

# Modeling of Coreference Relations in Multi-Object Constructions

Sandiway Fong

Jason Ginsburg

UNIVERSITY OF ARIZONA

UNIVERSITY OF AIZU

## 1. Introduction

In this paper, we demonstrate how a movement based account of traditional Binding facts, motivated by the need to obtain computational efficiency, accounts for coreference relations in a variety of multi-object constructions. The traditional Binding Conditions, given in (1), do a reasonably good job of describing a variety of coreference data.

- (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)

However, it is not clear why these types of Binding Conditions hold. In addition, there are a variety of data that pose problems for the Binding conditions. In this paper, we demonstrate how an analysis of pronoun-antecedent relations as doubling constituents, couched within Phase Theory (Chomsky 2000, 2001), accounts for coreference relations found in certain multi-object constructions.

The primary focus of our analysis is the following multi-object constructions.

(2) (a) I showed John<sub>i</sub> himself<sub>i</sub> in the mirror.

(b) I showed John<sub>i</sub> to himself<sub>i</sub> in the mirror. (Pesetsky 1995:125-126)

(3) (a) \*I showed himself<sub>i</sub> to John<sub>i</sub> in the mirror.

(b) \*I showed himself<sub>i</sub> John<sub>i</sub> in the mirror. (Pesetsky 1995:125-126)

(4) Pictures of himself<sub>i</sub> worry John<sub>i</sub>. (Belletti & Rizzi:317)

Coreference in (2a-b) and the impossibility of coreference in (3a-b) are expected in accord with Conditions A and C. However, the well-formedness of (4) is surprising since there is a Condition A violation, as *himself* is not c-commanded by its antecedent in surface structure.

In this paper, we present an account of these constructions that does not rely on representational-level Conditions A-C. In addition, our analyses have been implemented using a computer model, which automatically computes each stage of the derivation of a target sentence, and thus aids with theory development and verification.

The organization of this paper is as follows. In section 2, we present background information that this work follows. In section 3, we describe our proposals. Section 4 describes our model. Section 5 presents the derivations of

the target sentences. Section 6 is the conclusion.

## 2. Background

Kayne (2002) develops an analysis that attempts to explain traditional Binding facts without recourse to Conditions A-C. Kayne's primary proposal is that a pronominal element and an antecedent originate within a doubling constituent of the form [Spec Head] where the Spec is the antecedent and the head is the pronominal, as in (5).

(5) [John he]

In this account, a Spec can move out of a doubling constituent, but a head cannot. Also, the Spec can only move out of a doubling constituent if the doubling constituent has moved. Kayne proposes that a reflexive DP (e.g., *himself*) contains a position for the doubling constituent to move to, as shown in (6).

(6) [[John he]<sub>i</sub> [~~John he~~]<sub>i</sub>'s self]

However, Kayne's analysis faces a variety of problems. Most notably, it is not clear why a doubling constituent must move in order for the Spec to be extracted.<sup>1</sup> In this paper, we develop a revised version of Kayne's analysis.

Our analysis is framed within Phase Theory (Chomsky 2000, 2001, 2004,

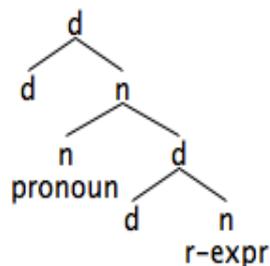
etc.). In Phase Theory, a derivation proceeds in phases. According to Chomsky (2004), when a phase head is Merged, the complement of a lower phase head is sent to Spell-Out, thereby becoming inaccessible to further operations.<sup>2</sup> This idea that, during the sentence generation process, portions of a derivation are closed-off (become inaccessible) plays an important role in determining the timing of movement in our analysis of coreference phenomena, which we turn to next.

### 3. Proposals

We assume a modified version of Kayne’s doubling constituent proposal. Our proposed doubling constituent structures are shown in (7a-b).

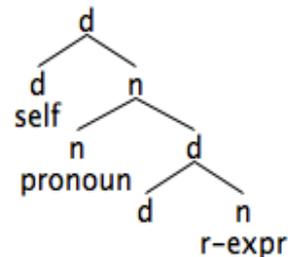
(7) (a) pronoun and r-expression

(non-phase)



(b) anaphor and r-expression

(phase)



A pronoun takes a DP r-expression complement (7a). An anaphor (7b) contains a D phase head *self*. Morphological affixation between *self* and a pronoun results in an anaphor. (7a) lacks the phase head *self* and is not a phase.

We also propose the Last Resort (LR) operation, given in (8).

(8) *Last Resort (LR)*

*A Lexical Item (LI) with an uninterpretable feature that is in imminent danger of falling outside of a probe-goal scope relation will, if possible, undergo internal merge into an available theta-position.*

The LR process makes it possible for a syntactic object to escape from a phase. For example, in (9), X\* and Y\* are phase heads, where \* signifies a phase. When X\* is Merged, the phase (or a portion of the phase) YP will be closed off and Z, a DP with an uninterpretable feature [*uFeat*], is contained within YP. Thus, Z is subject to LR and undergoes internal merge into an available theta-position in the higher XP. In this position, the uninterpretable feature is checked, resulting in [*iFeat*].

(9) Last Resort: [<sub>XP</sub> Z<sub>[iFeat]</sub> X\* ... [<sub>YP</sub> Y\* ... Z<sub>[uFeat]</sub> ...]]

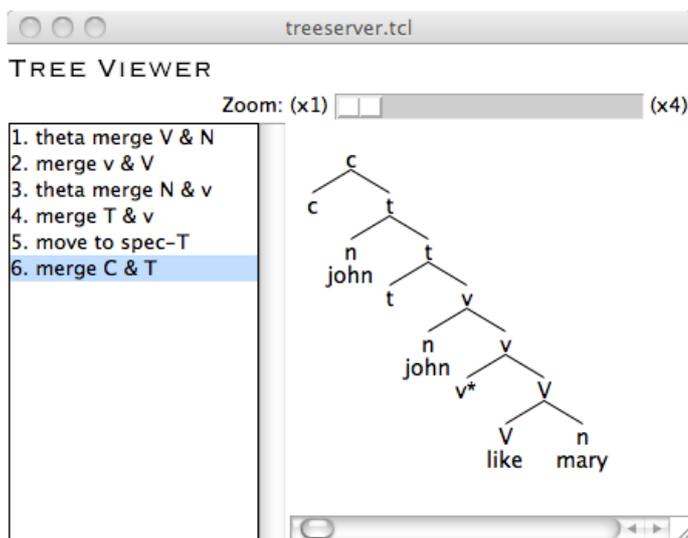
In this manner, LR-movement can save a derivation, under the appropriate circumstances.

#### **4. Model**

We tested our theory with a computer model. Our model incorporates the basic probe-goal search of Phase Theory, the possibility of LR theta-Merge when a phase is closed off, and the doubling constituent structures (7a-b). The model constructs derivations of sentences in a bottom up fashion and displays

each step of the derivation, as in (10), which shows the derivation for *John likes Mary*. The left side of the Tree Viewer describes each step. First there is theta-Merge between V and N,<sup>3</sup> followed by merge of v\* with V, etc.

(10)



In order to account for the multi-object structures that are the focus of this paper, we utilize complex-PP structures, following Pesetsky (1995), in which a PP has full argument structure with a specifier and complement. The PP in (2a), repeated below, has the structure in (11a-b), where G is a null preposition that has a theme complement and a goal specifier. *John* functions as the goal and *himself* as the theme. The theme receives Case from the preposition G.

(2) (a) I showed John<sub>i</sub> himself<sub>i</sub> in the mirror.

(11) (a) [<sub>VP</sub> V [<sub>PP</sub> Goal [<sub>P</sub> G] Theme]]

(b) [<sub>VP</sub> show [<sub>PP</sub> John [<sub>P</sub> G] himself]]

The PP in (2b) has a similar structure, but with the pronoun *to*, as shown in (12a-b). In this case, however, *to* takes a goal complement and a theme specifier.

(2) (b) I showed John<sub>i</sub> to himself<sub>i</sub> in the mirror.

(12) (a) [<sub>VP</sub> V [<sub>PP</sub> Theme [<sub>P</sub> to] Goal]]

(b) [<sub>VP</sub> show [<sub>PP</sub> John [<sub>P</sub> to] himself]]

Note that the thematic roles of *John* and *himself* differ between (2a-b), with *John* and *himself* functioning as the goal and theme in (2a), and as the theme and goal in (2b). Example (4), repeated below, has a structure with a CAUSp, which is a null P that assigns a causer theta-role but no Case. As shown in (13a-b), the CAUSp takes a theme as a complement and a goal as a specifier. There is also a CAUSv, which is a causative v that occurs in experiencer constructions. Since the theme does not get Case from the CAUSp, the theme later undergoes movement to subject position (13c).

(4) Pictures of himself<sub>i</sub> worry John<sub>i</sub>.

(13) (a) [<sub>VP</sub> CAUSv [<sub>PP</sub> Goal [<sub>P</sub>CAUSp] Theme]]

(b) [<sub>VP</sub> worry [<sub>PP</sub> John [<sub>P</sub>CAUSp] pictures of himself]]

(c) [<sub>TP</sub> pictures of himself T [<sub>VP</sub> worry [<sub>PP</sub> John [<sub>P</sub>CAUSp] ~~pictures of~~  
himself]]]

Crucially, we assume that P in these multi-object constructions is a phase head; it has full argument structure with a specifier and a complement (Chomsky 2000, 2001). This will play an important triggering role in determining the timing of the LR operation.

## 5. Derivations

In this section, we demonstrate how our proposals predict the relevant target data.

The well-formedness of (2a-b) results from the possibility of LR (see (8)), as constrained by Phase Theory. In these constructions, there is an anaphor with the structure in (7b) above. In (2a), repeated below, the doubling constituent originates as the complement to the null pronoun G, and the anaphor *himself*, which functions as the theme, is licensed in its base position.

(2) (a) I showed John<sub>i</sub> himself<sub>i</sub> in the mirror.

The doubling constituent is a phase. In addition, the preposition G is a phase head. When G is Merged, the lower doubling constituent phase will be closed off.<sup>4</sup> If *John* were to remain in the lower DP phase, the derivation would crash because *John* could not be licensed (it lacks a theta-role and case). Since the lower DP phase will be closed off, LR is able to apply, and the goal *John* is remerged in theta-position in [Spec, PP] (14a). The complete derivation is shown in (14b).



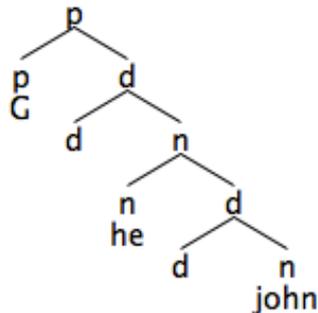


(16) (a) \*I showed John<sub>i</sub> him<sub>i</sub> in the mirror.

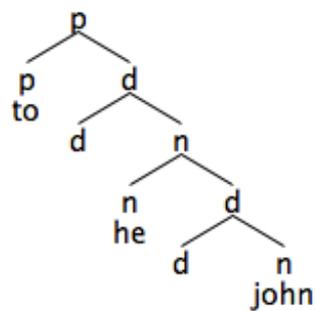
(b) \*I showed John<sub>i</sub> to him<sub>i</sub> in the mirror.

In these cases, the doubling constituent has the structure in (7a) above, which we assume is not a phase (the D head is not a phase head). Therefore, when P is Merged, there is no lower phase. See (17a-b) which correspond to (16a-b), respectively. Thus, LR cannot apply to *John*, and *John* cannot be Remerged in theta-position.

(17) (a)



(b)



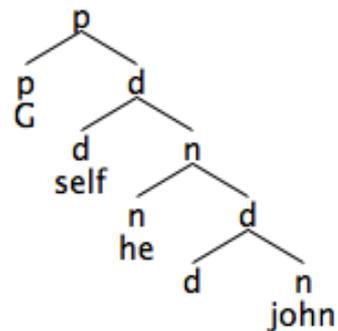
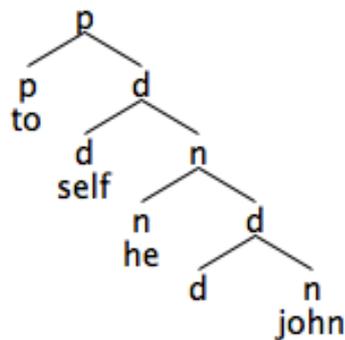
This analysis also accounts straightforwardly for the ill-formedness of (3a-b), repeated below.

(3) (a) \*I showed himself<sub>i</sub> to John<sub>i</sub> in the mirror.

(b) \*I showed himself<sub>i</sub> John<sub>i</sub> in the mirror.

Assuming the doubling constituent structure in (7b), in which *self* is the head of the doubling constituent, in (3a-b), when P is Merged (18a-b), the anaphor is licensed in its base position and has no reason to move. On the other hand, the r-expression *John* is subject to LR. As a result, (3a-b), which would require movement of *himself*, cannot be derived. Instead, these constructions will converge as (2a-b). While (2a-b) are well-formed, they are not the desired structures.

(18) (a) (b)



The ill-formedness of corresponding examples in which a pronoun c-commands the r-expression, (19a-b), are accounted for in a similar manner as (3a-b).

(19) (a) \*I showed him<sub>i</sub> to John<sub>i</sub> in the mirror.

(b) \*I showed him<sub>i</sub> John<sub>i</sub> in the mirror.

Assuming that the head of the doubling constituent is the pronoun, as in (17a-b),

the pronoun is not subject to movement, since it is licensed in its base position. Furthermore, the doubling constituent does not originate in a phase, since it is base generated as part of a non-phase structure (see the prototype in (7a)). LR cannot apply, and the derivation cannot converge.

The well-formedness of example (4), repeated below, results from the possibility of LR movement.

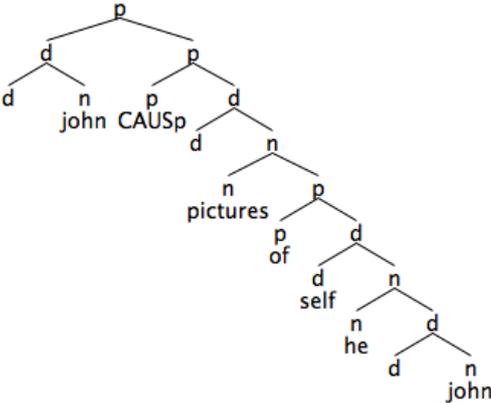
(4) Pictures of himself<sub>i</sub> worry John<sub>i</sub>.

The doubling constituent with the phase head *self* is base-generated within a *picture*-DP. The *picture*-DP Merges with CAUSp (a null P that assigns a causer theta-role but no Case). CAUSp is a phase head. As a result, the lower DP anaphor containing *John* will be closed off. Thus, the experiencer *John* undergoes LR to theta-position in [Spec, PP], as shown in (20a). LR applies because *John* lacks Case and a theta-role. Next, the head CAUSv (a causative v that occurs in experiencer constructions) is Merged into the derivation. Since CAUSv itself is a phase head dominating a lower PP phase, that lower PP phase will be closed off. Thus, the DP *pictures of himself*, which is contained within the PP that is about to be closed-off, undergoes LR to the specifier-of-CAUSv theta-position, shown in (20b). LR applies because *pictures of himself* lacks Case. The *picture*-DP gets a causer theta-role twice (once from CAUSp and once from CAUSv), but this is permitted because a DP can be assigned the same theta-role multiple times, following the analysis in Pesetsky (1995). The final

structure obtained is shown in (20c).

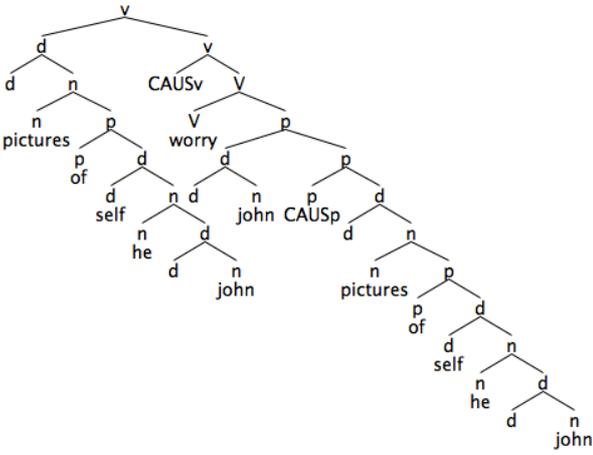
(20) (a)

1. merge D & N
2. merge N & D
3. merge D & N
4. theta merge p & n
5. merge N & P
6. merge D & N
7. theta merge p & d (causer)
8. LR theta move d (experiencer) to p
9. merge V & p
10. merge v\* & V
11. LR theta move to v\*
12. merge T & v
13. move to spec-T
14. merge C & T



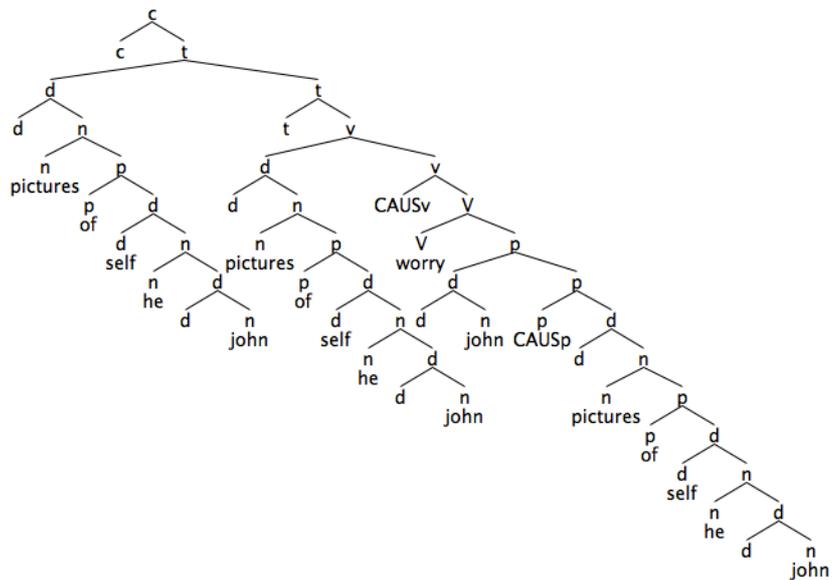
(b)

1. merge D & N
2. merge N & D
3. merge D & N
4. theta merge p & n
5. merge N & P
6. merge D & N
7. theta merge p & d (causer)
8. LR theta move d (experiencer) to p
9. merge V & p
10. merge v\* & V
11. LR theta move to v\*
12. merge T & v
13. move to spec-T
14. merge C & T



(c)

1. merge D & N
2. merge N & D
3. merge D & N
4. theta merge p & n
5. merge N & P
6. merge D & N
7. theta merge p & d (causer)
8. LR theta move d (experiencer) to p
9. merge V & p
10. merge v\* & V
11. LR theta move to v\*
12. merge T & v
13. move to spec-T
14. merge C & T

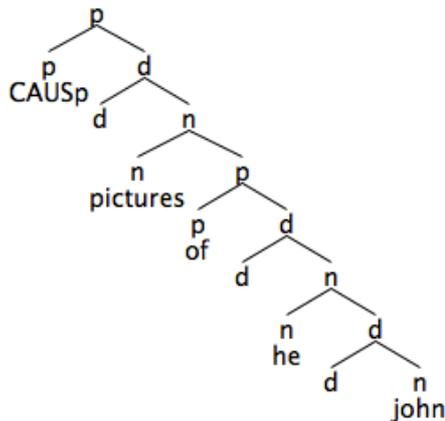


When the anaphor in (4) is replaced with a pronoun, resulting in (21), the result is ill-formed.

(21) \*Pictures of him<sub>i</sub> worry John<sub>i</sub>.

In this case, when the CAUSp phase head is Merged, the pronoun is not contained within a phase (22). Since the doubling constituent is not a phase, LR will not be triggered to bring *him* into theta-position, and the derivation crashes.

(22)



## 6. Conclusion

This paper demonstrates how a derivation based account, motivated by the need to preserve computational efficiency in probe-goal search, accounts for a variety of coreference facts involving constructions that underlyingly have multiple objects.

We can also extend this analysis to account for a variety of other data (see Fong and Ginsburg (To Appear)), such as (23a-b), which demonstrate the typical complementary distribution between pronouns and anaphors.

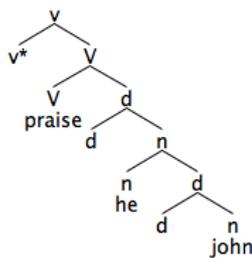
(23) (a) \*John<sub>i</sub> praises him<sub>i</sub>.

(b) John<sub>i</sub> praises himself<sub>i</sub>.

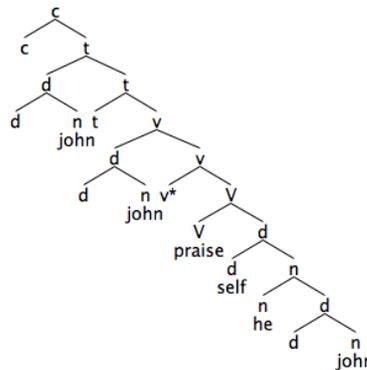
In (23a), when  $v^*$  is Merged, *John* is not contained within a phase that is about to be closed off and so LR cannot apply (24a). Thus, this is ill-formed. In (23b), when  $v^*$  is Merged, *John* is contained within a DP phase (due to the phase head

*self*). LR applies and *John* moves to subject theta-position (24b), thus resulting in a successful derivation.

(24) (a)



(b)



Importantly, this work not only demonstrates how a movement-based analysis, motivated by the need to achieve computational efficiency, accounts for a variety of coreference facts involving constructions that underlyingly have multiple objects, but it also demonstrates how a computer model can be used to confirm and verify how computation in the generative framework can be made precise.

## Notes

<sup>1</sup>See Ginsburg & Fong (To Appear) for further discussion of problems with Kayne's analysis.

<sup>2</sup>According to the Phase Impenetrability Condition (Chomsky 2000, 2001, 2004,

etc.), the complement of a phase head is sent to Spell-Out separately from the phase edge (the head and specifier, if present). Our proposals are able to account for the examples discussed in this paper regardless of whether or not a phase edge is closed-off (sent to Spell-Out) together with, or separately from, its complement. See Ginsburg & Fong (To Appear) for discussion of examples for which the issue of whether or not a phase edge is closed-off becomes important.

<sup>3</sup>For the sake of simplicity, this example uses N to represent a DP. The other examples in this paper use D.

<sup>4</sup>For the sake of simplicity, we refer to closing off of an entire phase. As noted in endnote 2, it does not matter for these examples if the complement of a phase head, or if an entire phase, is closed-off.

## References

- Beletti, A. & Rizzi, L. 1988. Psych-verbs and theta-theory. *Natural Language and Linguistic Theory* 6.291-352.
- Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
- Chomsky, N. 2000. Minimalist inquiries: The framework. In *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, ed. by R. Martin, D. Michaels and J. Uriagereka, 89-155. Cambridge, MA: MIT Press.
- Chomsky, N. 2001. Derivation by Phase. In *Ken Hale: A life in language*, ed. by M. Kenstowicz, 1-52. Cambridge, MA: MIT press.
- Chomsky, N. 2004. Beyond explanatory adequacy. In *Structures and beyond: The cartography of syntactic structures, volume 3*, ed. by A. Belletti, 104-131. Oxford, UK: Oxford University Press.

- Fong, S & J. Ginsburg. To Appear. Computation with doubling constituents: Pronouns and antecedents in Phase Theory. In *Towards a Bilingual understanding of grammar: Essays on interfaces*, ed. by A. M. Di Sciullo. Amsterdam: John Benjamins.
- Kayne, R. 2002. Pronouns and their antecedents. In *Derivation and explanation in the Minimalist Program*, ed. by S. D. Epstein & T. D. Seely, 133-166. Malden, MA: Blackwell.
- Pesetsky, D. 1995. Zero syntax: Experiencers and cascades. Cambridge, MA: MIT Press.

**(sandway@email.arizona.edu)**

**(jginsbur@gmail.com)**