Agree and Merge: Locality in copy-raising

Milan Rezac
University of Toronto & Université de Nantes/Naoned (A.A.I.)
mrezac@chass.utoronto.ca

Abstract

Chomsky (2000) proposes that movement is purely a sequence of the Agree and Merge operations. This paper solves the problem of how non-thematic base-generation in [Spec, HP] Merges a copy of the goal of the Agree by H rather than another DP linked to some free variable in H'. This is done by using Agree to transmit variable assignment indices as φ-features. The same Agree-Merge composition and interpretation is found in copy-raising, whose sole difference from movement rests in properties of the copy-deletion algorithm and its interaction with the binding theory. If a non-thematic DP is sister to a predicate whose head has not φ-Agreed, it is correctly predicted to be free to bind any variable, giving rise to broad subject and other resumptive constructions. Tough movement, however, is shown to be a species of copy-raising targeting a DP in an Ā-position (OP).

1 Introduction: The compositional theory of movement and Agree

This work develops the COMPOSITIONAL THEORY OF MOVEMENT (CTM) suggested in Chomsky (2000) [MI], where movement arises as a composition of Agree and Merge. MI's approach to CTM turns out to be unique, because of its distinctive ontology of the core syntactic dependency-forming relation Agree between two syntactic objects α and β: unlike Attract, Agree does not involve creating a representation where α and β stand in a local relation. This conclusion is based on Lasnik's (1999) demonstration that φ-Agree alone has no semantic consequences. Agree is therefore purely a feature-valuation mechanism with no rearrangement of syntactic atoms. This creates a problem for MI's CTM which does not arise for Attract, which always involves some movement with difference only of pied-piping, and for which locality conditions can be stated as conditions on movement. A way of composing the independent locality-sensitive Agree and structure-building Merge steps needs to be found, which does not resort to constructs such as derivational memory. This is proposed here, and it turns out to bring into the fold constructions which clearly show two independent Merge steps and yet are constrained by locality.

In movement, the copy in a non-thematic position must be interpreted in some way so that it "shares" the theta-role of some lower copy, where a thematic interpretation is just licensing by a lexically specified λ-operator. Heim & Kratzer (1998) develop a version of predication theory for this purpose. The sister of the moved constituent is treated as a derived λ-predicate with a variable bound by its λ-operator, identifying where the moved constituent substitutes. The mechanism converting a syntactic representation to a derived predicate must know the location of the variable; this, Heim & Kratzer encode in the singular operation Move which introduces an index at the top of the targeted constituent:

(1) a [seemed Kate, to leave her_{ij} house] →Move Kate, [i [seemed x, to leave x_{ij}'s house]]

b Translation (partial): Kate, λx, seemed x, to leave x_{ij}'s house

If movement is a primitive operation this is all that needs to be said. However, in MI movement is just Agree and Merge, and Agree is just feature valuation. This is what creates the problem I am interested: how does Merge know to add the goal identified by the earlier operation Agree? Agree itself leaves no track of the goal Kate at the target T except the valuation of T's φ-features from it. The Merge step that follows should in principle be able to Merge anything, and in fact it can Merge things other than Agree's goal: it can Merge an expletive. Yet other options are severally limited:

(2) *Every girl/*Kate, seems [IP [a picture of her,] to be on the wall].
This is the question of the Agree-Merge link I address here. The approach relies first on developing an account of copy-raising constructions, which show the potential of MI's CTM. Copy-raising clearly combines the separate steps of locality-sensitive Agree and structure-building Merge:

(3) Nolwenn, seems like she, left. (English copy-raising)

2 Copy-raising: Base-generating non-thematic subjects in [Spec, TP]

Four conclusions can be reached about copy-raising, slightly elaborating on earlier work: (i) the matrix verb assigns no theta-role its subject; (ii) the matrix T must have the same φ-features as the T of the complement CP; (iii) if the matrix subject is a DP, it must be interpretively linked to a pronoun in [Spec, TP] of the embedded clause (they); (iv) the subject (the books, there) is base-generated in the matrix [Spec, TP], and does not have a copy-theoretic representation for reconstruction elsewhere.

(4) a The books1 seem like they1 won't please anyone.
   b There seem/*seems like there are several books available.
   c The shoe1 seems like it1's on the other foot now.
   d Why do/*does there seem like there are several books on the table?

Conclusion (i) is evident from the possibility of expletives and idiom chunks, and conclusion (ii) from (4)b, (4)d. (iii) can be demonstrated by minimally contrasting a construction similar to copy-raising, where however the matrix subject can be linked to a pronoun anywhere or to none at all. This construction cannot support expletives or idiom chunks, showing the matrix verb is in fact responsible for thematic interpretation of its subject (as in individual level predicates): see Potsdam & Runner (2001):

(5) a This restaurant looks like the/its food is delicious.
   b *The other foot1 looks like the shoe is on it1 now.

Conclusion (iv) is demonstrable by a wide variety of anti-reconstruction tests whereby copy-raising contrasts minimally with normal raising, and which reduce to the fact that there is a copy-theoretic representation of the subject below the matrix [Spec, TP] in raising but not in copy-raising (op.cit.):

(6) a Two people, seem two people1 to have won the lottery. seem > 2, 2 > seem
   b It seems that two people have won the lottery. seem > 2
   c Two people are such that they seem to have won the lottery. 2 > seem

(7) a Two people, seem like they1 have won the lottery. *seem > 2, 2 > seem
   b =Two people are such that they seem like they have won the lottery. 2 > seem
   c ≠It seems like two people have won the lottery. seem > 2

Conclusion (iii) shows that the interpretation of the matrix subject in copy-raising depends on cross-clausal agreement, which can take place without copy-raising of any interpretable subject as in (4)b. Largely following Polinsky & Potsdam (2001), cross-clausal agreement has the regular locality of φ-Agree, which accounts for (ii). This is independently supported by the fact that copy-raising constructions can only occur with verbs that do not assign any theta-role to a DP; if a matrix DP were present it would block φ-Agree with any DP in the lower clause. Potsdam & Polinsky's conclusions are extended and applied to English in two ways. First, the φ-features of complementizers are crucial in determining the distribution of copy-raising: that is 3rd.sg. as a D-like C and φ-Agree cannot cross it, which distinguishes seem thatC from seem likeφ C and seem asφ if/thoughC constructions with φ-less C's. Second, Ā-movement past a complementizer with φ-features places a DP in a position where it is local to matrix φ-Agree, predicting a class of constructions where cross-clausal agreement and thus copy-raising are licensed by an Ā-goal; these will be tough-movement constructions in English.
The near-paradox of copy-raising is (iii), which implies that a DP base-generated in matrix \([\text{Spec, } TP]\) must be interpretively linked to the pronoun identified by the \(\varphi\)-Agree of its T. The problem of this Merge-Agree link is thus exactly the one which faces CTM.

### 3 Agree and Merge

I start by assuming that the subject DP in copy-raising receives its interpretation by predication, as all non-thematic DPs (Williams 1994). To account for the fact that a DP base-generated in a non-thematic \([\text{Spec, } HP]\) is restricted by the Agree-valued \(\varphi\)-features of H in both movement and copy-raising, I propose that it is \(\varphi\)-Agree which is responsible for indicating where within \(H'\) is the variable bound by the \(\lambda\)-abstract which translates \(H'\) (cp. Browning 1989). \(\varphi\)-Agree is thus an index-transmission mechanism.

The basic mechanism of predication theory is Predicate Abstraction \([\text{PA}]\) (Heim & Kratzer 1998:186):

\[(8) \text{PREDICATE ABSTRACTION (PA): Let } \alpha \text{ be a branching node with daughters } \beta \text{ and } \gamma, \text{ where } \beta \text{ dominates only a numerical index } i. \text{ Then, for any variable assignment, } a, [\alpha] a = \lambda x \in D_e [\gamma] [F_{[i]}].\]

It cannot be movement which introduces the index in copy-raising. I propose that an index, namely a value of the variable assignment function, is a \(\varphi\)-feature \([ix]\), of the same kind as number: interpretable on DPs as they are inserted into the derivation. \(\varphi\)-Agree values \([ix]\) in the same way as other \(\varphi\)-features; other instances of Agree (\(\tilde{A}\)-Agree) will also value \([ix]\) if present as a free rider.

\[(9) [\text{HP } H_{[\ldots,ix=\ldots]} \ldots \text{DP}_{[\ldots,ix=n\ldots]} \ldots] \rightarrow \text{Agree } [\text{HP } H_{[\ldots,ix=n\ldots]} \ldots \text{DP}_{[\ldots,ix=n\ldots]} \ldots]\]

On DPs, \([ix]\) is interpreted to provide a variable via Fox’s (2002) approach to the copy theory, which may then be bound by a higher \(\lambda\)-operator:

\[(10) \text{Trace Conversion (modified from Fox 2002)}\]

a. Variable Insertion: \((\text{Det}[g; \ldots,ix=x_0\ldots]) \text{ Pred } \rightarrow (\text{Det}[g; \ldots,ix=x_0\ldots]) [\text{Pred } \lambda y(y=x_0)]\]

b. Determiner Replacement: \((\text{Det}) [\text{Pred } \lambda y(y=x)] \rightarrow [\text{Pred } \lambda y(y=x)]\)

to illustrate for A-movement:

\[(11) a \quad \text{Before Agree: } T_{[g;\text{probe}]} \text{ left } [\text{every girl}]_{[3, sg, ix=7]}\]

\[b \quad \text{Agree & Merge: } [\text{every girl}]_{[3, sg, ix=7]} T_{[3, sg, liw=7]} \text{ left } [\text{every girl}]_{[3, sg, ix=7]}\]

\[c \quad \text{Trace Conversion: } [\text{every girl}]_{[3, sg, ix=7]} T_{[3, sg, liw=7]} \text{ left } [\text{the girl identical to } x]\]

PA can now be trivially modified to apply to these structures. However, I propose a departure with antecedent in traditional predication theory, where it is the base-generated subject which determines the index for PA of its sister predicate. There is a conceptual reason for this, because at the point PA applies, uninterpretable material should have deleted, including valued index probes (here on T). Formulating PA to depend on the index of the subject permits a natural solution to the perennial problem of predication theories, expletives, which may be assumed simply not to have an index and thus not to trigger PA.

However, to so formulate PA, there must be a condition which restricts a DP base-generated in \([\text{Spec, HP}]\) to the same index feature value as the Agree-valued index feature of H, because it is the latter which indicates the index of the variable to be bound within \(H'\). Such a condition is arguably needed independently to explain why the subject of copy-raising agrees for \(\varphi\)-features with its T, and correctly predicts the existence of \(\varphi\)-agreeing expletives which are found in Czech (Rezac 2004, chapter 4). The condition is the Match Condition, and PA may now be formulated as follows:

\[(12) \text{MATCH CONDITION: if Merge}(\alpha,\beta), \text{ then for any formal feature } F, \text{ the value of } F \text{ on the label of } \alpha \text{ and the label of } \beta \text{ do not differ.}\]

\[(13) \text{PREDICATE ABSTRACTION (PA): Let } \alpha \text{ be a tree dominating two sub-trees, } \beta \text{ and } \gamma, \text{ such } \beta \text{ has an index feature } [ix=i]; \text{ then for any variable assignment } a, [\alpha] a = \lambda x \in D_e [\gamma] [F_{[i]}(\beta)].\]
PA now correctly applies only when a predicate has an interpretable subject. A further advantage of formulating PA to refer to subjects will be seen below when we turn to broad subjects.

4 Copy-raising versus Movement

Copy-raising and movement on this account make use of exactly the same elements: Agree, which transfers the index of a predicate variable to H, the label of the to-be derived predicate; Merge of a non-thematic DP in [Spec, HP] under the Match Condition; and Predicate Abstraction. Their interpretive differences are descriptively summarized by saying that copy-raising necessarily involves a pronoun linked to the matrix [Spec, TP], while movement necessarily involves a copy and so allows reconstruction phenomena. This difference correlates with a formal one: a clausal boundary intervenes in copy-raising but cannot do so in movement, given assumptions about successive-cyclicity.

Beyond Agree and Merge, movement involves a third operation before pronunciation, the copy-deletion mechanism $\Delta$. $\Delta$ deletes a syntactic object $\alpha$ c-commanded by a higher object $\beta$ if the two are identical (Chomsky 1995:252-3), but the deletion must be limited to a certain domain, because it does not take place e.g. across an adjunct boundary or across the subject of a full CP/TP:

(14) He said that Nolwenn saw him/*him.

Consideration of the distance spanned by successive-cyclic A and $\bar{A}$-movement reveals that the domain of deletion is approximately that of Conditions A/B (the same fundamental generalization as the GB treatment of A-traces as Condition A anaphora). $\Delta$ must further be responsible for bleeding the binding theory: Conditions B and C must not apply between copies. Thus, for two identical elements $\alpha$ and $\beta$ within the domain of Conditions A/B of each other, $\Delta$ obligatorily applies and bleeds the binding theory; outside of this domain, it cannot apply, and the binding theory applies between $\alpha$ and $\beta$. These independent requirements suffice to completely determine the distribution of copies vs. pronouns in movement vs. copy-raising, given the distribution of CP/TP boundaries in the two.

To see this in more detail, consider first "movement", (15)a. One possible source is (15)b, interpreted as in the previous section; $\Delta$ here applies obligatorily and bleeds Conditions B and C. Source (15)c is ruled out for (15)a because $\Delta$ cannot apply (non-identity) and Condition B blocks the structure. A subset of the structures ruled out by condition B might be thought to allow a Condition A anaphor, as in (15)d, but that is not a possibility because there is no distinct co-argument to satisfy the anaphor's deficiency. Finally, choosing a non-identical non-pronominal DP as in (15)e violates Condition C.

(15) a Kate came.  
b Kate$_{\gamma}$ came Kate$_{\gamma}$ (obligatory $\Delta$)  
$\rightarrow$ PA & lexicon Kate$_{\gamma}$ $\lambda x. [\lambda y. y$ came] $\lambda x [x/7]$  
c Kate$_{\gamma}$ came she$_{\gamma}$ $\rightarrow$ Interpretation =b $\ldots$ Condition B  
d Kate$_{\gamma}$ came herself$_{\gamma}$ $\rightarrow$ Interpretation =b $\ldots$ anaphora distribution  
e Kate$_{\gamma}$ came the girl$_{\gamma}$ $\rightarrow$ Interpretation =b $\ldots$ Condition C

$\Delta$ however does not apply in copy-raising structures simply because a full CP boundary intervenes. Condition C rules out a non-pronoun linked to the matrix [Spec, TP], but Condition B is also blocked by the CP boundary, differentiating copy-raising from raising:

(16) a She$_{\gamma}$ seems like she$_{\gamma}$/*she$_{\gamma}$ is about to leave. $\ldots$ $\Delta$, Condition B inapplicable  
b She$_{\gamma}$ seems to be *she$_{\gamma}$/she$_{\gamma}$ about to leave. $\ldots$ $\Delta$, Condition B must apply  
c She$_{\gamma}$ seems like *the girl$_{\gamma}$/the girl$_{\gamma}$/Kate$_{\gamma}$ is about to leave. $\ldots$ Condition C

A simple theme runs through this explanation of the movement-copy raising difference. Application of Conditions B and C, and their suspension under $\Delta$, determine where copies are required and where they are not allowed. This accounts for the differences between copy raising and movement.
5 Merge without Agree: Broad subjects and resumptive relatives

The system developed has a significant consequence. Recall how the Merge-Agree link works: non-thematic Merger of a DP in [Spec, HP] is restricted by the Match Condition to the same \( \phi \)-features, including index feature, as that which H has receives via \( \phi \)-Agree. However, if H lacks a \( \phi \)-probe, the \( \phi \)-features of a DP base-generated in [Spec, HP] are not limited. The formulation of Predicate Abstraction as referring to the index of the DP subject rather than of H then predicts that such a non-thematic DP could bind any variable in its c-command domain.

(17) Syntax by Match Condition Semantics by PA

a \([_{HP} \text{DP}_{\forall}] [_{H} x_{\forall} \cdots x_{\forall} \cdots x_{\forall}] \cdots\] \( \text{DP} \lambda x. [_{H} T^{(*)}] \)

b \([_{HP} \text{DP}_{\exists}] [_{H} x_{\exists} \cdots x_{\exists} \cdots x_{\exists}] \cdots\] \( \text{DP} \lambda x. [_{H} T^{(*)}] \) or \( \text{DP} \lambda x. [_{H} T^{(*)}] \)

Exactly this phenomenon occurs in so-called broad subject constructions, discussed by Doron & Heycock (1999) for Japanese, Hebrew, and Arabic:

(18) John-i-ga  zibun-zisin-i-no  hisyo-ga  kubi-ni  natta  (koto)

John NOM self-GEN secretary GEN was-fired fact

John, [is such that] his secretary was fired. (Japanese, Doron & Heycock 1999)

(19) \([_{TP} \text{HP} \text{ broad subject}_{[\text{ix}=\text{i}]} \text{ [TP narrow subject}_{[\text{ix}=\text{n/i}]} [_{T} T_{[\text{ix}=\text{n}]} \cdots\] ]\]

Broad subjects are in an A-position between C and the lowest [Spec, TP], which must be filled by a regular agreeing DP subject if T has \( \phi \)-features, or may be empty if there are none. In either case, the sister of broad subject is either a projection of T without \( \phi \)-features (assuming Agree-valued features do not project further), or some head between T and C without \( \phi \)-features, depending on other assumptions. The Match Condition does not restrict the index of a broad subject because the label of its sister has no \( \phi \)-features, and PA will be able to translate this sister into a \( \lambda \)-abstract binding any open variable.

Exactly the same expectations arise for languages where T systematically has no index \( \phi \)-feature, such as Breton, which is subject to the Complementarity Principle. Unlike other Complementarity Principle languages like Irish, Breton offers a [Spec, TP] A-position, and as predicted it may host base-generated subjects which any pronoun in their c-command domain:

(20) An tiez_i a vo ret  prennañ o_i dorioù.

The houses particle will.be-3.SG necessary to.close their doors

It will be necessary to close the doors of the houses.

The houses whose doors it will be necessary to close. (Breton)

As the translation shows, the surface string is ambiguous between a matrix clause and a DP modified by a relative clause (with different analyses for a, T vs. C\(+\)T resp.). This shows how resumptive constructions in general fit into the system: the C in whose CP a null operator is base-generated in a resumptive construction has no index probe, unlike the regular C which triggers \( \tilde{A} \)-movement, leaving the operator (pro) free to link to any variable in its scope.

6 \( \tilde{A} \)-movement and copy-raising: English tough-movement

\( \tilde{A} \)-movement of a DP to the left periphery of an embedded clause may take it past the C of that clause, making it the closest goal for matrix \( \phi \)-Agree. This predicts that cross-clausal agreement and thus copy-raising will be able to occur with it. I argue this is the proper analysis of tough-movement in English:

(21) a The books\textsubscript{1} are easy for Kate\textsubscript{2} [\textsubscript{CP OP\textsubscript{1} PRO\textsubscript{2} to convince people to read e./"them\textsubscript{1}.}]

b It is easy for Kate\textsubscript{2} [\textsubscript{CP PRO\textsubscript{2} to convince people to read these books/\"them\textsubscript{1}.}].
Standard diagnostics show that the matrix subject in *tough*-movement does not reconstruct in copy-theoretic terms below [Spec, TP], so it is base-generated there. The subject must necessarily link to the site identified by Ā-movement of a null operator to the edge of the embedded clause, with its special properties studied in Browning (1989). Φ-Agree by the matrix T reaches either the embedded CP if there is no OP-movement, or OP if there is one, as dictated by locality:

(22) a. It is easy [CP C PRO, to convince Nolwenn]
    b. Nolwenn is easy [CP OP, C PRO, to convince t_i]

Unlike in regular copy-raising, *there*-expletives and idioms are both excluded in *tough*-movement, but this is predicted. The matrix subject cannot be a *there*-expletive because its associate is OP, which is pro (Browning 1989) and thus definite. An idiom cannot be broken up by Ā-movement of a subconstituent.

Copy-raising fed by the Ā-system in English is only possible when the Ā-goal is a null operator, rather than an overt *wh*-word. There are no constructions of the following sort:

(23) *Kate, was asked/wondered/understood who i Nolwenn saw t_i.

However, these are in fact ruled out by of Condition C, blocking *who*. Therefore, the present account makes the prediction that Ā-movement feeds copy-raising only of the moving Ā-word is not subject to Condition C, which in English limits it to OP. In this way, all the core properties of *tough*-movement are predicted by the minimal account constructed for movement and copy-raising.

7 Conclusion

The emergent theory of Agree-Merge interaction has better empirical coverage and yet less stipulation than a singulary theory of Move. It unites a fully compositional theory of movement, copy-raising, broad subject constructions, and *tough*-movement, not reducing any of these to the other. The interaction of Agree and Merge with the copy-deletion algorithm and the binding theory at the interfaces gives the correct properties, and correctly correlates them with the characteristic structure, for each “construction”.

References


