggesting that it is the syntactic context (non-local subject-antecedence and position in embedded [Spec, TP]) that is involved.

- In languages like Hungarian, European and Brazilian Portuguese (Szabolcsi, 2009; Barbosa, 2009) and Tamil (Sundaresan, To Appear), the subject of a nonfinite clause may be overt. To explicate this point further, consider the following Hungarian example taken from Szabolcsi (2009) (formatting mine):

  (23) senki nem akart csak ö leü l-ni
  nobody not wanted.3SG only he/she sit-INF
  “Nobody wanted it to be the case that only he/she takes a seat.”

  CONTEXT: A group of friends are in a crowded bus and there is only one available seat.

Crucially, the complement of the matrix control verb akart (wanted) is an infinitival clause (overtly marked as such by the infinitival suffix ni on the embedded verb). However, the embedded subject is overt and is, furthermore, obligatorily controlled by the matrix subject senki (nobody); Szabolcsi states that this pronoun “acts as a variable bound by the matrix subject; moreover it has the same de se interpretation that [subject-]controlled PRO classically receives” [p. 2]. This again supports a contextualist hypothesis for de se effects.

- Even subject-controlled PRO is not always interpreted obligatorily de se. Thus, in the following example (24), subject-controlled PRO is not interpreted de se:

  (24) Bill’s agent, neglected [PRO, to inform Bill of his upcoming promotional meeting.]

Other embedding control predicates like forget, and under certain circumstances, manage also seem to behave similarly. To the extent that the de se effect of OC PRO is cancellable by the syntactico-semantics of the embedding verb, this also supports a contextualism
hypothesis for de se.\textsuperscript{9}

I take these arguments to show that the de se effect of subject-controlled PRO must be taken to be independent of the inherent nature of PRO itself – rather, it must follow from independent facts about its syntactico-semantic environment.

4.3 Morphological properties

OC PRO and SELF-anaphors appear to make a different number of morphological distinctions, at least on the surface. But as I show here, this is not actually problematic for the idea that they are underlyingly featurally identical.\textsuperscript{10} With respect to their surface morphophonology, SELF-anaphors can show full (e.g. English) or partial (e.g. Japanese, Malayalam, Norwegian) surface marking for person, number, and gender. But PRO, being silent, trivially lacks all surface marking to differentiate these categories, perhaps suggesting that it is lacking the relevant features. That is, SELF-anaphors and PRO have distinct sets of surface shapes – suggesting that they may not be underlyingly identical after all. However, there is strong empirical evidence to suggest that these elements might be underlyingly much more similar (perhaps even identical) than their surface forms suggest. In (25) below, from Icelandic (from Sigurðsson, 2008), the floating quantifier báðir is marked nominative, masculine, and plural - reflecting the case, gender, and number the nonfinite subject (here, PRO) would bear if it were overt:

(25) Bræðrunum likaði illa [að PRO vera ekki báðir kosnir]  
brothers.D.M.PL liked ill [to PRO be not both.N.M.PL elected]

‘The brothers\textsubscript{i} disliked [\textit{CP PRO, not being both elected}]’

Crucially, the nominative case on this quantifier cannot have been inherited from the matrix DP which has a different Case-value (a quirky dative). Thus, Sigurðsson (1991, 2008) argues that the other (= φ-)features on the floating quantifier haven’t been inherited from the matrix DP but must be associated with the features of its clausemate, nonfinite subject, PRO.\textsuperscript{11}

\textsuperscript{9}I am very grateful to Ken Safir (p.c.) for alerting me to this data.

\textsuperscript{10}Note: feature identity does not refer to identity for feature-values but for feature-attributes. I.e. the claim here is not that OC PRO and SELF-anaphors both have feature values: 3rd, singular, and masculine, say, but that they both have feature attributes for person, number, and gender.
In (26) below, the controlled PRO subject of the control clause (strong CP-phase) binds the overt SELF-anaphor *himself*.\(^{11}\) We would only expect this to be possible if the binder PRO were at least as \(\phi\)-featurally specified as its bindee *himself* (at the point of binding, anyway).

(26) Ezekiel\(_i\) tried \([\text{CP PRO}_i \text{ to kill himself}_i]\]

Data like these strongly suggest that PRO can bear case, gender and number features. This would mean that the lack of surface agreement for PRO is not due to its underlying lack of \(\phi\)-features but due to its surface, phonological silence (which must be due to independent factors). It crucially also leaves open the possibility for the following strong hypothesis, namely that PRO and SELF-anaphors, while distinct on the surface, might be featurally fully identical underlingly – especially given their similarity in other areas.

### 4.3.1 Feature-composition of an anaphoric DP (DP\(_{\text{anaph}}\))

Let us assume, based on the discussion above, that OC PRO and SELF-anaphors are featurally fully identical. But what feature-specifications does the anaphoric DP (henceforth DP\(_{\text{anaph}}\)) that underlingly represents both these elements start out with? Here, I present evidence to demonstrate that, contrary to many proposals which claim that nominal anaphoricity is a direct function of \(\phi\)-defectiveness, the essential featural component of an anaphoric DP is an unvalued R feature.

The idea has been floated in some recent research (Sigurdsson, 2008; Kratzer, 2009; Reu-land, 2001) that what makes a DP anaphoric is its lack of \(\phi\)-features. Indeed, crosslinguistically, many SELF-anaphors do seem to be \(\phi\)-featurally deficient: e.g. Dutch *zichzelv* (lacks number and gender), Japanese *zibunzisin* (lacks person and gender). However, there are empirical reasons to question this line of reasoning and to believe, instead, that \(\phi\)-defectiveness is neither a necessary nor a sufficient condition for (bound-variable) anaphoricity.\(^{12}\)

The first piece of counter-evidence comes in the form of non-anaphoric but \(\phi\)-defective DPs in natural language. For instance, “quirky expletives” – (non-anaphoric) datives in Per-

---

\(^{11}\) Notice that a direct binding relationship between the matrix DP *Ezekiel* and the embedded object *himself* cannot obtain because this would be non-local and violate the PIC.

\(^{12}\) The role, if any, of \(\phi\)-features in anaphoric phenomena does still need to be worked out. If \(\phi\)-feature deficiency is not required for anaphoricity, why is it so many anaphors across languages lack \(\phi\)-features?
son Case Constraint (PCC) phenomena that are $\phi$-deficient but nevertheless non-anaphoric (Richards, 2008).

Conversely, anaphoric DPs that are fully $\phi$-specified are also attested. For instance, the anaphor $ta(a)n$ (lacking person) in Tamil alternates with the fully $\phi$-specified $avan/avai$ (he/she) (Annamalai, 1999) even in local binding contexts:

(27) raman tann.ai/avan.ai aįittu.ko.ηt qa.n aan raman.NOM self.ACC/he.ACC hit.REFL.PST.3MS
    “Raman, hit himself”

Conceptually, as well, it seems clear that there is more to reference/denotation than $\phi$-feature-specification. $\phi$-features restrict the domain of referents for a DP (male vs. female, singular vs. plural, 1st-person vs. 3rd person) (Heim and Kratzer, 1998; Heim, 2008). Nevertheless, they are not enough to pick out a single individual or group of individuals – from among a set of others that all have the same $\phi$-features (see Hicks, 2009, for further arguments along these lines). Based on empirical as well as theoretical arguments like those outlined above, I propose that the essential and definitional property of this DP$_{anaph}$ is its referential defectiveness. This is characterized by a variable index on D (Heim and Kratzer, 1998). I propose that the syntactic correlate of this referential defectiveness is an unvalued [R] feature (Reinhart and Reuland, 1993; Landau, 2004). In the narrow syntax, this flags the DP as a Probe for an Agree relation which will be interpreted as variable binding at LF.

5 The finiteness issue

Recent empirical evidence has shown that, while DP-distribution might well still be a function of clausal “finiteness” in some form as per Minimalist claims (Martin, 2001), we nevertheless need to go beyond a binary finite/non-finite distinction to a more graded scale which takes into account the different flavors of infinitives, participial clauses and subjunctives crosslinguistically. More nuanced and finely structured models (as in the cartographic approaches of Rizzi, 1997; Cinque, 1999, and subsequent) achieve this by formally encoding degrees of clausal dependency in terms of the feature specification/lack thereof in an “exploded” C-domain. Thus,
the feature-specification of the various C-heads (Fin, Force, etc) in the C-layer is responsible for the types of TP complements they may embed.

Following work in Sigurðsson (2004) and especially recent work in Giorgi (2010) within such a cartographic approach, I propose that one (but perhaps more) of the C-heads in the C-domain that embeds control clauses is featurally deficient. For now, I just term this C_{def} without going into which specific head is deficient (Giorgi, 2010, identifies this as Fin, but there might be reason to think that Force is involved as well – as I briefly discuss later). Specifically, C_{def} is deficient in that it lacks temporal, spatial and speech-participant information – pertaining to the utterance context/speech event. In contrast, fully “finite” CPs (e.g. that-CPs in English) would be featurally specified for speech-event information. Subjunctive- and gerundival clauses would, presumably, fall somewhere in between these two opposite ends of the finiteness cline.

Returning to control CPs, the speech-event information deficiency (syntactically encoded as a unvalued or uninterpretable features on C_{def}) is responsible for temporal anaphoricity as in the following example:

(28) John_{i} tried [_{CP}PRO_{i} to go to school (*tomorrow)]

Significantly, it is also responsible for nominal anaphoricity as manifested by OC PRO and local and long-distance anaphors. Both types of anaphoricity arise because the embedded clause can’t anchor its tense and reference in the speech-context and thus has to anchor them in the matrix clause.\(^{13}\) I exploit this correlation in my proposal below – though I restrict my discussion to nominal anaphoricity effects, specifically as they pertain to OC PRO and SELF-anaphors.

6 Proposal: contextual allomorphy, locality, cyclic Spell-Out

The key claim that I make is that OC PRO and SELF-anaphors are underlyingly identical. Differences between the two elements fall out naturally as a function of their syntactically-conditioned
allomorphy. This claim rests on the following central assumptions about the nature and architecture of the grammar. First, I assume that the grammatical derivation is Y-modular: there is a “narrow” syntax that serves as input to both the LF and PF modules. This in turn implies that morphophonological realizations as well as LF-semantic processes will be a function of previous processes in the narrow syntax, but the syntax cannot “see” and thus will not care about requirements at the PF or LF interfaces. Spell-Out is the intersection-point at which the paths to LF and PF diverge from one another. Crucially for this proposal, Spell-Out is cyclic; specifically, I assume that each Spell-Out cycle corresponds to the computation of a single phase.\textsuperscript{14} Additionally, I assume, crucially for this proposal, that all phonological material is inserted late (Halle and Marantz, 1993) – the syntax consists of unary features or feature-bundles arranged in a systematic relational hierarchy to one another. Finally, the C-domain is exploded in the sense of Rizzi (1997) and is crucial in mediating the syntactico-semantics of the embedded clause with that of the embedding one.\textsuperscript{15}

Below, I outline the remaining details of the allomorphy hypothesis for OC PRO and SELF-anaphors.

6.1 Different syntactic positions pre-SpellOut → complementarity

Essentially, a DP\textsubscript{anaph} that is the same minimal phase as a potential antecedent at SpellOut will be pronounced as a SELF-anaphor. For a DP\textsubscript{anaph} to be pronounced as PRO, the following two conditions must jointly hold. First, the DP\textsubscript{anaph} must not be in a different phase from its antecedent at SpellOut. In addition, it must be in a TP that is embedded under a C-layer that contains a C\textsubscript{def} which is featurally deficient with respect to speech-event information, as described above.

\textsuperscript{14}Although cyclic Spell-Out is assumed in most versions of Minimalism, there is less universal agreement about what derivational chunk constitutes a Spell-Out cycle. Probably the simplest version is that which I assume here, namely that the chunk that gets spelled out in each cycle is equivalent to a phase – though this is by no means the only possibility. See, for instance, Hicks (2009) for a proposal that syntactic chunks (phases) used for LF processing are of a different size than those that are shipped to the PF interface.

\textsuperscript{15}Note that the idea that control is mediated by C/a series of heads in the C layer is not a new one but has been made previously (Borer, 1989; Landau, 2008). However, the idea that this mediation is a function of contextual features encoded on one or more C heads, as suggested here, is a relatively new one and is borrowed from recent works (see among others Giorgi, 2010; Sigurðsson, 2004; Bianchi, 2003; Ritter and Wiltschko, 2010).
6.1.1 Minimality vs. locality revisited

I observed earlier that, while \textsc{self}-anaphors are sensitive only to locality, \textsc{oc pro} is sensitive not only to (anti-)locality but also to minimality. Chomsky (1995, section 4.10) makes the following proposal concerning intervention effects for movement: in contexts where more than one binder is present in a local relationship with the bindee, Minimality/intervention effects do not obtain - because the two binders, being both equally local to the anaphor (by being in the same minimal phase as it), count as equidistant from the anaphor for purposes of binding. However, in contexts where more than one non-local binder is present, the binders are not similarly equidistant from the bound variable anaphor. Thus, Minimality effects obtain and the closer binder acts as an intervener to variable-binding by the higher binder. Crucially, this account also serves to derive the exclusive subjecthood restriction on \textsc{oc pro}. Below, I sketch different scenarios of anaphoric DP-distribution and show how the anaphoric DP gets spelled out as \textsc{oc pro} in all and only those cases where it occupies the embedded subject position.

**Scenario 1:** \textsc{dp anaph} is in matrix object position:

\[
[CP \ CP_{subj} \ldots (\text{DP}_1) \ldots [VP V \text{DP}_{anaph}]]
\]

There is a higher DP (DP\textsubscript{subj}) in the same phase which will count as the DP\textsubscript{antecedent}. \textsc{dp anaph} has access to the $\phi$-featural content of DP\textsubscript{antecedent}. Result: \textsc{dp anaph} is spelled out as a \textsc{self}-anaphor, not as \textsc{oc "pro"}.

**Scenario 2:** The anaphoric DP is in embedded object position:

\[
[CP \ldots \text{DP}_1 \ldots [CP \ CP_{subj} \ldots [sP \text{DP}_{anaph} \ldots [VP V \text{DP}_{anaph}]]]]
\]

The minimal c-commanding antecedent to \textsc{dp anaph} is the nonfinite \textsc{dp subj} which starts out in the same minimal domain (VP) as \textsc{dp anaph} and serves as an intervener to any higher potential antecedents (e.g. \textsc{DP}_1). \textsc{dp anaph} has access to the $\phi$-features of DP\textsubscript{antecedent}. Result: \textsc{dp anaph} gets spelled out as a “\textsc{self}-anaphor”.

**Scenario 3:** The anaphoric DP is in embedded subject position:

\[
[CP \ CP_{subj} \ldots \text{DP}_1 \ldots [CP C [TP \text{DP}_{anaph} \ldots [VP \text{DP}_{anaph} \ldots [VP \text{DP}_{anaph} \ldots ]]]]]
\]
$\text{DP}_{\text{anaph}}$ is in embedded [Spec, TP]. Any c-commanding antecedent to $\text{DP}_{\text{anaph}}$ is thus in a higher phase relative to it. Given PIC, $\text{DP}_{\text{anaph}}$ cannot know the $\phi$-features of a $\text{DP}_{\text{antecedent}}$ at Spell-Out. Thus, the condition for Spell-Out as a SELF-anaphor will not be satisfied and $\text{DP}_{\text{anaph}}$ will be spelled out as OC PRO. Conditions of minimality will further ensure, at LF, that the closer antecedent - namely $\text{DP}_1$ - will variably-bind the anaphor, and not a higher antecedent, like $\text{DP}_{\text{subj}}$ - yielding, in the typical case, object over subject control.

**SUMMARY:** Whenever an anaphoric DP is merged in embedded subject position of a certain type of CP, it is always and only spelled out as OC PRO.

### 6.2 Syntactic allomorphy conditions Spell-Out rules

Spell-Out rules are conditioned by Minimalist Agree: the **Probe** is $\text{DP}_{\text{anaph}}$ for its unvalued [R: ] feature. Potential Goals are c-commanding, referential DPs, potentially via intermediate functional heads (v, T, C) (see also Heinat, 2008; Kratzer, 2009). Note, by the way, that c-command relations between Probe and Goal are reversed here – (see Zeijlstra, 2010, and others for independent proposals regarding upward probing).

#### 6.2.1 Building “SELF”

In simplex transitive and ECM clauses, the anaphoric DP ($\text{DP}_{\text{anaph}}$) and its DP antecedent are in the same minimal phase pre-SpellOut.\(^\text{16}\) Notice that a phase-local c-commanding antecedent DP would itself only be available to $\text{DP}_{\text{anaph}}$ if the latter were in object position. If the $\text{DP}_{\text{anaph}}$ were in canonical subject position at [Spec, TP], a structurally c-commanding antecedent DP would have to be outside the CP phase. Since $\text{DP}_{\text{anaph}}$ has an unvalued R feature, it probes upward in the phase to get this feature valued through Agree. Since there is a DP with a valued R feature in the same phase, it agrees with this DP for the unvalued R feature and potentially

\(^\text{16}\) Note, that I am not claiming that the $\text{DP}_{\text{anaph}}$ has to be merged in the same minimal phase as its antecedent DP. The important point for this proposal is the weaker point that $\text{DP}_{\text{anaph}}$ ends up in a phase-local relationship with its antecedent before Spell-Out. This accommodates structures like (i) involving passivized control complements:

i. John, tried [PRO\(_t\) to get slapped \(_t\)]
also any unvalued $\phi$-features. At LF, this results in variable-binding of the DP$_{anaph}$ yielding obligatory coreference and sloppy identity effects. At PF, the DP$_{anaph}$ is spelled out with the $\phi$-features that match those of its antecedent and the case-feature corresponding to its position in the structure (and perhaps other language-specific factors). Consider the simplex (29) and ECM (30) structures below: 17

Narrow Syntax: $\text{DP}_{[R:]}$ AGREES WITH $\text{DP}_{ant}$

PF: $\text{DP}_{[R:]} \rightarrow \text{himself}$

LF: VARIABLE-BINDING OF $\text{DP}_{[R:]}

17This model is nevertheless still very simplified and many issues need to be investigated further. For instance, if, as I assume, the DP$_{anaph}$ might already have some pre-valued $\phi$-features, what is the formal process that ensures that its $\phi$-features match those of its antecedent? Is there a filtering mechanism at the interface that weeds out all the mismatched cases and spells out only the matched ones or is the process at least partly “constructivist” in the sense of Kratzer (2009) and crash-proof syntax in general? What about the “SELF” part of the anaphor – where does this come from? I leave these questions for future research while nevertheless acknowledging their relevance.
6.2.2 Building “PRO”

(31) a. Phase 1: $\text{DP}_{\text{anaph}}$ Agrees with $\text{C}$

\[
\begin{array}{c}
\text{CP} \\
\text{C}_{\text{def}} & \text{TP} \\
\text{DP}_{\text{anaph}} & \ldots \\
[\text{R: }] \\
\end{array}
\]

b. Phase 2: $\text{C}$ cyclic Agrees with $\text{DP}_{\text{ant}}$

\[
\begin{array}{c}
\text{vP} \\
\text{DP}_{\text{ant}} & \text{v'} \\
\text{Marlon} & \text{v} & \text{VP} \\
\text{tried} & \text{v} & \text{V} & \text{CP} \\
\text{tried} & \text{C}_{\text{def}} & \text{TP} \\
[\text{R: }] \\
\end{array}
\]
For DP$_{\text{anaph}}$ in [Spec, TP] in a control infinitive CP (31), all potential DP-antecedents are non-(phase-)local and Agree is less straightforward.

Specifically, direct, long-distance Agree with a DP for [R]/$\emptyset$-features is impossible, so Agree must proceed cyclically via a c-commanding head in the CP layer. But the C-domain of control CPs lacks features (tense, person, mood etc.) relating to the speech-event. Thus, the subject anaphor never has inherent reference (i.e. is always (referentially) anaphoric). It does not have its [R:] valued by the relevant C$_{\text{def}}$, meaning these aren’t valued yet when the DP$_{\text{anaph}}$ is spelled out. The result of this series of computations is the Spell-Out as a silent, default element: “PRO”. At the computation of the higher phase (Phase 2), Agree does obtain between a higher DP (DP$_{\text{ant}}$) and C$_{\text{def}}$. At LF, both phases are computed together, yielding long-distance variable binding. This in turn yields obligatory coreference and sloppy-reading effects.

### 6.3 Testing the hypothesis: focused C-domain?

The proposal made here thus essentially treats the phonetic silence of PRO as a direct result of its inability to Agree with a c-commanding Goal. This is due to the featural defectiveness of the C-domain and phasally-cyclic Spell-Out which forces it to be shipped to Spell-Out nevertheless. This suggests that if the feature-defectiveness in the C domain were to be somehow eliminated, then PRO should be capable of being overt.

Interestingly, there is potential evidence for such a structure in Tamil (32):

```
(32) raman$_i$       PRO saadatt.ai saappi.i]a paa.tt.aan
    raman.NOM PRO rice.ACC  eat.INF  try.PST.3MSG
    “Raman$_i$ tried [CP PRO$_i$ to eat the rice.”]
```

In the structure in (32) above, the nonfinite subject of the try-class control complement can take an obligatorily-controlled PRO subject – as expected. But if the subject is contrastively focused, the long-distance nominative anaphor taan may surface in this position. This embedded anaphor is obligatorily coreferent with the matrix subject and, crucially furthermore, is interpreted de se, just like OC PRO in this position:
(33) raman, taan saadatt.ai saappi.t, a paa.tt.aan
    raman.NOM self rice.ACC eat.INF try.PST.3MSG
    “Raman, tried $[CP$ for himself, to eat the rice.”$]

Crucially, no non-anaphoric overt DP (focused or otherwise) may occupy this position:

(34) *raman, krishnan saadatt.ai saappi.t, a paa.tt.aan
    raman.NOM krishnan rice.ACC eat.INF try.PST.3MSG
    “Raman tried $[CP$ for Krishnan$_{foc}$ to eat the rice.”$]

Such data are not restricted to Tamil or the languages of the Dravidian family. Szabolcsi (2009) mentions very similar examples in control structures in Hungarian and in other works on the subject (see (23) above). Szabolcsi (2009) presents detailed independent evidence that the overt pronoun is indeed inside the infinitival clause and that it, furthermore, is the subject of the infinitival clause. Crucially for this current discussion, these embedded nonfinite overt subjects must always be pronominal DPs which furthermore, in Szabolcsi’s terms, agree for person and number with the matrix verb. Though important differences may remain, this pattern strikingly similar to the pattern seen in the Tamil structures in (32) - (34) above.

Although more research needs to be done before this can be formalized, the mere possibility of structures like (33) and (23) above, demonstrates that the featural deficiency of the C-layer in control clauses induces these two key restrictions on the subject of its TP complement. One is a semantic restriction on the referential (in)dependence of the subject – specifically, the subject DP must be anaphoric (i.e. incapable of bearing independent reference) – characterized here as [R: ]. Another is a morphophonological restriction (in the default case) that results in the subject’s being spelled out as a silent element at PF. Crucially, structures like (33) show that these two restrictions are independent from and strictly orthogonal to each other: i.e. the phonological silence of PRO must be dissociated from its semantic anaphoricity. Note, furthermore, that such a result would be less straightforward to achieve within a movement analysis in the Hornstein (1999) (and subsequent) tradition.

Such data thus provides promising new evidence in support of the idea that the covertness of PRO has something to do with the featural deficiency of a functional head in the exploded C domain. It is important to observe, in this context, that one of these heads is supposed
to be Focus (Rizzi, 1997) that is responsible for encoding information-structural information (including focus) in the syntax.

### 6.4 Potential problems and open questions

Nevertheless, questions and issues remain – both specific to the implementation proposed here and in general with regards to the PRO-SELF allomorphy thesis as a whole.

One of the possible objections to the cyclic Agree model developed for OC PRO in Section (6.2.2) above is this: if the C head (labelled C_{def}) in the lower phase (Phase 1 in (31)) is indeed featurally deficient as I claim, how does it trigger Agree with the Probe DP in [Spec, TP]? One option would be to pursue a feature-sharing approach along the lines of Pesetsky and Torrego (2007) (and prior). Under this model, we could propose that both C_{def} and DP_{anaph} share the same unvalued [R] feature in Phase 1 – possibly C_{def} inherits the unvalued R feature from DP_{anaph}. At the computation of the next phase, C_{def} would its unvalued [R] feature checked by the c-commanding DP in Phase 2 – under the feature-sharing model, this would serve to simultaneously check the unvalued feature on DP_{anaph} as well. However, crucially for this story, DP_{anaph} cannot “wait” until this valuation happens in Phase 2 but must already be shipped to Spell-Out at the end of Phase 1 – this is what results in its being spelled out as a silent element, namely “PRO”. A logical alternative to the feature-sharing approach would be to claim that C_{def} first inherits the relevant valued feature from the antecedent DP in Phase 2 so that, when the DP_{anaph} agrees with it, it is no longer itself unvalued for the R feature. But there are two non-trivial problems with such an approach. First, given a strictly derivational model and the PIC, such a derivation would be counter-cyclic. Second, under such a story, DP_{anaph} would Agree directly (rather than cyclically, in two steps, as proposed) with C_{def} which would itself serve to check its unvalued [R] feature. But if this is indeed the case, there is no longer any significant distinction between the Agree operations in (31), on the one hand, and (29) and (30) on the other. Both would involve a single-step, phase-local Agree operation – thus, there would no longer be a way to motivate the Spell-Out distinctions between “SELF” and “PRO” based on the type of Agree operation alone.

Other gaps and open questions remain. For instance, the current proposal says nothing
about the derivation of long-distance (“SE-”) anaphors – how do these fit into the allomorphy paradigm sketched here? Logically, the easiest option would be to simply extend the paradigm to include SE-anaphors as a third type of allomorph. But there is an immediate problem with calling a SE-anaphor an independent, third type of allomorph namely that, crosslinguistically, SELF-anaphors seem to morphologically contain SE-anaphors: thus, we have the Japanese *zi-bunzisin - zibun*, Scandinavian *seg - segselv* (and its variants), and so on. Any theory that deals with SE-anaphors must, thus, be able to capture this potential morphological correlation. Another gap in the current analysis has to do with arbitrary PRO (PRO_{arb}) – can this proposal be enriched to handle PRO_{arb} and, if so, how? I acknowledge the empirical and theoretical significance of these points but have nothing more concrete to say about them at present and leave them for future research.

7 Conclusion

The proposal here has argued for a reductionist treatment of OC PRO and SELF-anaphors in terms of syntactically-conditioned allomorphy. The various empirical properties of OC PRO and SELF-anaphors are then shown to follow from a systematic division of labor across the various grammatical modules: locality and cyclic SpellOut in the syntax, and variable binding at LF. If this proposal is on the right track, Binding Condition A will be merely an epiphenomenon of these derivations and the control module will be rendered unnecessary.

Acknowledgements

I am very grateful to Thomas McFadden and Gillian Ramchand for detailed comments and discussions, to Idan Landau for detailed written feedback on an earlier draft and to Ken Safir for generously sharing unpublished work and quirky binding data. I also thank Artemis Alexiadou, Peter Svenonius, and the audiences at ConSOLE XVIII in Barcelona, the 1st LOT conference in Cyprus, and Workshop on Theoretical Morphology 5 in Wittenberg, Germany for the op-

18Another interesting question in light of the allomorphy proposal made here is whether there is an overt-anaphor equivalent to PRO_{arb}. It is possible that the impersonal *oneself* (and crosslinguistic variants thereof) is this equivalent – but this needs to be further tested.
portunity to present earlier versions of this idea and for useful feedback. Finally, thanks to the excellent organizers of the LOT conference and to Kleanthes Grohmann for making this journal article possible. Any blame for errors, misunderstandings and misrepresentations, however, undoubtedly goes to me alone.

References


Zeijlstra, Hedde. 2010. There is only one way to Agree. Talk given at GLOW 33, Wroclaw, Poland.