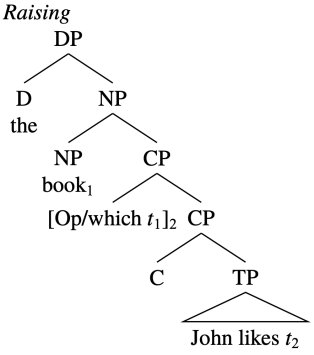
Dissolving Matching[1]

Shrayana Haldar
Massachusetts Institute of Technology

1. Introduction

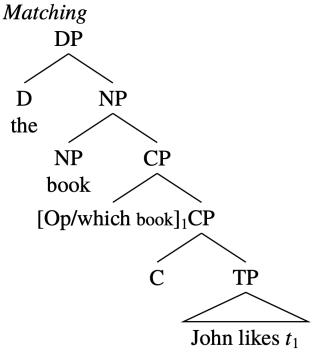
This paper concerns relative clauses (henceforth, "RCs") and their structures.[2] There's a long tradition in the literature on RCs to posit a distinction between what are known as a Raising and a Matching structure. The following are two representative examples of these two structures (1).

(1) a. *Raising*



(adapted from Bhatt 2002, 2)

b. *Matching*



(adapted from Bhatt 2002, 3)

Against this background, I've three principal goals in this paper: (a) present a unified analysis of RCs that does not posit a Raising v. Matching distinction by extending Wholesale Late Merge (henceforth, "WLM"), as presented in Takahashi and Hulsey (2009) (henceforth, "T&H") and (b) show an existing attempt at such unification (Henderson 2007) is inadequate. The following is the

[1] The content of this paper was conceived in an email conversation with Kyle Johnson. I'm grateful to him for engaging in the discussion that eventually led to this. I'm also indebted to Amir Anvari, Athulya Aravind, Bronwyn Bjorkman, Rajesh Bhatt, Danny Fox, Kyle Johnson and David Pesetsky for several discussions on topics related to this and for helping me work through the examples necessary for what happens to be extremely hard and fleeting judgements. I'm also grateful to Devon Denny, Adèle Mortier, Christopher Romero, Rachel Stacey, Anastasia Tsilia for the many judgements they provided and also to the GLOW in Asia XIII audience. This work is currently at its most initial stage and all errors are mine.

[2] All throughout this talk, I will only concern myself with restrictive RCs. So, unless otherwise mentioned, "RCs" is to be taken to refer to restrictive RCs.

structure of this paper. In §2, I summarize the motivation behind the Raising v. Matching distinction. In §3, I explain why this distinction doesn't serve us explanatorily, through an exploration of extraposition data. In §4, I show why Henderson's (2007) account of the data shouldn't be adopted. In §5, I present my implementation of WLM to substitute for the problematic part of Henderson's account. In §6, I present Johnson's (2018) multidominant model of DP Movement within which my account will be couched and how RCs will be derived given this model and my implementation of WLM. In §7, I show how this entire account can derive the extraposition data. §8 concludes the paper.

2. Raising v. Matching

It's by now widely accepted in the literature that Raising accounts for reconstruction effects. *E.g.,* in (2a), *headway* needs to reconstruct into the object position of *made* to be interpreted idiomatically inside the idiom chunk *make headway.* In (2b), *aspect of herself* needs to reconstruct for *herself* to be able to be bound by *Alma* and satisfy Condition A. But, in (3), if Raising applies, then we should predict a Condition C violation, which we don't, that is, Matching applies here. This is the basic logic of the argument and I would refer the reader to the larger literature on this for more details (Bhatt 2002, Sauerland 1998, 2000, 2003, Hulsey and Sauerland 2006, henceforth "H&S").

(2) ***Raising → reconstruction***

a. the headway that John made

b. the aspect of herself$_1$ that Alma$_1$ despises

(3) ***Matching → antireconstruction***
the aspect of Alma$_1$ that she$_1$ despises

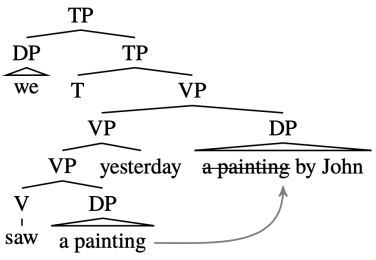
3. The Inadequacy of the Distinction

3.1. Hulsey and Sauerland (2006): Argument in Favor of the Distinction

According to F&N, we have the derivation in (4b) for an example like (4a): the DP is QRed, the higher copy is deleted because the movement is covert, and then the modifier *by John* is late merged to the restrictor of the higher copy. See F&N for extensive evidence in favor of such an analysis.

(4) a. We saw a painting yesterday by John.

b.



(Fox and Nissenbaum 1999, 3: 3)

Given this analysis, we make a prediction: if, instead of *by John,* we have a modifier which cannot be late merged, for some reason or another, that is, if it must be present in the structure throughout the time the modified DP is also in the structure, then that modifier shouldn't be able to be extraposed because extraposition of adjuncts, within this framework, must arise through Late Merge. Incidentally, Raising RCs happen to be exactly that sort of modifiers because, as we saw in (2a), the interpretation of the heads of Raising RCs is entirely dependent on its possibility of reconstruction into its internal structure, which is impossible if that internal structure is not present. That is, we predict that Late Merge/extraposition are incompatible with Raising RCs.

This prediction is borne out by examples given in H&S. Witness (5-6). ("$\#_{IC}$" stands for "infelicitous with idiomatic interpretation".) All of these examples involve Raising RCs (idioms chunks in 5: *make headway* and *take advantage;* reflexives in the RC heads in 6: *himself* and *herself*) because, in order to ensure good semantics, the heads of these RCs must be able to reconstruct, which requires Raising. And whenever we force extraposition with the adverbs shown within parentheses, ungrammaticality is incurred.

(5) a. Mary praised the headway ($\#_{IC}$ last year) that John made.
b. I was shocked by the advantage ($\#_{IC}$ yesterday) that she took of her mother.

(H&S, 9; Sportiche 2017, 32-33)

(6) a. I saw the picture of himself$_1$ (*yesterday) that John$_1$ liked.
b. Mary discovered the book about himself$_1$ (*yesterday) that Bob$_1$ wrote.

(H&S, 13; Sportiche 2017, 29-30)

3.2. Henderson (2007): Argument against the Distinction

However, Henderson (2007) shows that what actually makes these examples ungrammatical is whether a good interpretation is possible in both the higher matrix clause (henceforth, "MC", i.e., everything but the RC in the sentence/the part of the sentence from its left edge to the right edge of the crucial adverb signaling extraposition) and the RC. See the following examples that point precisely to this effect (8). *Heed, advantage,* and *headway* can be interpreted as part of the idiom chunks *pay heed, take advantage,* and *make headway* in the MC in (8a-c), respectively.

(8) a. John paid the same heed last year that Mary paid.
b. John took the same advantage last week that Mary took.
c. John made the same headway last year that Mary made.

(9a from Henderson 2007, 28: 215; the rest mine)

This shows that the Raising v. Matching distinction doesn't serve us explanatorily because, otherwise, all the examples in (8) would be bad because the Raising RC must be late merged in these extraposition examples. Henderson's account for this has two parts. I will first present

evidence against and reject the one in which he applies Vehicle Change and then preserve the other one.

4. Henderson (2007): A Quick Overview

Henderson's (2007) idea was to posit Raising for all RCs, eliminate Matching, and, in A'-Movement, to have the option to apply Vehicle Change to every copy but the highest one. This enables a derivation like (9b) for (9a), which obviates the Condition C effect that would have raised without Vehicle Change. However, this doesn't help with (10a), because the interpretation in which the universal quantifier outscopes the existential one, the disjoint reference effect persists, which means that the derivation in (10d) isn't allowed; only the one in (10c) is. And that means that Henderson's account doesn't generalize to all A'-Movements.

(9) a. the aspect of Alma_1 that she_1 despises

b. $[_{\text{DP}}$ the $[[_{\text{NP}}$ aspect of $\text{Alma}_1]$ $[_{\text{CP}}$ $[_{\text{NP}}$ aspect of $\text{her}_1]$ that $[_{\text{TP}}$ she_1 $[_{\text{VP}}$ likes $[_{\text{NP}}$ aspect of $\text{her}_1]]]]]]$

(10) a. *Someone talked to him_1 about every relative of John_1.

b. *Only interpretation considered* $\rightarrow \forall > \exists$

c. *Derivation without* Vehicle Change
*$\langle$Every relative of $\text{John}_1\rangle$ someone talked to him_1 about every relative of John_1?

d. *Derivation with* Vehicle Change
✓$\langle$Every relative of $\text{John}_1\rangle$ someone talked to him_1 about every relative of him_1?

So, I will now show how WLM can be used to substitute this Vehicle Change idea.

5. Wholesale Late Merge

Given the problematic aspect of Henderson's Vehicle Change idea, my alternative to Matching is WLM. (11) is T&H's concept of WLM in my words.

(11) Wholesale Late Merge (WLM) *(first attempt)*
The merging of the NP restrictor of a DP *can* be delayed until it reaches its final Case position, that is, the Case position in a derivation after which it's never merged into a Case position again.

This insight was an extension of the idea of late merging adjuncts, explored in Lebeaux (1988 *et sequentia*), F&N and Fox (2002). Instead of adjuncts, T&H late merges the NP restrictor of a DP. T&H exploits WLM to delay the merging of the NP restrictor in the A-Moving DPs in (12) to account for the fact that Condition C effects are unattested here, which would have been attested if the NP could be merged in embedded SpecTP because the c-commanding and coreferent

pronouns in the *to*-phrases. But by the time the derivation gets to the matrix SpecTP, which is the final Case position, the NP must be merged, which is why we get the Condition C effect in (10a), where, in order to obviate the Condition C effect, we must be able to delay the merging of the NP until SpecCP, which c-commands the final Case position, *i.e.,* the object position of *bring*.

(12) a. [The claim that John$_1$ was asleep] seems to him$_1$ to be correct.
(Chomsky 1993: 37; T&H, 8a)

b. [Every argument that John$_1$ is a genius] seems to him$_1$ to be flawless.
(Fox 1999: 192; T&H, 8b)

c. [This aspect of Sue$_1$] seemed to her$_1$ [___ to be beyond reproach].
(Pesetsky 2022, 15b)

However, this same antireconstruction effect is observable in A-Scrambling, where movement happens not *to* Case positions, but *from* them. It's widely known, from the Scrambling literature, that clause-internal movement and extraction from non-finite clauses are instances of A-Scrambling, while extraction from finite clauses are A'-Scrambling, and this split is consistent with the general profile of the A-v-A'-Movement distinction. (Hindi: Keine 2016, 2018a,b, Keine and Dash 2018, following Gambhir 1981, Déprez 1989, Mahajan 1990, 1994, Gurtu 1992, Dayal 1994, Kidwai 2000, *inter alia,* Japanese: Saito 1985, 1989, 1992, 2003, 2004, Saito and Hoji 1983, *inter alia,* Korean: Tossups 1999, 2003, 2007, *inter alia*). For instance, in the Hindi examples below, (13-14) constitute A-Movement, creating a new antecedent (*raam aur prataap* "Ram and Pratap") for binding the reciprocal (*ek-duusre-kii* "each other's"), while (15) constitutes A'-Movement and doesn't create new antecedent for binding.

(13) *Hindi*
Scrambling inside a clause

a. *[**ek-duusre-kii$_1$** bahinõ-ne] [**raam aur prataap]** -ko$_1$ maaraa.
each other's sisters-ERG Ram and Pratap -ACC hit
"*Each other's$_1$ sisters hit [Ram and Pratap]$_1$."

b. [**raam aur prataap]** -ko$_1$ [**ek-duusre-kii$_1$** bahinõ-ne] t_1 maaraa
Ram and Pratap -ACC each other's sisters-ERG hit
"[Ram and Pratap]$_1$, each other's$_1$ sisters hit t_1."

(Keine 2018a, 11: 6)

(14) *Hindi*
Scrambling out of a non-finite clause

a. *[**ek-duusre-kii$_1$** bahinõ-ne] [$_{TP}$[**raam aur prataap]** -ko$_1$ maarnaa] caahaa.
each other's sisters-ERG Ram and Pratap -ACC hit.INF wanted
"*Each other's$_1$ sisters wanted to hit [Ram and Pratap]$_1$."

b. [**raam aur prataap]** -ko$_1$ [**ek-duusre-kii$_1$** bahinõ-ne] [$_{TP}$ t_1 maarnaa] caahaa.
Ram and Pratap -ACC each other's sisters-ERG hit.INF wanted
"[Ram and Pratap]$_1$, each other's$_1$ sisters wanted to hit t_1."

(Keine 2018a, 14: 7)

(15)

Hindi
Scrambling out of a finite clause

*[	**raam**	**aur**	**prataap]**	-ko$_1$	[	**ek-duusre-kii$_1$**	bahinõ-ne]	socaa	[$_{CP}$	ki	sangiitaa-ne	t_1
	Ram	and	Pratap	-ACC		each other's	sisters-ERG	thought		that	Sangita-ERG	
	maaraa].											
	hit											

Intended → "[Ram and Pratap]$_1$, each other's$_1$ sisters thought that Sangita had hit t_1."

(Keine 2018a, 16: 8)

This means that WLM should be generalized to movement to A-positions, and not just to Case positions and then we will be able to delay the merging of the crucial NP long enough for it to escape the c-command domain of the reciprocal. Here's an attempt to capture that (16).

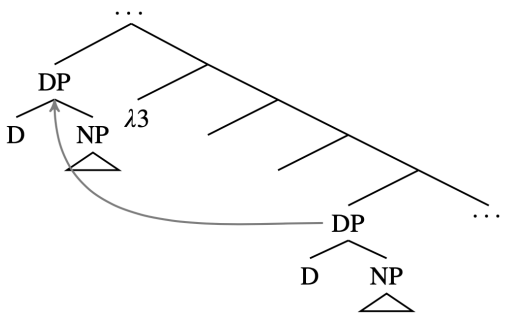
(16) WHOLESALE LATE MERGE (WLM) *(second attempt)*
The merging of the NP restrictor of a DP *can* be delayed until it reaches the final A-position, that is, the A-position in a derivation after which it's never merged into another A-position again.

This can be exploited in RCs because the RC-internal head resides in an A-position inside the RC and then, after its movement, the whole relativized DP with the moved NP head is in an A-position in the sentence that this relativized nominal is part of. That is, this is precisely the kind of situation where WLM is licensed. The specifics of that will depend on the properties of Johnson's (2018) multidominant model of DP movement that I will now introduce.[3]

6. Multidominance and RCs

I will now present the parallelism between the traditional Copy Theory of DP movement and its multidominant variant. (17) is how movement is envisioned within Copy Theory, where the lower copy is interpreted via Trace Conversion, as given in (18).

(17)



[3] I will hint at the necessity for multidominance as I conclude the paper. In short, a multidominant framework allows us to state generalizations invoking whether nodes are completely or incompletely dominated (in the sense of O'Brien 2017) by other nodes.

(18) Trace Conversion

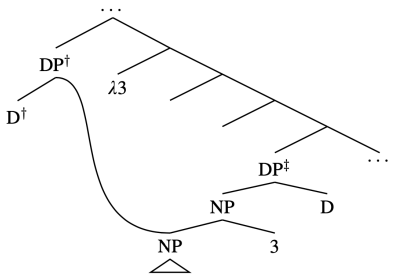
a. *Variable Insertion:*
(Det) Pred → (Det) [Pred λy . $y = x$]

b. *Determiner Replacement:*
(Det) [Pred λy . $y = x$] → the [Pred λy . $y = x$]

(Fox 2002, 10: 67)

Johnson's (2018) idea — also developed and used in Johnson (2009, 2012), Fox and Johnson (2016), O'Brien (2017) and Poole (2017) — was to do away with this separate notion of Trace Conversion and build the *result* of it into the structure itself. For instance, (19) would be the multidominant equivalent of (17). DP‡, here, should be treated as the converted trace, "D" as the replacing predicate — with a null phonology and the same semantics as English *the* — and "3", as having the semantics, "λy_e . $y = x_3$". Then, the NP undergoes Parallel Merge with D†, the determiner that gets spoken, which can be any quantificational determiner like *every, few, no, etc*, and DP† is then merged into the position where the DP is actually pronounced.[4]

(19)

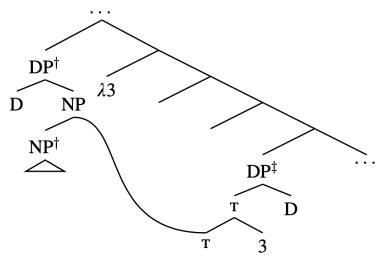


Thus, it should be evident why (17) and (19) are semantically completely equivalent, except that Trace Conversion is yet to apply to (17).

Given this conception of DP movement, where the DP itself never moves — but its NP restrictor is remerged — (20) is how WLM must be envisioned within this multidominant system to create the effect of the restrictor being absent in the "lower copy". ([[T]] has the semantics "λx_e . $x = x$". The need for this tautological predicate becomes apparent later, in (24a). In its absence, the only other node that could act as the hinge of multidominance would be "3". But then, "λx_e . $x = g(3)$" would be part of the meaning of the RC head. That means that the restrictive meaning of the RC is lost because the head is already specified as the distinct entity that g maps 3 to. That is, we end up predicting that restrictive RCs don't exist and that these RCs can only have appositive readings, which is clearly not the case. To prevent this, I use this tautological predicate.) Moreover, since no DP moves here, the WLM in (16) doesn't work. (21) is the final, revised version. (Thanks to Danny Fox for suggesting this wording.)

[4] The notion of Parallel Merge was developed in Barbara Citko's work in a different context (Citko 2005 *et sequentia*). But it's quite useful in Johnson's implementation too.

(20)

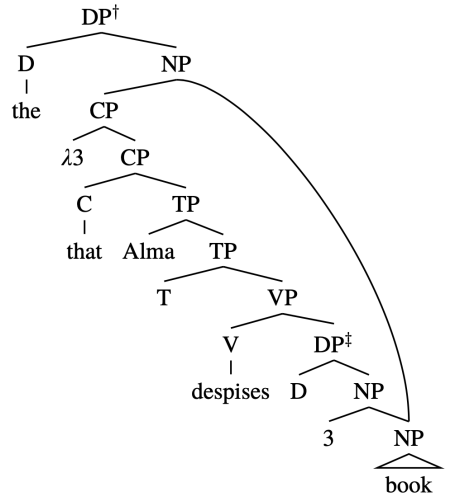


(21) Wholesale Late Merge (WLM) *(final version)*
An NP must have at least one path extending from it to the root such that the first DP in this path, counting from the NP, is in an A-position.

If DP† is in an A-position, then (21) is satisfied, which accounts of the antireconstruction data in A-Movement (12-15), while if it's in an A'-position, then (21) isn't satisfied and such an instance of WLM is banned, which accounts for the reconstruction effect in A'-Movement (10a).

In RCs, the RC-internal head NP is shared as in Johnson's work and adjoins to the CP to create the larger modified NP (22). There are two paths extending from the multidominated NP: one contains just DP† and another, DP† *and* DP‡. Both of these DPs are in A-positions. So, if we apply WLM during the remerging of the NP, then we'll be able to create a structure similar to (21) and, because of the two positions being A-positions, the WLM condition will be satisfied (22).

(22) The book that Alma despises is on the table.
Structure for the relativized DP

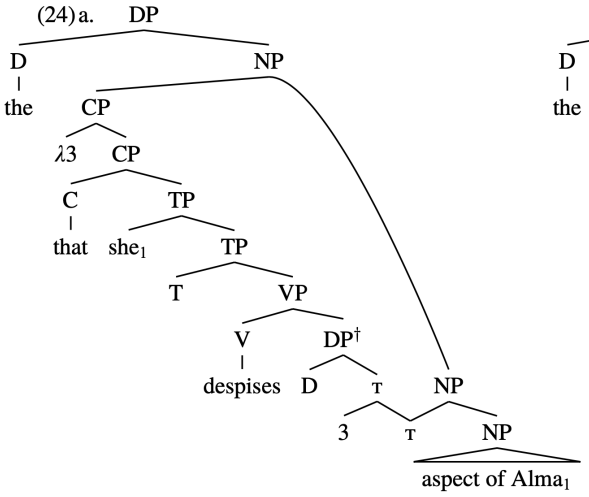


Therefore, for (23), both (24a) (with WLM) and (24b) (without WLM) will be available in principle. However, in (24b), Condition C is violated because of *Alma* being c-commanded by a coreferent *she*, so this structure will crash. For (25), (26a) will crash because *herself* isn't bound in it and Condition A is thus violated, while (26b) will survive because there is no such violation

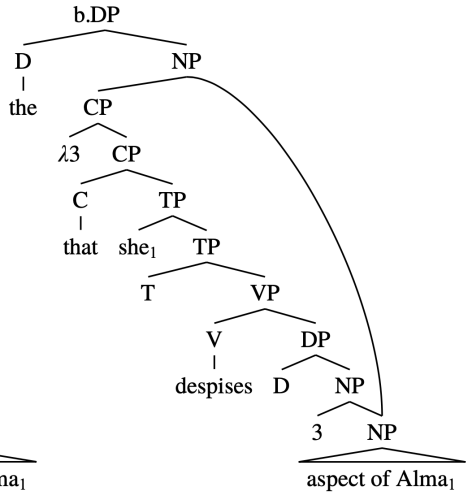
of Condition A. (27) works the same way (25) does: since "reconstruction" of *headway* is needed, the WLM derivation in (28a) crashes, the one without it in (28b) doesn't.

(23) the aspect of Alma$_1$ that she$_1$ despises

(24) a.

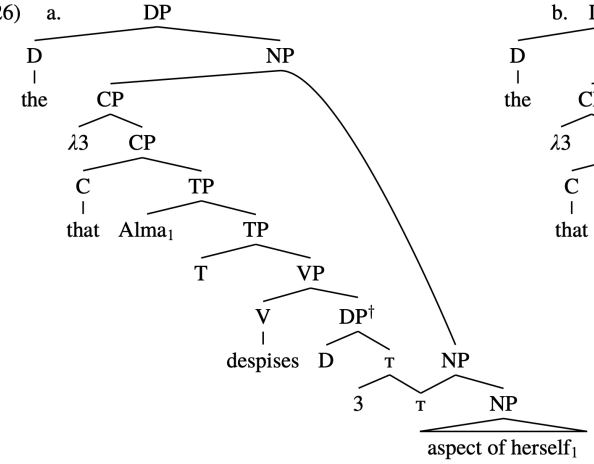


b.

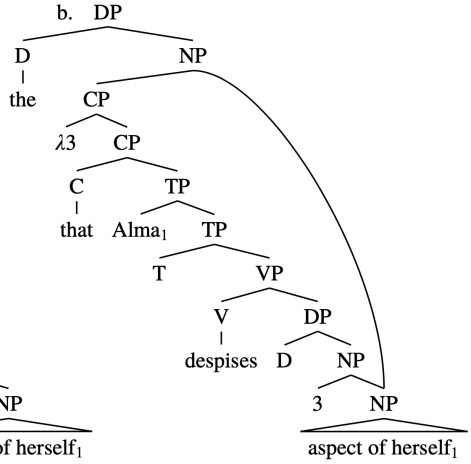


(25) the aspect of herself$_1$ that Alma$_1$ despises

(26) a.

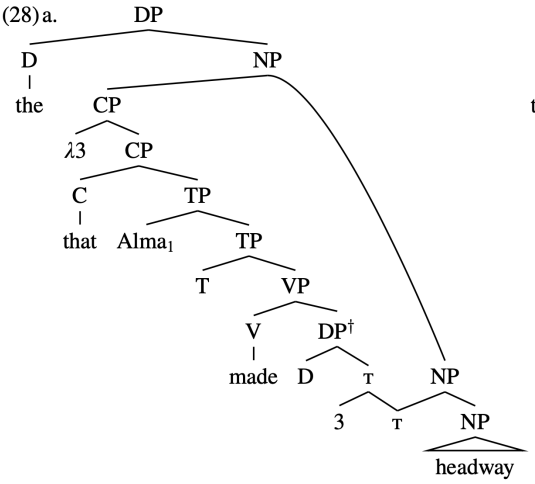


b.

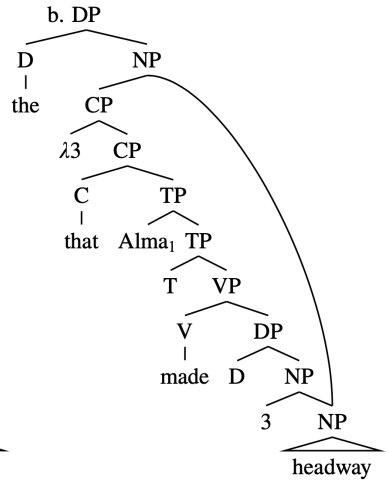


(27) the headway that Alma made

(28) a.



b.



7. Extraposition

Recall the H&S examples repeated here in (29-30), which initially showed the incompatibility between Late Merge and Raising, which were then refuted by Henderson's examples in (8) above. We're now in position to understand this phenomenon.

(29) a. Mary praised the headway ($\#_{IC}$ last year) that John made.
b. I was shocked by the advantage ($\#_{IC}$ yesterday) that she took of her mother.

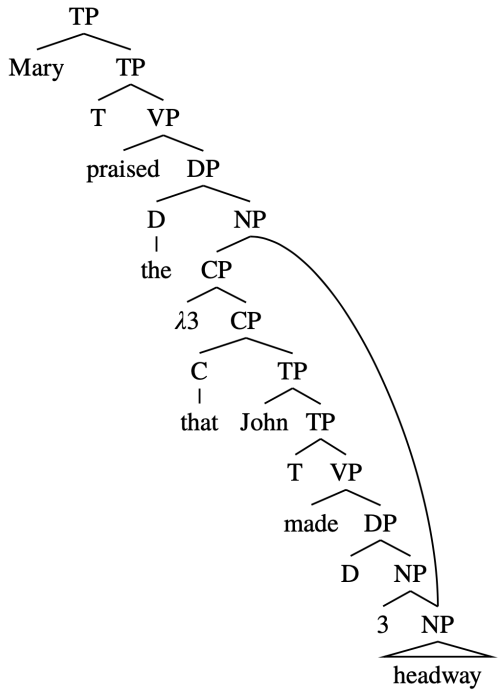
(H&S, 9)

(30) a. I saw the picture of himself$_1$ (*yesterday) that John$_1$ liked.
b. Mary discovered the book about himself$_1$ (*yesterday) that Bob$_1$ wrote.

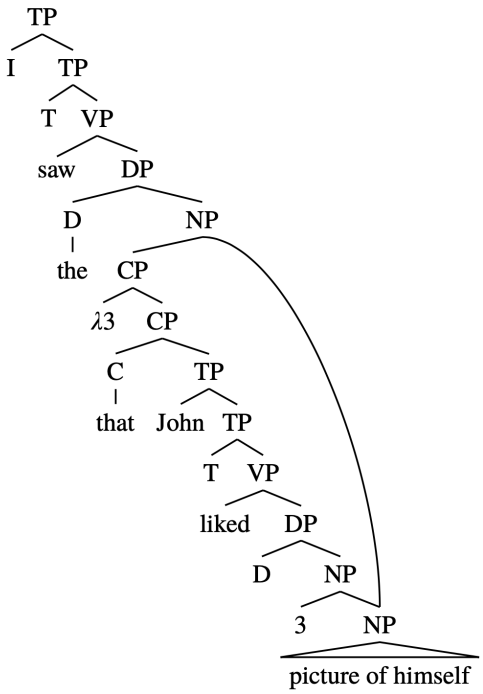
(H&S, 13)

The structures of the non-extraposition versions of (29a) and (30a) are given below for convenience.

(32) Mary praised the headway that John made.

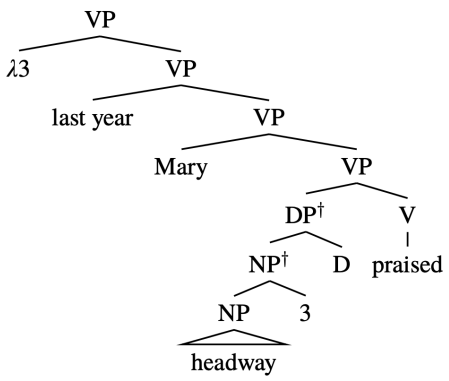


(33) I saw the picture of himself that John liked.



Given the dissolution of the distinction between Raising v. Matching in my account of RCs, we would now want to understand what makes (29-30) bad, which we can, by trying to deriving them. Recall that we're assuming Johnson's rendition of F&N's account for extraposition. Given that, for (29a), for instance, the VP in (31) is first going to have to be built — to the edge of which a larger DP will merge, just as in (4b).

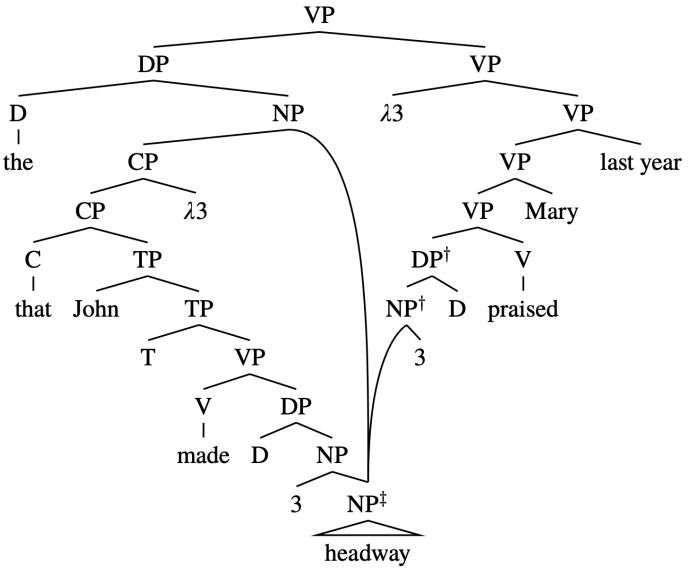
(31)



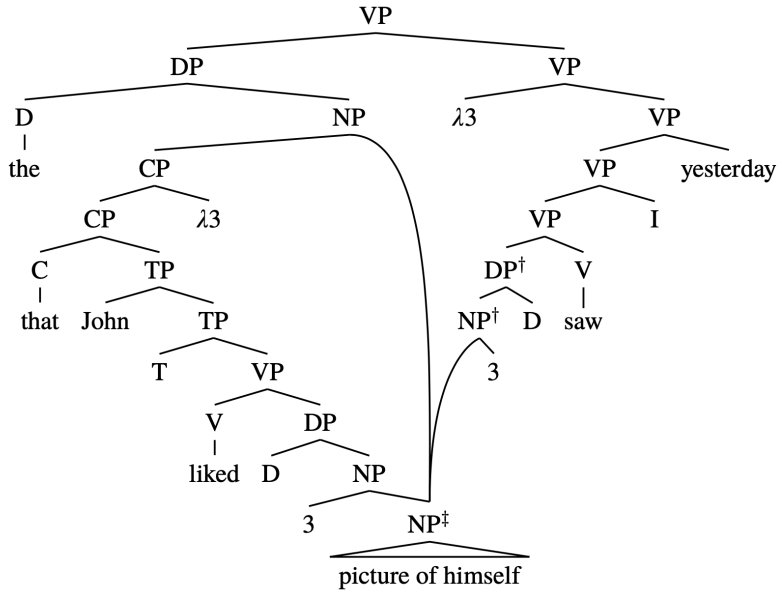
It's crucial to understand the following at this point of the derivation: the NP restrictor *headway* must be present, as in, merged into, the MC, by which I mean the object position of *praised* in the clause *Mary praised the headway*. That is because otherwise it will have no way to be pronounced there. In other words, we cannot have a derivational route in which, as in an instance of WLM, the merging of the NP restrictor *headway* is delayed because, in a later step of the derivation, we'll have an A-position where this NP can still be merged and satisfy the WLM condition, because of its consequences for linearization, regardless of what linearization algorithm we espouse.

After (31), the derivational step I propose is also one advocated for by Henderson: Sideward Movement (Nunes 2001 *inter alia*) — in order to bring what is supposed to become the head of the RC, *headway,* into the RC itself. That means, within our multidominant system, treating DP† as the converted trace of the movement and share its NP restrictor *headway* accordingly, just as in (19). When that and the following steps of Merge to build the whole RC and the relativized DP are completed and then this relativized DP is merged to the edge of the VP in (31), as in the case of extraposition in (4b), we obtain (32) for (29a). An identical derivation with different lexical items yields (33) for (30a).

(32) Mary praised the headway ($\#_{IC}$ last year) that John made.



(33) I saw the picture of himself$_1$ (*yesterday) that John$_1$ liked.



This means that we have a way to derive the sentences in (29-30). Then the question is: what makes them bad? I argue, again concurring with Henderson in this respect, that what makes these sentences bad is not the syntax but the semantics: *headway,* in (32), can't be interpreted as the object of *praise* in the MC idiomatically, which prevents the idiomatic interpretation, while *himself,* in (33), is free in the MC, because *I* can't bind it there, violating Condition A. These interpretive conflicts between the RC and the MC are what makes these sentences anomalous.

As a result of that, once the MC is enabled to provide an appropriate semantics unlike in (29-30), we obtain grammaticality, which was shown in Henderson's own examples in (8). This way of understanding RCs and extraposition thus allows us to explain the discrepancy between the examples in (29-39) and (8).

8. Concluding Remarks

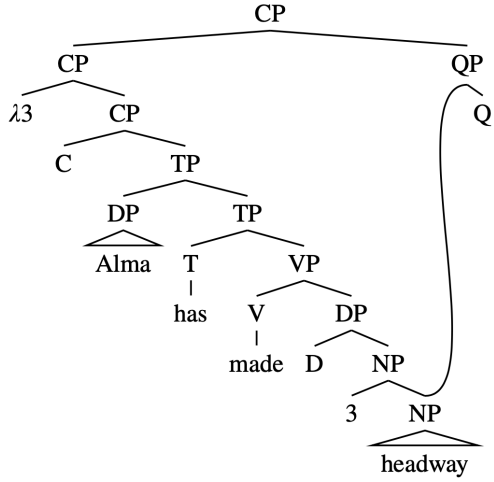
The purpose of this paper was to provide an account of the antireconstruction effect in RC heads that would eliminate the additional operation of Matching so far accepted in the literature and explain certain somewhat ignored extraposition data. It does so by proposing WLM to derive RCs across the board, thus eliminating Matching and then applying this to Henderson's (2007) approach to extraposition. This entire account is couched within a multidominant framework for DP movement developed by Johnson (2018) *inter alia*.

One thing that's often a cause for alarm when it comes to multidominant structures like the ones I have above is linearization. I have some thoughts on how to proceed on that front: Johnson (2012) develops a linearization algorithm to tackle multidominant structures of extraposition, which can be exploited to linearize these structures. Johnson didn't have Raising, *i.e.,* WLM and Sideward Movement the way I have shown, but his fundamental concepts will be useful. How, I

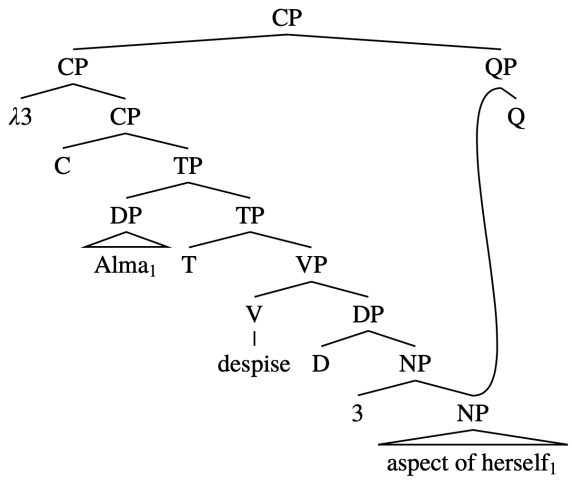
cannot fully implement and elaborate on here for reasons of space. I will leave that for future research.

There's one rather significant issue I must address in conclusion, concerning the interpretation of NPs. Observe in (32-33) that what we must say in order to ensure that these structures yield anomalous semantics is in effect that an NP must be interpreted in every single position where it sits in a structure. Otherwise, we could count only the position of the shared NPs inside the RCs for the interpretation and the critical position in the MC where these NPs cannot receive a proper interpretation could be disregarded, as just another position it occupied during the derivation — which, as it turns out, is what we ought to do in the derivations of the *Wh*-Questions given below, where we must disregard the higher position in order to account for the reconstruction effect helping the idiomatic interpretation (34) and the satisfaction of Condition A (35). So, the question we're now faced with is: why is it that we must entertain both positions occupied by the NP in (32-33) for the purposes of interpretation, but can't, in (34-35)?

(34) What headway has Alma made?



(35) Which aspect of herself$_1$ does Alma$_1$ despise?



This is probably the most interesting question this paper makes us contemplate and following this cue will lead us to understand intriguing patterns in the interpretation of DPs and their NP restrictors. I have begun to develop a preliminary sketch of an answer to this question, which involves adjusting structures like (34-35) somewhat and understanding Condition A — and conditions on the interpretation of DPs in general — in terms of complete v. incomplete dominance, to which I've alluded in footnote 3 above. I will leave that to be developed in future research and conclude this paper here.

References

Bhatt, Rajesh. 2002. The raising analysis of relative clauses: evidence from adjectival modification. *Natural Language Semantics* 10:43–90.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In *The view from Building 20*, ed. Ken Hale and Jay Keyser, 1–52. Cambridge, Massachusetts: MIT Press.

Citko, Barbara. 2005. On the nature of merge: External merge, internal merge, and parallel merge. *Linguistic Inquiry* 36:475–496.

Déprez, Vivian. 1989. On the typology of syntactic positions and the nature of chains. Doctoral Dissertation, MIT, Cambridge, Massachusetts.

Fox, Danny. 1999. Reconstruction, binding theory, and the interpretation of chains. *Linguistic Inquiry* 30:157–196.

Fox, Danny. 2002. Antecedent-contained deletion and the copy theory of movement. *Linguistic Inquiry* 33:63–96.

Fox, Danny, and Kyle Johnson. 2016. QR is restrictor sharing. In *Proceedings of the 33rd West Coast Conference on Formal Linguistics*, ed. Kyeong-min Kim, Pocholo Umbal, Trevor Block, Chan Queenie, Tanie Cheng, Finney Kelli, Mara Katz, Sophie Nickel-Thompson, and Shorten Lisa, 1–16.

Fox, Danny, and Jon Nissenbaum. 1999. Extraposition and scope: a case for overt QR. In *Proceedings of the West Coast Conference on Formal Linguistics 18*, ed. Sonya Bird, Andrew Carnie, Jason D. Haugen, and Peter Norquest, 132–144. Somerville, Massachusetts: Cascadilla Press.

Gambhir, Vijay. 1981. Syntactic restrictions and discourse functions of word order in Standard Hindi. Doctoral Dissertation, University of Pennsylvania, Philadelphia, Pennsylvania.

Gurtu, Madhu. 1992. *Anaphoric relations in Hindi and English*. New Delhi: Munshiram Manoharlal.

Henderson, Brent Mykel. 2007. Matching and raising unified. *Lingua* 117.1:202–220.

Hulsey, Sarah, and Uli Sauerland. 2006. Sorting out relative clauses. *Natural Language Semantics* 14:111–137.

Johnson, Kyle. 2009. Why movement? Unpublished manuscript, University of Massachusetts at Amherst.

Johnson, Kyle. 2012. Toward deriving differences in how Wh movement and QR are pronounced. *Lingua* 122:529–553.

Johnson, Kyle. 2018. Movement as Multidominance. Lecture notes from a course taught at the University of Crete, Rethymnon.

Keine, Stefan. 2016. Probes and their horizons. Doctoral Dissertation, University of Massachusetts at Amherst.

Keine, Stefan. 2018a. Case vs. positions in the locality of A-movement. *Glossa* 3.1:1–34.

Keine, Stefan. 2018b. Selective opacity. *Linguistic Inquiry* 50.1:13–61.

Keine, Stefan, and Bhamati Dash. 2018. The ups and downs of agreement. Ms. Los Angeles, CA: University of Southern California.

Kidwai, Ayesha. 2000. *XP-adjunction in universal grammar: scrambling and binding in Hindi-Urdu*. Oxford: Oxford University Press.

Lebeaux, David. 1988. Language acquisition and the form of the grammar. Doctoral Dissertation, University of Massachusetts, Amherst.

Mahajan, Anoop K. 1990. The A/A-bar distinction and movement theory. Doctoral Dissertation, Massachusetts Institute of Technology.

Mahajan, Anoop K. 1994. Towards a unified theory of scrambling. In *Studies on scrambling: Movement and nonmovement approaches to free word-order phenomena*, ed. Norbert Corver and Henk van Riemsdijk, 301–330. Berlin: de Gruyter.

Moulton, Keir. 2009. Natural selection and the syntax of clausal complementation. Doctoral Dissertation, University of Massachusetts at Amherst.

Nunes, Jairo. 2001. Sideward movement. *Linguistic Inquiry* 32:303–344.

Pesetsky, David. 2022. Spring 2022 lecture notes.

Poole, Ethan. 2017. Movement and the semantic type of traces. Doctoral Dissertation, University of Massachusetts at Amherst.

Saito, Mamoru. 1985. Some asymmetries in Japanese and their theoretical implications. Doctoral Dissertation, Massachusetts Institute of Technology.

Saito, Mamoru. 1989. Scrambling as semantically vacuous A'-movement. In *Alternative conceptions of phrase structure*, ed. Mark R. Baltin and Anthony S. Kroch, 182–200. Chicago: University of Chicago.

Saito, Mamoru. 1992. Long distance scrambling in Japanese. *Journal of East Asian Linguistics* 1:69–118.

Saito, Mamoru. 2003. A derivational approach to the interpretation of scrambling chains. *Lingua* 113:481–518.

Saito, Mamoru. 2004. Japanese scrambling in a comparative perspective. In *Peripheries: Syntactic edges and their effects*, ed. David Adger, Cécile de Cat, and George Tsoulas, 143–163. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Saito, Mamoru, and Hajime Hoji. 1983. Weak crossover and move α in Japanese. *Natural Language and Linguistic Theory* 1:245–259.

Sauerland, Uli. 1998. The meaning of chains. Doctoral Dissertation, Massachusetts Institute of Technology.

Sauerland, Uli. 2000. Two structures for English restrictive relative clauses. In *Proceedings of the Nanzan GLOW*, 351-366.

Sauerland, Uli. 2003. Unpronounced heads in relative clauses. In *The interfaces: deriving and interpreting omitted structures*, ed. Kerstin Schwabe and Susanne Winkler, 205–226. John Benjamins Publishing Company.

Sportiche, Dominique. 2006. Reconstruction, binding and scope. In *The Blackwell companion to syntax*, ed. Martin Everaert and Henk van Riemsdijk, volume IV, 35–93. Oxford: Blackwell.

Sportiche, Dominique. 2017. Relative clauses: promotion only, in steps. Unpublished manuscript, UCLA.

Takahashi, Shoichi, and Sarah Hulsey. 2009. Wholesale late merger: Beyond the A/A' distinction. *Linguistic Inquiry* 40:387–426.

Tsoulas, George. 1999. On long-distance scrambling in Korean. In *Harvard studies in Korean linguistics*, 217–302. Hanshin Publishers.

Tsoulas, George. 2003. Base generation of Korean long distance scrambling: A re-examination of the arguments and their consequences. Manuscript, University of York.

Tsoulas, George. 2007. Long distance scrambling and anaphora. In *Selected papers on Theoretical and Applied Linguistics*, ed. Eleni Agathopoulou, Maria Dimitrakopoulou, and Despina Papadopoulou, volume 17, 220–232. Aristotle University of Thessaloniki.