Computation with doubling constituents: Pronouns and antecedents in Phase Theory

Sandiway Fong Jason Ginsburg
University of Arizona University of Aizu

Abstract

In this paper, we develop a computational implementation of (syntactic) Binding Theory compatible with basic assumptions from the Minimalist Program. We take as our starting point the basic Merge/Move operations plus the probe-goal Case agreement system of Chomsky (2000, 2001) together with Kayne’s (2002) proposal that pronoun-antecedent coreference relations originate with a base-generated pronominal/r-expression doubling constituent. We propose that an anaphor and r-expression antecedent originate in a DP doubling constituent that is a (strong) phase, with the phase head ‘self’, and a pronoun and r-expression antecedent originate in a DP doubling constituent that is not a phase. Under standard assumptions about locality and Phase Theory, the r-expression half of the doubling constituent must separate from the pronominal component in a timely fashion and undergo movement to be licensed independently (i.e. receive its own theta role and have its case feature valued). We propose that movement of the r-expression is constrained by an operation of Last Resort, whereby the r-expression, under appropriate circumstances, can be remerged into a derivation and licensed. We demonstrate, via a computational implementation, how our proposals derive the classic asymmetry in distribution between pronouns and anaphors for mono- and bi-clausal sentences, ECM, picture DPs and other constructions.

1 Introduction

As part of the biolinguistics perspective, the Minimalist Program (MP) (Chomsky 1995) focuses on simple and optimal solutions to the problem of the nature of human language. It is expected that considerations of efficient computation (within the constraints of the biological system) should contribute to and help explain the shape or restrictions on the space of possible human languages (cf. Chomsky 2005). An uncontroversial property of this computational system is that it should be recursive, i.e. in principle allow the unbounded combination of a limited number of lexical primitives to form an infinite variety of specific structures that encode the variety of language phenomena. Within the strictures of the MP, it is proposed that simple recursive merge (external and internal, i.e. displacement) suffices to generate an arbitrary number of structures, and the agree relation that obtains between probes and goals serves to restrict the possible instances (Chomsky 2001). Given that recursive merge can generate arbitrarily complex structures, efficient computation dictates that
probe search be restricted in scope to cyclic domains known as phases. In this paper, we show that this computationally-motivated limit on search also serves to explain a variety of Binding Theory (Chomsky 1981, 1986) facts.

Much research has focused on the various restrictions that language imposes on coreference relations. These restrictions have traditionally been accounted for by Binding Theory, exemplified by the famous Conditions A-C (1a-c).

(1)  
   (a) An anaphor must be bound in a local domain.  
   (b) A pronoun must be free in a local domain.  
   (c) An r-expression must be free. (Chomsky 1995:96)

These Binding Conditions accurately describe a wide variety of binding phenomena. However, these Conditions have, since their formulation, been recognized to be inadequate to account for the variety of coreference phenomena found in language (e.g., see Reuland & Everaert 2001 and references cited within, among many others). In addition, while going a long way towards describing coreference phenomena, they do not explain why the phenomena are the way they are (e.g., why does Condition A hold?). Within the MP, in order to achieve a more adequate account of pronoun-antecedent relations and explain away the Binding Conditions, there have been some attempts to formulate Binding Theory in terms of movement (cf. Hornstein 2001, Kayne 2002, Zwart 2002, Heinat 2003). This paper builds on these analyses, particularly on the work by Kayne (2002). In this work, we describe a computer model that relies on independently motivated elements of Phase Theory (Chomsky 1999, 2000, 2001, 2004, 2006) to account for a variety of pronoun-antecedent relations in a computationally efficient manner.

The organization of this paper is as follows. In section 2, we present an overview of Kayne (2002), which forms the foundations of our analysis. In section 3, we describe our proposals and the computer model that we used to implement these proposals. In section 4, we demonstrate how our computational model accounts for a variety of coreference facts. In section 5, we give our conclusions.

2 Doubling constituents

This paper builds on work by Kayne (2002) that assumes certain coreference relations result from movement of an r-expression out of a doubling constituent. Kayne, building on movement analyses of control phenomena (O’Neil 1995, 1997, Hornstein 1999) and Hornstein’s (2001) movement analysis of certain pronoun antecedent relations, develops a doubling constituent analysis of pronoun antecedent relations. Kayne proposes that a pronominal element and an antecedent originate within a doubling constituent of the form [Spec Head], such as ‘[John he]’, where the Spec is the antecedent and the head is the pronominal. The Spec can move out of a doubling constituent, but the head cannot. The head of the doubling constituent is licensed in its final surface position and thus has no need to move, whereas the Spec needs to move to obtain a theta-role and case. A crucial component of Kayne’s analysis is that the Spec can only move out of a doubling constituent if the doubling constituent has undergone movement. In addition, a reflexive has a structure in which a doubling constituent moves, as in (2).
These proposals derive some typical Condition A-C effects.

The notion that only a Spec can move out of a doubling constituent accounts for Condition C effects. An r-expression cannot be c-commanded by a pronominal element because the pronominal is the head of the doubling constituent. The head is licensed (gets a theta-role and case) in its base position, whereas the Spec needs to move to get licensed. For example, the ill-formed (3a) cannot be derived, assuming the base structure in (3b). (3a) is ill-formed because the head of the doubling constituent ‘he’, not the Spec ‘John’, has been extracted. ‘He’ cannot move because then it would get a second theta-role.

(3)  
(a) *He$_i$ thinks John$_i$ is smart. (Kayne 2002:137)  
(b) thinks [John he] is smart.

Certain Condition B effects are derived as follows. In (4a), with the derivation in (4b), the doubling constituent moves from its base position to [Spec, TP]. This frees the Spec ‘John’ for movement to theta-position. Note that ‘he’ remains free in the lower clause, thus satisfying Condition B.

(4)  
(a) John$_i$ thinks he$_i$ is smart. (Kayne 2002:146)  
(b) [John thinks [TP [John he] is smart [John he]]]

On the other hand, in (5a), assuming the base structure in (5b), the doubling constituent has nowhere to move before the subject theta-position can be filled. Thus, the Spec cannot move out of the doubling constituent and the subject theta-role is not assigned.

(5)  
(a) *John$_i$ praises him$_i$. (Kayne 2002:146)  
(b) [praises [John him]]

This analysis also accounts for why an anaphor can be local to its antecedent, in accord with Condition A. In the derivation of (6a), as shown in (6b), the doubling constituent moves within the anaphor, since Kayne assumes that a reflexive has a position within it to which a doubling constituent moves (see (3) above). This movement somehow frees the Spec John for movement to theta-position. Thus, the anaphor is bound locally.

(6)  
(a) John$_i$ praises himself$_i$.  
(b) [John praises [[John him] [John him] self]]

Kayne’s system, however, faces some problems. First of all, a crucial component of Kayne’s analysis is the requirement that a doubling constituent move in order for the Spec to be extracted. This accounts for the Condition A-C effects in examples (2-6). However, it is not clear why a doubling constituent must move in order for its Spec to move.

In addition, there are some basic data that are problematic for Kayne’s analysis. Kayne’s analysis appears to predict the opposite grammaticality judgments for (7) and (8). Example (7a) is well-formed, indicating that the Spec John has moved out of the doubling constituent. However, as shown in (7b), there does not appear to be any position that the doubling constituent can move to so that the Spec John can be extracted.

Note that Kayne’s analysis does not account for why an anaphor must generally be local to its antecedent. See example (8).
(7) (a) Johni thinks that Mary likes himi.
    (b) \([v*P \ldots v* \text{ thinks } [CP \text{ that } Mary \ T \ [v*P \text{ Mary likes } [John \ him]]]]\)

Consider (8). Kayne’s analysis predicts that it should be possible for John to move out of
the doubling constituent and into theta-position, since the doubling constituent moves within
the anaphor, as shown in (8b). But the ill-formedness suggests that this is not the case.

(8) (a) *Johni thinks that Mary likes himselfi.
    (b) \([v*P \text{ John } v* \text{ thinks } [CP \text{ that } Mary \ T \ [v*P \text{ Mary likes } [DP \text{ [John him] } [John him] self]]]]\)

A possible explanation for the ill-formedness of (8) is that Mary blocks movement of John,
but if that were the case, then it is not clear why Mary would not block movement of John
in (7). Note that Kayne’s analysis thus fails to account for why an anaphor must generally
be local to its antecedent.

Next, consider the ECM construction (9a). In the partial derivation (9b), the doubling
constituent moves within the lower clause and then to the higher clause. This movement
should free the Spec John for movement to theta-position, thus predicting, contrary to fact,
that (9a) should be well-formed.

(9) (a) *Johni considers himi to be intelligent. (Kayne 2002:146)
    (b) \([v*P \text{ John considers } [John him] \ [TP \text{ [John him] to be } \text{intelligent } [John him]]]\)

Kayne’s explanation for the ill-formedness of (9a) is as follows.

\[
\ldots \text{‘raising to object’ must apply first and […] once it does [John him] is too high in the structure for there to be any available intermediate position above it, yet below the subject theta-position of consider. (Kayne 2002:146)}
\]

However, it is not clear why raising to the object position of an ECM verb, as well as raising
to the subject position of the embedded clause, should not count as movement that frees the
Spec of a doubling constituent.

Another problem for Kayne’s analysis can be found with possessive determiner phrases.
In the well-formed possessive (10a), it is not clear where the doubling constituent can move
to in order to enable extraction of the subject ‘John’.

(10) (a) Johni likes hisi dog.
    (b) likes [John he]’s dog

On the other hand, Kayne’s analysis predicts that the ill-formed (11a) should be well-formed,
since the doubling constituent can move within the anaphor, thereby allowing ‘John’ to move
to subject theta-position.
(11)  (a) *John$_1$ likes himself$_1$’s dog.
    (b) John likes $[_{DP} [\text{John him}] [\text{John him}] \text{self}]$’s dog

In this paper, we also adopt a version of Kayne’s doubling constituent proposal. Also see Zwart (2002)$^2$ and Heinat (2003)$^3$ for related analyses. Our proposal, however, predicts data that Kayne’s analysis can account for, as well as data that are problematic for Kayne’s analysis, without the stipulation that a doubling constituent move, and in a computationally efficient manner. We next turn to our proposals.

3 Proposals

We follow Kayne (2002) in assuming that a pronoun and antecedent originate as a doubling constituent, and we also follow Kayne’s idea that constraints on movement out of a doubling

---

$^2$Zwart (2002) assumes (following Kayne 2002), that coreference is the result of merge of an antecedent and a variable referential element. Zwart (2002:274) writes “A pronoun $\alpha$ is coreferential with $\beta$ iff $\alpha$ is merged with $\beta$. Unlike in Kayne’s analysis, though, the pronominal element that originates in a doubling constituent must be an anaphor, and all other forms of pronoun-antecedent coreference are considered to be accidental. On the one hand, this analysis does away with the troubling requirement that a doubling constituent move in order for the Spec to be extracted, but on the other hand, it does not provide a principled account of pronoun coreference, which Zwart suggests is subject to certain pragmatic constraints. For example, it is not clear how to account for traditional Condition C effects. Zwart points out that (i) is ill-formed because if ‘John’ and ‘him’ originate in the doubling constituent ‘[John him]’, then it must be Spelled out as shown in (ii).

(i) John$_i$ loves him$_i$.
(ii) John$_i$ loves himself$_i$.

But then it is not clear why (i) cannot result from accidental coreference, as in (iii).

(iii) John$_i$ thinks that Mary$_i$ loves him$_i$.

$^3$Heinat (2003) presents a revised version of the doubling constituent proposal in which he too assumes base generation of a doubling constituent and he does not follow Kayne’s requirement that a doubling constituent move. Rather, the form of a pronominal (whether it is pronounced as a pronoun or anaphor) depends on whether or not it is sent to Spell-Out at the same time as its antecedent. Heinat’s proposal relies on the Phase Theory (Chomsky 2000, 2001) view that when the edge of a phase is reached, the complement of the head of the phase is sent to Spell-Out. When the antecedent and the pronominal, which Heinat refers to as a PRONOUN, are sent to Spell-Out simultaneously, the PRONOUN is pronounced as an anaphor. When the PRONOUN is sent to Spell-Out before the antecedent, the PRONOUN is pronounced as a pronoun. One potential problem with this proposal is that it requires there to be object shift in an English construction with an anaphor object. If (i) has the underlying structure in (ii), then the PRONOUN is sent to Spell-Out when the v$^*$P edge is reached, and ‘Mary’ will be sent to Spell-Out later - thus the PRONOUN should be pronounced as ‘her’ contrary to fact.

(i) Mary$_i$ likes herself. (Heinat 2003)

(ii) $[_{CP} \text{Mary} \ [_{vP} \text{Mary likes [Mary PRONOUN]}]]$
To get around this problem, Heinat takes the position that the doubling constituent undergoes object shift to the v$^*$P edge, as in (iii) below.

(iii) $[_{CP} \text{Mary likes } [_{vP} \text{[Mary_copy] PRONOUN }]_{vP} \ldots$ (adapted from Heinat 2003)
But this means that the verb must also move out of the vP or else the object ‘herself’ would precede ‘likes’. Yet, it is not clear where the verb would move to. Furthermore, it is not clear that there is Object Shift in this construction. Heinat presents Chomsky’s (2001) claim that Object Shift can occur in English, as in (iv).

(iv) (guess) what$_{Obj}$ [John$_{Subj}$ T [$_{vP}$ l$_{Obj}$ [$_{Subj}$ read l$_{Obj}$]]] (Chomsky 2001:26)
However, Chomsky claims that when there is Object Shift in languages like English “the object must move on beyond the position of OS [Object Shift] (Chomsky 2001:26).” In (iv), Chomsky claims that the object moves to [Spec, CP]. Furthermore, even if there is Object Shift in a relative clause such as (iv), it is not clear that there is Object Shift in a simple transitive clause such as (i).
constituent account for pronoun-antecedent facts. However, we differ from Kayne with respect to the structure of a doubling constituent and with respect to how movement occurs out of a doubling constituent.

We assume a view of Phase Theory whereby a derivation is broken up into phases that are formed via selection and merge of Lexical Items (LIs) from a numeration. Strong phase heads are C (complementizer), v* (little v) and certain D (determiner) heads (self or possessive ‘s’). We propose a revised view of the Phase Impenetrability Condition (PIC). In the view of the PIC in Chomsky (2004), when a phase head is merged, the complement of a lower phase head is sent to Spell-Out. In our view, when a phase head is merged, an entire lower phase (not just the complement of the phase head), if present, is closed off, where ‘closing off’ refers to a domain that becomes unaccessible. In (12), when the phase head C is merged, the lower v*P phase is closed-off and becomes unaccessible to further syntactic operations.

\[(12) \quad [c C T [v* v* V \ldots]]\]

This closing off results from the need to restrict the domain of probe search. Beyond two phases, memory limitations make search inefficient. We demonstrate in this paper that, for the constructions discussed in this paper, there is no need to access phase edges. We propose the structures in (13a-b), where (13a) is a doubling constituent consisting of a pronoun and antecedent r-expression (‘r-expr’) and (13b) is an anaphor and antecedent r-expression.

\[(13) \quad \begin{align*}
& (a) \text{pronoun and r-expression} \\
& (b) \text{anaphor and r-expression}
\end{align*}\]

In (13a), the pronoun (e.g., ‘he/him’) is a a noun with an r-expression DP complement (e.g., ‘John’). We propose that the DP in an anaphor, (13b), is a (strong) phase DP with the D phase head ‘self’. Morphological affixation between ‘self’ and a pronoun results in a reflexive; for example; ‘self’ combined with ‘him’ results in ‘himself’ at Spell-Out. The DP in (13a) lacks ‘self’ and is not a (strong) phase.

In addition, we incorporate the idea that the grammar makes available an operation of Last Resort. Last Resort (Chomsky 1995) was originally formulated as an operation that requires movement to involve a feature checking operation. Our version of Last Resort also involves feature checking. Crucially, though, we propose that Last Resort arises when an LI that needs to undergo a feature checking relation is in danger of falling outside the scope of a goal - this condition arises when the LI is contained within a phase that is about to be closed off to higher operations. Last Resort (LR) is defined as follows.

\[\text{If our analysis is on the right track, it suggests that there may not be any need to use phase edges as escape hatches, a view which could lead to a simplification of Phase Theory. However, we leave for further analysis investigation of whether or not there is a need to access phase edges in derivations.}\]
(14) Last Resort (LR): an LI with an unvalued feature that is in imminent danger of falling outside of a probe-goal scope relation can:

(a) preferentially undergo internal merge into an available theta-position
(b) or, go into a subject buffer for later external merge
(c) external merge to a theta-position has priority over Last Resort

If an unlicensed DP remains in a closed off domain, then a derivation will crash. To avoid this non-optimal outcome, the human grammar module makes available LR as an escape mechanism. The DP, if possible, is theta-merged with the head of a tree (14a). This is the most optimal operation, since the LI is reused immediately. If that is not possible, which is the case if the head of the tree does not select for a theta-position, then the LI will be set aside for further merge. The grammar has a special buffer that houses a subject and that the LI can be placed in (14b). From the subject buffer, the LI can later be selected and merged into subject position. There is a catch though - this LR operation is blocked by external theta-merge (14c); if external merge into theta-position of a phrase from the numeration is possible, then LR is blocked. This arises out of the need for efficient computation - external merge is more economical than the LR movement process.

The LR process, combined with our view that a phase is closed-off when a higher phase head is merged, accounts for the possibility of remerge as in (15a-b), where ‘[F:−]’ represents an unvalued feature of an r-expression. Assume that the r-expression lacks case and/or a theta-role.

(15) (a) \[v^* \ldots [d \text{self} \ldots \text{r-expr}_{F:−} \ldots]]\]
(b) \[c \ldots [v^* \ldots \text{r-expr}_{F:−} \ldots]]\]

In (15a), ‘self’ and v* are phase heads. When v* is merged, the lower DP phase will be closed off. Since the r-expression contains an unvalued feature, it is subject to the LR process and it will be immediately remerged with v*; that is, if external theta-merge is not an option (i.e., there is no other DP available within the numeration). In (15b), both C and v* are phase heads. When C is merged, the lower v*P will be closed off. In this case, the r-expression thus is subject to the LR process. Note, though, that C does not select for an argument; theta-merge with C is not possible. Therefore, the r-expression will go directly into a subject buffer, from which it can later be remerged into a derivation.

The computational system has a choice when it comes to selecting between theta merge operations: (a) external merge, (b) LR operation of immediate remerge, and (c) LR operation of insertion into a subject buffer. Our proposals, namely options (b) and (c), do not compromise the optimality of Chomsky’s system: in other words, theta merge options are managed such that these options are in complementary distribution, and extra choice points (that weaken the optimality of derivations) are never introduced. Furthermore, the order of preference between these options is directly tied by computational efficiency concerns. The first option (a), external merge directly into theta position, is a primitive of Chomsky’s system, and thus the simplest possible operation. It is therefore preferred (when available) over the two remaining possibilities.\(^5\) The second option, a direct theta merge operation

\(^5\)See (28a) and (37) below for examples where direct theta merge blocking LR is crucial to our analysis.
after search to the limit of the probe mechanism, is computationally more complex than option (a), and thus is dispreferred. The last option, namely storage into a special buffer for delayed theta merge, still achieves theta merge but is more complex still. It involves additional “hardware” (i.e. the buffer) and multiple copy operations (i.e. copy into and out of the buffer).

We implemented this LR process via a computer model. The model automatically constructs the derivation of a sentence from a numeration that it is fed. LIs are selected from a numeration and merged. Example (16) is a screenshot of the derivation of the simple sentence ‘John likes Mary’. The left side of the Tree Viewer lists each step of the derivation. The middle of the tree viewer shows a tree diagram of each step of the derivation. The right side shows each step of the derivation in bracket format. Initially, case on ‘Mary’ and ‘John’ is unvalued, as represented with ‘n!case’. When case becomes valued, the ‘n!case’ disappears - it is valued and eliminated. V merges with ‘Mary’. Then v* merges and assigns ‘Mary’ case. This is followed by theta merge of ‘John’, etc.

(16)

This computer model incorporates the LR process. Under the appropriate circumstances, LR allows a derivation to converge. Under other circumstances, an LI with an unvalued feature is stranded in a closed-off phase and a derivation crashes. We next demonstrate how this model constructs the derivations of sentences with pronoun-antecedent relations.

4 Derivations

In this section, we demonstrate how the LR process accounts for a variety of pronoun-antecedent data. We examine mono-clausal constructions, bi-clausal constructions, possessive DPs and picture DPs.

4.1 Pronoun-antecedent relations in mono-clausal constructions

First of all, we account for the complementary distribution of pronouns and anaphors in local relations, as in (17a-b), originally presented as (5) and (6) above.

(17) (a) *John, praises himi.

(b) Johni, praises himselfi.
The distinction between these examples results from the (im)possibility of the r-expression ‘John’ to undergo the LR process, which depends on whether or not the r-expression is base-generated within a DP phase.

The derivation of the ill-formed (17a) proceeds as shown in (18a-e), which are screenshots of the derivation as it is constructed automatically by our model.

(18) (a)  (b)  (c)  (d)  (e)

Initially, D and ‘John’ are merged (18a). This is followed by merge of the pronoun ‘he’ with the DP (18b), and then merge of a higher D, to form the complete doubling constituent (18c). Next, the V ‘praise’ is merged (18d), followed by v* (18e). When v* is merged, there is no lower phase present and thus no phase is closed off. At this point, the DP ‘John’ has unvalued features, since it lacks case and a theta role (it is not licensed). However, since it is not contained within a phase that is about to be closed off, it cannot undergo the LR process. As a result there is no DP available for merge with v* and the derivation crashes. Thus, we are able to account for the inability of a pronoun to be bound locally - a Condition B effect.

The derivation of the well-formed (17b) proceeds as in (19a-i). Crucially, the doubling constituent in this example corresponds to an anaphor and antecedent.

(19) (a)  (b)  (c)  (d)  (e)
The doubling constituent is formed in steps (19a-c), and forms a DP phase after merge of the D ‘self’. The V ‘praise’ is merged (19d), followed by merge of v* (19e). Crucially, the DP ‘John’ lacks case and a theta-role. Thus, when v* is merged, since ‘John’ is contained within a lower DP phase that is about to be closed off, and ‘John’ contains unvalued features, it undergoes the LR process. Since v* selects for a subject DP and there is no LI available for external merge, ‘John’ is immediately remerged with v*, thereby landing in subject theta-position (19f). This is followed by merge of T (19g). Then ‘John’ undergoes EPP driven movement to [spec, T] (19h). Lastly, C is merged (19i) and the derivation converges successfully. We thus account for a Condition A effect, since the anaphor is bound locally.

4.2  Pronoun-antecedent relations in multi-clausal constructions

We next turn to pronoun-antecedent relations in multi-clausal constructions. Our analysis accounts for why a pronoun and antecedent can co-occur if they are in separate clauses. Consider (4), repeated below.

(20) John, thinks he, is smart.

The derivation for this example is shown in (21). The left side of the screenshot lists each step of the derivation.
The doubling constituent ‘[he John]’ is base generated within the adjectival predicate. Then it undergoes EPP-driven movement to subject position of the embedded clause. When the matrix \( v^* \) is merged, crucially, the lower CP phase must be closed-off. Thus, ‘John’ which contains unvalued features (it lacks a theta-role and case) undergoes the LR reinsertion process and it is immediately remerged in subject-theta position in the matrix [spec, \( v^* \)]; \( v^* \) selects for DP and there is no other DP available for external merge.

Our analysis also straightforwardly accounts for the impossibility of a pronoun binding an r-expression as in (3), repeated below.

(22) *He, thinks John, is smart.

Following Kayne’s proposal that the head of a doubling constituent must be a pronoun, (22) is underivable. The doubling constituent must have the form (simplified) ‘[he John]’, and thus, it can only result in the LR movement of ‘John’, thereby producing example (20) above. Thus, the impossibility of binding an r-expression in this example, a Condition C effect, is accounted for.

In addition, we account for the possibility of long-distance co-reference relations (Condition B effects). Consider (7), repeated below, which as noted in section 2, is problematic for Kayne’s analysis.

(23) John, thinks that Mary likes him,.

The derivation is shown in (24).
When the lower v* is merged, there is no lower phase present, and thus ‘John’ cannot undergo the LR process. When the embedded C is merged, the lower v*P will be closed off. Thus, at this point, ‘John’, which contains unvalued features, is subject to the LR process. However, the head of the tree at this point, C, does not permit theta-merge. As a result, ‘John’ cannot be used immediately and be remerged into the derivation. Thus, ‘John’ is inserted into a subject buffer. When the matrix v* is merged, ‘John’ is selected and merged in subject theta-position, and the derivation converges successfully. Long distance coreference, as in (25) is accounted for in a similar manner, as shown in (26).

(25) John, thinks that Peter thinks that Mary thinks that Bill likes him,
In this construction, when the most embedded C is merged, the lower v*P will be closed off. Thus, ‘John’ undergoes the LR process. Again, since C does not permit theta-merge, ‘John’ goes into the subject buffer and then is remerged in matrix subject position.\(^6\)

We also are able to account for the impossibility of long distance binding of an anaphor, as in (8), repeated below, which we noted is problematic for Kayne’s analysis.

\[(27) *\text{John}, \text{thinks that Mary likes himself.}\]

The relevant parts of the derivation are shown in (28a-b).

\(^6\)Once a DP is inserted into a subject buffer, it can potentially be inserted into any subject position. However, there appears to be a preference for the matrix subject position. In this example, the most natural interpretation is one in which the pronoun refers to the matrix subject. We leave for further analysis why this might be the case.
The doubling constituent crucially originates within a DP phase. Thus, when v* is merged, the lower DP will be closed off. The DP ‘John’ contains unvalued features. However, in this case, the direct external merge of the subject ‘Mary’ blocks the LR process, since external merge is a simpler and more efficient operation than remerge of ‘John’ in a theta-position or insertion of ‘John’ into a subject buffer. As a result, ‘John’ remains in-situ, and ‘Mary’ is merged in subject position (28a). Then, when the matrix v* is merged (28b), there is no subject available and the derivation crashes. The impossibility of local binding of an anaphor in this construction, a Condition A effect, thus falls out of our analysis.

ECM constructions, which are problematic for Kayne’s analysis, also are accounted for. Consider (9), repeated below.

(29) *John\textsubscript{i} considers him\textsubscript{i} to be intelligent.

The relevant part of the derivation is shown in (30). The doubling constituent raises from within the adjectival predicate to [spec, T] of the embedded clause. Assuming that an ECM lacks a CP (Chomsky 1981), when the matrix v* is merged, the lower TP, which contains the doubling constituent, is not a phase, and thus will not be closed off. As a result, ‘John’ is unable to undergo the LR process, leaving no matrix subject available, and causing the derivation to crash.

---

7This blocking action is crucial. For example, if the subject buffer is available to ‘John’, the illicit derivation will go through.
We also predict the possibility of an anaphor in subject position of an ECM clause, as in (33).

(31) John considers himself to be intelligent.

Unlike in the derivation of (29) above, ‘John’ is contained within a DP phase doubling constituent. As shown in (32), the DP doubling constituent raises to the embedded [spec, TP]. When the matrix v* is merged, the DP phase will be closed-off. As a result, ‘John’ undergoes the LR process. Since v* selects for a subject and there is no other subject available for theta-merge, ‘John’ is reused immediately and remerged in subject theta-position.

(32)

Next, consider the following example, in which ‘heself’ occurs in subject position of an embedded clause.

(33) *John thinks heself is smart.

As shown in (34), when the embedded C is merged, ‘John’ will undergo the LR process and be inserted into a subject buffer (since theta-merge is not possible), after which it can be merged in subject theta-position.
We propose that this example is ill-formed due to a lexical gap - English simply does not have a nominative reflexive (e.g., there is no ‘heself’ in English). This example crashes because the embedded subject anaphor cannot be spelled out.

We have applied our analysis to a variety of pronoun-antecedent facts. We have accounted for the same data that Kayne’s analysis can account for (3-6). In addition, we have demonstrated how our analysis accounts for data that are problematic for Kayne’s analysis (7-9).

4.3 Pronoun-antecedent relations in possessive DPs

We next turn to possessive DPs. As discussed in section 2, the possessive DPs in (10) and (11), repeated below, are a problem for Kayne’s analysis. Our analysis, however, predicts these judgments.

(35) (a) John₁ likes his₁ dog.
    (b) *John₁ likes himself₁’s dog.

Example (35a) is interesting because the pronoun ‘he’, assuming that ‘he’ + ‘’s’ = ‘his’, appears to be bound locally - a Condition B violation. Our analysis accounts for the well-formedness of this construction as resulting from the possessive DP being a phase. The derivation of (35a) is shown in (36).
When $v^*$ is merged, the possessive DP phase will be closed off. Thus, LR applies to ‘John’. The r-expression ‘John’ undergoes LR and is remerged immediately with $v^*$ - a subject theta-role is assigned and the derivation converges.

The derivation of the ill-formed (35b) is shown in (37).

Our model constructs the subject doubling constituent separately from the rest of the possessive DP. The possessive D ‘s’ selects for the subject doubling constituent. When the subject DP, which is a phase, merges with the possessive D, which is also a phase head, the subject DP will be closed off - it is a phase contained within the possessive D phase. Thus, before it is merged, the LR process could apply to ‘John’, which has yet to be licensed. However, theta-merge blocks LR. In this case, there is competition between theta-merge of the subject doubling constituent and LR of ‘John’ from within the subject doubling constituent. Theta-merge wins - the subject doubling constituent is theta-merged with the possessive D. LR is blocked. As a result, when $v^*$ is merged, there is no subject available and the derivation crashes.

### 4.4 Pronoun-antecedent relations in picture DPs

We next demonstrate how our model can be extended to account for certain picture DP constructions. An experimental study of native speaker judgments of picture DPs by Keller
Asudeh (2001) found that pronouns and anaphors are in complementary distribution in some cases (38a-b) and (38c-d), but not in others (38e-f). Examples (38e-f) are interesting in that subjects found both the pronoun and anaphor to be generally acceptable.

(38) (a) ?*Hannah, found a picture of her.<br>  (b) Hannah, found a picture of herself.<br>  (c) ?*Hannah found Peter,’s picture of him.<br>  (d) Hannah found Peter,’s picture of himself.<br>  (e) Hannah, found Peter’s picture of her.<br>  (f) Hannah, found Peter’s picture of herself.<br>

The complementary distribution between (38a-b) depends on the availability of LR movement, as determined by whether or not the r-expression is base generated in a DP phase. In (38a), as shown in (39), when v* is merged, the r-expression ‘Hannah’ is not contained within a phase that is about to be closed-off. Therefore, it cannot undergo the LR process and no subject theta-role is assigned, causing the derivation to crash.

(39) In (38b), the r-expression originates within a DP phase. Thus, when v* is merged, since the DP doubling constituent will be closed off, ‘Hannah’ undergoes the LR process. In this case, it is remerged immediately in subject theta-position, and the derivation converges, as shown in (40).

---

8The judgments that we give reflect those presented by Keller & Asudeh’s (2001), as resulting from their study.
9Phrases containing “picture”, generally referred to as picture NPs or DPs, have been problematic for Binding Theory since its inception (e.g. see Chomsky 1981, Chomsky 1986).
10We note that (38f) does not seem to be completely acceptable to us - see the discussion of (46). Also, the traditional Binding Theory predicts that this example should be ill-formed (cf. Keller & Asudeh 2001).
The complementary distribution between a pronoun and anaphor in possessive picture-DPs, as in (38c-d), repeated below, is also accounted for. Crucially, a possessive D ‘s’ is a phase head.

(41)  
(a) ?*Hannah found Peter’s picture of him,i. 
(b) Hannah found Peter’s picture of himself,i.

In the ill-formed (41a), when the possessive D ‘s’ is merged, the r-expression, not being contained within a lower phase, is unable to undergo the LR process and a possessive subject theta-role is not assigned, causing the derivation to crash.

(42)  
In the well-formed (41b), ‘Peter’ is base generated within a DP phase doubling constituent. When the possessive D ‘s’ phase head is merged, ‘Peter’ undergoes LR and is remerged in subject theta-position of the possessive DP since ‘s’ selects for a subject. The derivation converges, as shown in (43).
Example (38e), repeated below, is also accounted for.

(44) Hannah, found Peter’s picture of heri.

As shown in (45), the doubling constituent is not a DP phase. However, it is contained within a possessive DP phase. Thus, when the matrix \( v^* \) is merged, ‘Hannah’ undergoes the LR process and is remerged in subject theta-position.

(45)

Lastly we turn to (38f), repeated below.

(46) Hannah, found Peter’s picture of herselfi.

Our model produces the following derivation, which crashes.
When the possessive D phase head ‘"s’ is merged, the lower DP phase containing ‘Hannah’ will be closed off. ‘Hannah’, which contains unvalued features, is blocked from undergoing the LR process because of the possibility of external theta-merge to subject position of the possessive DP. Thus, when v* is merged, there is no subject available and the derivation crashes. This is the derivation at work for people who find this example ill-formed; our intuition is that this example is perfect. However, Keller & Asudeh (2001) found that subjects considered this example to generally be well-formed. For those who find this well-formed, we suggest the following possibility; for some people, only theta-merge to v* blocks LR. Theta-merge to a possessive D, as in this example, is not sufficient to block the LR process. However, we leave this issue for further analysis.

5 Conclusion

Computationally speaking, a system that exhibits unbounded discrete infinity can still exhibit operational efficiency if its primitive operations have bounded or limited scope and there are no unnecessary choice or computational branching points in the system. The merge/agree system, as proposed in Chomsky (2001), exhibits both these properties. The probe-goal mechanism that underlies the agree relation is efficient in the (first) sense since goal search is phase bounded. Our introduction of Last Resort displacement directly to an edge theta position is similarly a local (and thus) efficient operation since it operates just within the same probe-goal search domain. Structure building in Chomsky’s system is also efficient in the (latter) sense since selecting whether to externally merge or displace is fixed by the choice of LIs and disambiguating maxims such as Maximize Matching Effects or Merge Over Move. Our proposals continue to preserve this efficiency. Last Resort creates no new computational choice points since by definition it is available only when no other operations apply. Our novel buffer mechanism facilitates long distance displacement of antecedent r-expressions to a target theta position in a single paired operation without recourse to iterated movement. It also does not introduce any extra choice points into the system since the competing operation of external theta merge takes precedence over and blocks its application.

In conclusion, this analysis accounts for more data than Kayne (2002) with fewer stipulations. A wide variety of coreference facts result from base generation of a pronoun and
antecedent within a DP; and the possibility of Last Resort movement (that can carry an r-expression into theta-position) as determined within the limits of Phase Theory.

References


Harvard University dissertation.
