

Movement and crossover in three languages.

Abstract: In recent work, Bhatt and Keine (2019, 2023) have shown that Hindi local scrambling obviates Secondary Weak Crossover (SWCO) but not Secondary Strong Crossover (SSCO). The purpose of this paper is two-fold. The first is to show that this property is quite widespread, as it applies in a virtually identical manner to Italian Clitic Left Dislocation (CLLD) and to PP-fronting in English, facts that had gone unnoticed so far. The second is to try to understand why such a generalization is so pervasive as to characterize persistently widespread forms of dislocation across such diverse languages. It is argued that the generalization in question can be understood in terms of two assumptions: (i) the binding relation involved in SSCO and SWCO is *indirect* binding (i.e. it involves the so called E-type strategy) and (ii) WCO can be explained on the basis of the idea that traces and pronouns not only have a different syntax, but also a different semantics (i.e. the ‘bimodal’ approach to crossover – Buring 2004, Chierchia 2020). Under this view, the generalization unveiled by Bhatt and Keine reduces to a straightforward application of Principle C. A more speculative conclusion we will offer is that Principle C itself may be in turn eliminated in favor of a generalized form of Minimality, à la Rizzi. Comparisons with other proposals (including, of course, Bhatt and Keine’s) will be discussed.

1. On the interaction between movement and binding.

The following paradigm displays an interesting effect of dislocation in English:¹

- | | | |
|-----|--|----|
| (1) | a. i. Some committee member burst into tears at every student’s defense.
ii. * He ₃ burst into tears at every student ₃ ’s defense.
b. ? * His ₃ advisor burst into tears at every student ₃ ’s defense.
c. * At every student ₃ ’s defense he ₃ burst into tears.
d. At every student ₃ ’s defense, his ₃ advisor burst into tears.
e. Every student’s defense should be a happy moment for him. | ∇E |
|-----|--|----|

Let us slowly walk our way through these examples to appreciate why they are interesting. First, wide scope construal of the possessor embedded in the PP over the subject is clearly possible in (1a.i). But scope does not feed binding: the subject pronoun in (1a.ii) cannot be bound from the lower quantifier. This looks like a straightforward Strong Crossover (SCO) effect, with a twist: the potential binder is a possessor, embedded within a containing DP. Postal (1993) dubbed this configuration *Secondary SCO* (SSCO). For the time being, I will be using terms like SCO or SSCO

¹ The paradigm in (1)-(3) is modelled on the Hindi data discussed in Bhatt and Keine (2019, 2023), which constitute one of the main sources of inspiration for the present work.

I will alternate in my examples masculine vs. feminine pronouns as much as possible. When feasible, I will do so ‘block by block’. E.g. in (1) I only use masculine pronouns, in (2) only feminine, and so on.

descriptively, as an empirical characterization of classes of phenomena. Continuing with our paradigm above, sentence (1b) shows that an embedded quantifier also cannot bind a pronoun embedded within the subject: your prototypical Weak Crossover (WCO) configuration, again with the added twist that the potential binder is a possessor. By analogy with (1a.ii), we refer to (1b) as Secondary WCO (SWCO).² In (1c) and (1d), we front the adverbial PP. And here something very interesting happens: fronting does NOT ameliorate SSCO. Sentence (1c) remains severely deviant on the intended interpretation. However, it does ameliorate significantly SWCO (1d). The asymmetry between (1c) and (1d) is particularly puzzling in view of the fact, that possessor binding is in general possible when no crossover is involved, as is well known, and illustrated in (1e). The embedded quantifier in (1e) can reach out not only for scope but also for binding. It is not trivial to understand, then, why binding is possible in (1d), but not in (1c). The intuition is that it should have something to do with SCO: (1c) should be ruled out just like its source (1a.ii). But this clearly does not happen for its WCO counterpart (1d). The question is why; and this question is what the present paper addresses.

The examples in (2) and (3) illustrate how systematic the paradigm in (1) is and how it generalizes to other cases of binding out of a DP, such as inverse linking (3).

- (2) a. i. Some painting will definitely require a burglar alarm in every collector's house. $\forall\exists$
 ii. * She₃ will definitely require an alarm in every collector₃'s house. SSCO
 b. * Her₃ insurance will definitely require an alarm in every collector₃'s house. SWCO
 c. * In every collector₃'s house, she₃ will definitely require an alarm.
 d. In every collector₃'s house, her₃ insurance will definitely require an alarm.
 e. Every collector₁'s house is an insurance nightmare to her₁.
- (3) a. i. Some childhood friend made a speech at the wedding of every relative of mine $\forall\exists$
 ii. * She₃ made a speech at the wedding of every relative₃ of mine SSCO
 b. * Her₃ best friend made a speech at the wedding of every relative₃ of mine SWCO
 c. * At the wedding of every relative₃ of mine she₃ made a speech
 d. At the wedding of every relative₃ of mine, her₃ best friend made a speech
 e. The wedding of every relative of mine was a happy moment for her whole family.

While the status of (a)-(b) in (1)-(3) is familiar, the systematic amelioration of SWCO effects by PP-fronting vs. the total absence of such amelioration in SSCO configurations has never been systematically discussed, to my knowledge. The pattern in (1)-(3) does not only constitute an interesting conceptual puzzle. It also obviously raises learnability issues, as it often happens with linguistic data that are both intricate and systematic. It is hard to imagine on what basis the learner would converge on the judgements that obtain in the adult grammar.

The purpose of this paper is two-fold. The first is to show that the paradigm in (1)-(3) is not limited to a corner of English grammar. It involves, instead, a wide range of unrelated constructions across languages. In particular, I will show that it holds of Italian Clitic Left Dislocation (CLLD) – generalization which has also gone unnoticed so far. And it furthermore holds, in a virtually unchanged guise, for local scrambling in Hindi, as discovered and discussed in Bhatt and Keine

² The terms ‘Strong’ vs. ‘Weak’ were inspired by a perceived difference in the strength of the violation. We notate this in marking ‘*’ SCO violations and ‘*?’ WCO ones. The fact that there is a gradience in S- vs. WCO has sometimes been questioned. Recent experimental confirmation of the gradience in strength can be found in Ross et al. (2023). We won’t have anything to say in this paper about gradience.

(2019, 2023). The crosslinguistic robustness of the paradigm suggests that the differential pattern of amelioration of S- vs. WCO is unveiling important properties of the grammar of binding and dislocation. Our second goal is to figure out what those properties might be. We will argue that the pattern of dislocation observed above follows from a very basic version of Principle C together with a ‘bimodal’ approach to crossover along the lines developed in Buring (2004) and Chierchia (2020). The bimodal approach to crossover is based on a sharp differentiation of the semantics of traces from that of pronouns (which on most approaches are unified). In fact, the pattern under consideration offers a pretty good testing ground for theories of crossover in general. Consider, for example, leftness - based theories, modern versions of which can be found in Shan and Barker (2006) or Bumford and Charlow (2023). While we don’t want to put the cart of theory comparisons in front of the horses of theory building, it is hard to miss that leftness based approaches may have a particularly hard time with the pattern in (1)-(3). The problem is that in the minimally contrasting pair (1c) vs. (1d), the possessor binders uniformly occur to the left of the pronoun at surface structure. If we adopt a movement analysis for PP fronting, the pronoun will occur to the left of the base position of its potential binder in both instances and hence, an account based on ‘leftness’ should rule out both sentences alike. If we don’t adopt a movement analysis and base generate PPs directly in the left periphery, the pronoun is to the right of the potential binder and both sentences should be good.

Let me give a preview of the crosslinguistic robustness of the paradigm under discussion. In Italian CLLD, constituents get fronted, leaving behind a clitic pronoun, obligatorily if the dislocated constituent is a direct argument (subjects and objects), optionally, if it is an oblique. CLLD dislocation is restricted to definites, specific indefinites and generics. The dislocated constituent has roughly the discourse function of a topic and hence CLLD resembles English topicalization. The following examples illustrate.

- (4) a. Gianni, tutti *(lo) amano
Gianni, every body (him) loves ‘Everybody loves Gianni’
b. A Maria (le) hanno dato un bel premio
To Maria pro to-her gave a nice prize ‘They gave a nice prize to Gianni’
c. *? Ogni studente Maria lo aiuta molto
every student, Maria him helps a lot
Intended: Maria helps every student a lot

The clitic is obligatory in (4a), optional in (4b); (4c) shows how CLLD of strong quantifiers is deviant. And here is the CLLD-based Italian counterpart of the paradigm in (1)-(3):

- (5) a. * Pro₁ ha presentato le proposte di ogni ricercatore₁.
(he₁) presented the proposal of every researcher₃
b. * Il suo₁ direttore di ricerca ha presentato le proposte di [ogni ricercatore]₁.
His₁ research director presented the proposals of every researcher₁
c. * [Le proposte di ogni ricercatore₁]₂ pro₁ le₂ deve controllare una per una
[the proposals by every researcher₁]₂ he₁ them₂ must control one by one
d. [Le proposte di ogni ricercatore₁]₂ il suo₁ direttore di ricerca le₂ deve controllare una per una.
[the proposals by every researcher₁]₂ his₁ research director them₂ must contro one by one.
For every researcher x, x’s proposals are such that x’s research director must control them one by

one.

- e. i. Qualche commissario è stato critico verso le proposte di ogni ricercatore ∀E
 Some committee member was critical towards every researcher's proposals
 ii. [Le proposte di ogni ricercatore₁] sono state illustrate da lui₁ personalmente
 The proposals by every researcher₁ were illustrated by him₁ personally

Sentence (5a) is a typical SSCO violation. Sentence(5b) a SWCO violation. In (5c), the object undergoes CLLD and the contained possessor DP cannot covary with a subject pronoun, though the sentence becomes grammatical if there is no co-variation. Sentence (5d) shows that co-variation becomes possible as soon as the pronoun is embedded within the subject position and does no longer C-command the trace of the dislocated constituent. (5e.i-ii) are there to show that facts about possessor binding and inverse linking are like in English: A low quantifier can scope over a subject and an embedded possessor can bind pronouns when crossover doesn't get in the way. Thus, the paradigm in (5) is wholly parallel to that in (1)-(3).

Finally, let us consider local scrambling in Hindi, which works similarly to Italian CLLD, but it involves no resumption and can target also strong quantifiers. The following data set, from Bhatt and Keine (2019, 2023) – B&K henceforth - displays the same pattern of grammaticality as the English and the Italian cases:

- (6) a. * [us-ne₁ [har larke-kii₁ behin-ko]₂ dāātaa]
 [he-ERG₁ [every boy₁-GEN sister-ACC]₂ scolded]
 For every boy x, x scolded x's sister
 b. * [us-ke₁ dost-ne] [har larke-kii₁ behin-ko]₂ dāātaa
 [he-GEN₁ friend-ERG] [every boy₁-GEN sister-ACC]₂ scolded
 For every boy x, x's friend scolded x's sister
 c. * [har larke-kii₁ behin-ko]₂ [us-ne₁ ___₂ dāātaa]
 [every boy₁-GEN sister-ACC]₂ [he₁-ERG₁ ___₂ scolded]
 For every boy x, x scolded x's sister
 d. [har larke-kii₁ behin-ko]₂ [us-ke₁ dost-ne] ___₂ dāātaa
 [every boy₁-GEN sister-ACC]₂ [he₁-GEN friend-ERG] ___₂ scolded
 For every boy x, x's friend scolded x's sister

Sentences (6a) and (6b) constitute SSCO and SWCO cases respectively. Sentences (6c-d) are their scrambled counterparts. Local scrambling of the object over the subject in a SSCO configuration remains deviant (6c). But the same local scrambling significantly ameliorates SWCO (6d). The facts about possessor scope and binding are, otherwise, as in English and Italian.

In sum, there can be little doubt that we are dealing with a steady pattern across many constructions involving overt fronting, constructions that are all topic oriented, but otherwise quite different in their syntax and semantics. Two things seem to be clear at the outset. First, with respect to dislocation, SCO and WCO come apart. Hence no theory that relies on a single constraint to rule them out (as, say, Reinhart's (1983) classical take) stands much of a chance at capturing this paradigm -- a point duly underscored in B&K. And second, the binding relation between the moved DP and the pronoun should per se be as unproblematic as canonical cases of possessor binding or inverse linking:

- (7) a. Every Englishman's mother loves him

b. Every Englishman's mother wants his happiness

So, the problem cannot be in the relation between the possessor and the pronoun; it must concern the relation between the pronoun and the trace of the moved constituent, as B&K again correctly point out. This won't be easy to handle for theories that do not have traces in the base positions of these constructions, like, for example, an approach that base generates PPs in English directly in the left periphery, or theories of CLLD where the relation between the fronted constituent and the pronoun in the base position is just one of resumption, without movement. In fact, it would also *prima facie* seem that traces would have to be fairly structured to be useful, as in, e.g., the following abstract representation of the Hindi example (6b) from B&K. of which we report here the English gloss (recall that Hindi is an SOV language):

(8) * [every boy₁'s sister]₂ [he₁ **every boy₁'s sister**]₂ scolded] = (6c)

Since Chomsky (1981), Principle C of the Binding Theory has become the standard way in which SCO violations are handled. In (8), the constituent in bold face, i.e. the lower copy, is the trace of the scrambled DP. The pronoun *he₁* appears to be in a 'Principle C' configuration with respect to the *every boy₁* part of the lower copy: an 'R-expression' winds up being coindexed with a C-commanding DP (namely *he₁*) in an A-position. B&K call the configuration in (8) 'Principle C-connectivity' and propose that it is responsible for the ungrammaticality of (6c). However, if we consider *how* Principle C derives SCO several issues arise. A fairly widespread and reasonable assumption is that traces undergo a 'trace conversion' that turns them semantically into variables (see, e.g., Fox 2002, 2003). Trace conversion consists in erasing the quantifier in its base position (where it is uninterpretable) and interpreting the rest as a description linked to the index of the trace. The following simple example is an illustration:

- (9) a. * He₁ saw everyone₁
 b. Everyone₁ [he₁ saw **everyone₁**] Quantifier Raising
 c. Everyone₁[he₁ saw the unique person z such that z = x₁] Trace Conversion

If we want the subject to co-vary with the prepositional object in (9), we must coindex them. If they are coindexed, QR + Trace Conversion and standard semantic interpretive principles enforce semantic co-variation between the two DPs in (9a). Principle C, by ruling (9a) as syntactically ill formed, blocks the only way of getting the subject bound from the object position in (9a).³ But how exactly is Trace Conversion or some variant thereof going to apply to configurations involving complex DPs and possessor binding, like (8)? It is not obvious. Are we going to need complex variables formed by multiple indices? And how does resorting to Principle C-connectivity impact on Italian style resumption? These are some of the questions we will have to address. We are going to try to make a case that a version of Principle C that *doesn't* involve complex traces suffices to account for the pattern under discussion. Our main hypothesis is simple: A'-movement

³ On this interpretation of the binding theory, co-indexing says nothing about coreference. Two counterindexed referential DPs like (a) can still co-refer.

(a) He₁ saw John₂

To rule out co-reference between a C-commanding pronoun and a referential DP in its C-command domain something like Rule I is needed (see, e.g., Reinhart 1983, Grodzinsky and Reinhart 1993, Heim 2007, a.o). We will come back to the role of Rule I later on, in Section 2.1. and in Section 4.2.

comes in two varieties, one which activates ‘Discourse Referents’ (DRs) at the landing site and therefore bleeds WCO, and one which doesn’t. A’-movement of both varieties can never bleed Principle C.⁴ We will speculate that this is so because Principle C is a manifestation of and can be subsumed under Minimality (roughly in the sense of Rizzi 1990) which applies to all sorts of movement (albeit in parametrized forms). We will also speculate on what drives this split in A’-dependencies (DR-introducing vs. non DR-introducing). But our account stands, we submit, even if these speculations turn out to be wrong, and it will help clarify the typology of movement in both its syntactic and its semantic aspects.

We are going to proceed as follows. What one says about the interaction of dislocation and crossover of the primary and secondary kind clearly depends largely on background assumptions about (i) possessor binding and (ii) crossover itself in general. Our take on both together with our main hypothesis will be laid out in Section 2, the background section. Then (Section 3) we are going to look at our three languages, starting with Italian, followed by Hindi and finishing with perhaps the hardest to come to grips with, namely English. In Section 4 we draw some comparisons, outline some further perspectives for the Binding Theory, and conclude. In the Appendix we offer a simple formal proof of the consistency of our proposal using dynamic semantics (leaving to others the burden of proving that similar effects can be obtained through other semantic frameworks).

2. A base line on possessor binding and crossover.

In this section we lay out our cards about possessor binding and crossover respectively. We also spell out in its basic form our main hypothesis. Broadly speaking the landscape of anaphora can be described as falling within three areas. The first is what any basic theory of quantifiers and pronouns must deliver, i.e. the syntax and semantics of sentences like (10a-b). This is often called ‘C-command’ anaphora, but we will refer to it here as ‘Basic Anaphora’; it encompasses what falls under the tradition of the ‘Binding Theory’ (BT) of Chomsky (1981) and much subsequent work. Crossover phenomena, such as those in (10c-d) were identified and studied within this tradition.

(10) Basic Anaphora/Direct Binding

- a. Every cat_i grooms its_i whiskers
- b. Which cat_i doesn’t groom its_i whiskers?
- c. * His_i professor failed to be present at every student_i’s defense.
- d. * At whose_i defense did his_i professor fail to present ___?

The second major category of quantification and binding phenomena is constituted by what has come to be known as ‘donkey anaphora’ or ‘non C-command’ anaphora. We will refer to it here as ‘Extended anaphora’; It concerns primarily indefinites; adverbs of quantification play a key role in it:

⁴ This is modulo ‘late merge’ phenomena for adjuncts, like those in (b) vs. (a), familiar from Lebeaux (1988), that are irrelevant to the issue at hand:

a. *[Whose claim that John is nice] did [he believe t]
 b. [Which story that John wrote] did [he like t]

Argument clause display Principle C effects (a); adjunct clauses do not (b).

(11) Extended Binding/Anaphora.

a. Cross sentential

I presented a new book₁ at our book club. Mary presented its₁ author's biography

b. Relative clauses

Every one who presented a book₁ presented also its₁ author's biography.

c. Conditionals and Adverbs of quantification.

i. If you run into a good book₁, you wish you could meet its₁ author.

ii. If you run into a good book, you rarely get to meet its author.

This is the traditional hunting ground of Situation Semantics, Dynamic Semantics and their derivatives.⁵ Finally, we have a 'residue' of what might be called 'Functional Anaphora' that includes phenomena that are not readily reduceable to Basic or Extended Binding, where pronouns pick out their antecedents as descriptions/functions of something else:

(12) Functional Anaphora/Indirect binding.

'Paycheck' anaphora:

a. John invested his bonus₁ in the company. Everybody else used it₁ for their debts.

b. i. Every employee who got his pay on time used it for his mortgage

ii. Every employee who got it on time used his pay for his mortgage

'Bach/Peters' sentences:

c. i. Every pilot₁ that shot at it₂ hit the mig₂ that was chasing him₁

ii. [every pilot that shot at it₂]₁ [t₁ hit [the mig that was chasing him₁]₂]

In all the sentences in (12), the antecedent, e.g., *his bonus* in (a) does not seem to be able to directly bind the pronoun. It looks like the pronoun goes 'proxy' for its antecedent, i.e., the pronoun stands for some definite description prompted by the context.⁶ As it is not obvious how to extend the treatment of Basic Anaphora or of Extended Anaphora to Functional Anaphora/Indirect Binding, it is very likely that a specific theory of the latter is going to be necessary. In fact, we may *not* wish to extend Basic Binding to Functional Anaphora, for these two areas have a very different profile. At the same time, having both direct and indirect binding will unavoidably raise the issue of how to slice the pie between the two. Be that as it may, out of prudence, let's work out a base line theory of indirect binding, so that we have something to fall back on when our favorite theory of binding reaches its outmost limits.

We will now sketch an approach to indirect binding that can give sufficient concreteness to our take on movement and crossover, but that is also meant to be tentative and negotiable in its details. Let us follow Postal (1969) and some of its more recent developments, like e.g. Elbourne

⁵ It is controversial whether Extended Binding should be considered direct anaphora, like Basic Anaphora, or not. On a Situation Theoretic take it typically would not be direct binding, because pronouns would be interpreted via functions from situations into individuals (or equivalently, as descriptions parametrized to situations). On a Dynamic Binding take Extended Anaphora would be a case of direct binding, because pronouns would be interpreted as a kind of bound variables.

⁶ Notice that this type of anaphora seems to be equally felicitous in the canonical right-to-left direction (e.g. 12a,b) as in the less canonical, cataphoric, left-to-right directions (12c,d). Functional Anaphora has been investigated within variable free semantics in e.g. Jacobson (2000). The proposal we will put forth shortly shares with Jacobson's proposal the treatment of pronouns as functions. It differs from it in that our approach is not variable free.

2.1. Possessor binding and E-type pronouns.

There are two main ways to treat possessor binding and binding from ‘inverse linking’ structures:

- | | |
|--|-------------------------|
| (15) a. Every Englishman’s mother loves him. | Possessor Binding |
| b. I just put every freshman’s paper in his mailbox. | Possessor Binding |
| c. The mayor of every city must be familiar with its problems. | Inverse Linking Binding |

The first way is by extracting at Logical Form the embedded quantifier out of the DP that contains it, adjoining it to the whole clause and having it bind simultaneously the trace and the pronoun (the direct binding approach):

- (16) [Every Englishman]_i [t_i’s mother loves him_i]

The other is to assume that the actual binder of the pronoun is not the embedded DP but the *embedding* one, i.e. the DP headed by *mother*, which would require no extraction. In this case, the pronoun has to be interpreted as an E-type pronoun (the ‘indirect binding’ approach). The interpretation of the pronoun is informally indicated in (17):

- (17) Every Englishman’s mother loves [her son(s)]⁹

The pronoun *him* here is interpreted as a covert description containing an anaphoric element (*her* in (17)) which *is* directly bound by the subject. The link between the subject and (the anaphoric element contained within) the E-type pronoun brings about an indirect co-valuation between the possessor and the pronoun in object position, which results in the intended truth-conditions.¹⁰ In what follows, I will give three reasons to go for the indirect approach for possessor binding.

For one thing, locality considerations militate against the direct approach. Such an approach requires extracting something out of a DP, where DPs are tendentially islands. It is interesting in this connection that a scrambling language like Hindi *disallows* overt extraction of the possessor.

- (18) a. Kal [Ram-kii behin-ne] Anu-ko dāātaa (B&K)
 Yesterday [Ram-GEN sister-erg] Anu-ACC scolded
 Yesterday Ram’s sister scolded Anu

Anvari (2024). Furthermore, I believe that E-type pronouns should be semantically number neutral, following Neale (1990). The reasons for this will become clear when we deal with possessor binding: see in particular example (17) below and fn. 9. Neale’s approach has been criticized by Kanazawa (2001); Chierchia (2022) addresses Kanazawa’s criticisms, by elaborating on recent theories of homogeneity for definite descriptions and arguing that singular morphology on pronouns forces a distributive interpretation of semantically number neutral descriptions. In the present context, these remarks are admittedly obscure promissory notes. But I see no way of delving further into these matters within the limits of this paper.

⁹ This is where one needs to assume that E-type pronouns are semantically number neutral, à la Neale (1990), if (17) is to be equivalent to *for every x, x’s mother loves x*. The number neutral object must be interpreted distributively. See Chierchia (2022) for how this can be obtained on the basis of recent accounts of homogeneity.

¹⁰ There may be a third way to deal with possessor binding which amounts to having a binary quantifier developed in Bumford and Charlow 2023; but I don’t understand it yet well enough to comment on it.

- b. * Ram-kii kal [____₂ behind-ne] Anu-ko dāātaa
 Ram-GEN yesterday [____₂ sister-erg] Anu-ACC scolded

There is, therefore, no reason to expect that possessor extraction is allowed covertly. Generally speaking, scope assignment mechanisms for quantified DPs are strictly local (see, e.g. Cecchetto 2004).

A second factor that may make us lean in favor of an indirect approach is simplicity (though admittedly simplicity matters are typically theory laden and hence never decisive). Consider a sentence like (19a):

- (19) a. The president spoke with some representative from every city
 b. i. [some representative from [every city]₁]₂ the president spoke with [**some representative from [every city]₁**]₂
 ii. every city₁ [**some representative from [~~every city~~]₁**]₂ the president spoke with [**some-representative from [~~every city~~]₁**]₂

In (19b) I sketch a possible derivation of the LF for (19a), assuming Quantifier Raising (QR) and the copy theory of traces. First, we apply QR to the whole prepositional object in (19b.i); then we subextract the DP-embedded quantifier in (19ii). Lower copies (i.e. traces) are marked in boldface. I erase from each copy the part that cannot be interpreted where they occur. This approach requires that we deal with complex traces containing two quantifiers (uninterpretable in their base position) and two indices, for which a non stipulative compositional interpretive procedure should be provided, not a trivial task. As we will see shortly, the indirect approach allows us to do away with complex traces of this sort.

A third reason to think that possessor binding is indirect comes from non-coreference. Consider the contrast in (20b-c), in light of Rule I, informally sketched in (20a). The ‘spirit’ of Rule I is that binding should be used if possible, and thus coreferential interpretations are generally not intended when binding is structurally available.¹¹

- (20) a. Rule I (informal)
 In a structure [... XP... YP...], if one of YP or XP can be replaced with a variable bound by the other, then co-reference between XP and YP is NOT intended.
 Examples:
 b. i. * He₂ likes John₂’s father
 ii. He₂/every man₂ likes his₂ father
 c. i. Her₂ mother loves Mary₂ very much
 ii. Her₂/every English woman₂’s mother loves her₂

There is a sharp contrast in acceptability between the coreferential interpretation of (20b.i) vs. (20c.i). Why? The deviance of (20b.i) is typically attributed to Rule I, informally stated in (20a): if we replace *John* in (20b.i) with something that can be bound, the result is well formed, on the

¹¹ I illustrate my point with respect to an informal version of Rule I only for simplicity. Fox (2000) proposes a related ‘Rule H,’ through which he derives an account of Dahl’s generalization for VP ellipsis. Büring (2005) carefully discusses the relationship between Rule I and Rule H and proposes a synthesis. All I have to say is neutral with respect to these developments and compatible with them. Marty (2017) provides an interesting way of formalizing Rule I and its derivatives in terms of an implicature based approach.

co-valuated reading. Hence, according to Rule I, coreference in (20b.i) is *not* intended. However, the same holds true in (20c): replacing *Mary* with a pronoun yields a good co-bound interpretation; and yet coreference here is possible. If the binding relationship between the subject and the lower pronoun in (20b.ii) is the same as that in (20c.ii), this contrast in judgement wouldn't make sense. But if the co-variation in (20b.ii) between the possessor and the pronoun is *not* direct binding, the reasons for the difference in judgement become quite clear: Rule I doesn't apply to indirect binding. Such a rule essentially states that an encapsulated grammatical mechanism (direct binding) is preferred over coreference, which is essentially a pragmatic interface strategy. But indirect binding also relies on the pragmatic interface (for the interpretation of the function) and hence has no clear advantage over a co-referential interpretation.

In light of these considerations (locality, complexity and coreference), we follow Buring (2004) and many others in adopting a version of the indirect approach.¹² The LF a sentence like (17) is as follows:

- (21) a. $[_{DP} \text{Every boy's mom}]_2 [t_2 \text{ loves } [_{DP} \text{him}_{E-2} \text{ boy}]]$
 b. $[_{DP} \text{him}_{E-2} \text{ boy}] \Rightarrow f_{\text{BOY}}(x_2) = \text{the boy(s) that are sons of } x_2$

In (21), I attach to E-type pronouns the prefix *E-* (for E-type) as a reminder of their special interpretation, but no particular importance or significance should be attached to this mnemonic notation. The anaphoric index of the pronoun is bound in the 'normal' way by an accessible antecedent (in the case at hand, the subject). We will call this the *direct* antecedent of the pronoun and we'll call the source of elision, the *indirect* antecedent of the pronoun.¹³ Notice that E-type pronouns do not obey Principle B of Chomsky's (1981) Binding Theory: the anaphoric index in (21) is arguably not free in its local domain. This is presumably because the pronouns in questions are interpreted as descriptions, and the function-variable necessary to interpret them creates a 'distance' sufficient to obviate Principle B. E-type pronouns will, however, be subject to Principle C. This turns out to be crucial to our treatment of SSCO. The fact that E-type pronouns are not subject to Principle B, but are subject to Principle C has an arguably reasonable explanation: The link created by indirect binding is semantic in nature and does not care about locality. Hence, only the principles of the Binding Theory that are *not* subject to locality, namely Principle C, matter to E-type anaphora.¹⁴

For the internal syntax of the complex DPs that host possessors and/or inversely linkable DPs, we follow a proposal put forth in May (1985) and adopted since by many, including Buring (2004). Namely, we assume that the internal DP is adjoined to the outer one as in (22)

¹² For a situation theoretic derivation of *every boy's mom loves him*, see Buring (2004).

¹³ Non E-type (C-command bound) pronouns have the same syntax as e-type ones. But instead of being interpreted as Skolemized choice functions they are interpreted as identity maps and the direct and indirect antecedents coincide. For example:

(a) Every cat likes its whiskers

(b) Every cat₃ likes [it₃ ~~eat~~]'s whisker's where [it₃ ~~eat~~] $\Rightarrow$ the cat identical to x_3

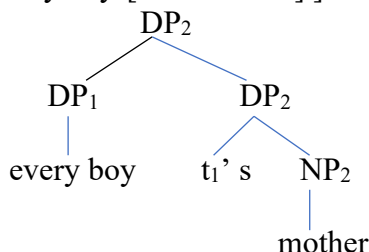
We are following here again Elbourne (2005); we are not going to adopt, however, Elbourne's situation theoretic semantics for E-type pronouns.

¹⁴ I am grateful to one of the anonymous referees for pointing out this line of explanation.

(22) a. Every boy's mother

b. i. $[_{DP} [_{every\ boy_1} [_{t_1}'s\ mother}]]_2$

ii.



c. $\lambda P \forall x [boy(x) \rightarrow P(ty.mother(y)(x))]$ 'the set of properties that every boy's mother has'

This creates a DP-internal chain $\langle everyboy_1, t_1 \rangle$ and a DP like (22a) winds up with two indices. The 'outer' one is associated with the head of the extended projection of the noun (i.e. *mother* in (22)); this index does the binding and constitutes the direct antecedent of the E-type pronoun. The inner one, obtained by adjoining the inner DP to the outer one, can only bind its own trace. The effect of this adjunction operation is to link the scope of the possessor to that of the outer DP.¹⁵ Putting it all together, we get the following derivation for (17):

(23) a. $[_{DP} Every\ boy_1\ t_1's\ mother]_2\ [t_2\ loves\ [_{DP}\ him_{E-2}\ boy]]$

b. $[t_2\ loves\ [_{DP}\ him_2\ boy]] \Rightarrow x_2\ loves\ f_{BOY}(x_2)$

c. $\|(a)\|^g = \lambda P \forall x [boy(x) \rightarrow P(ty\ mother(y)(x))](\lambda x_2. x_2\ loves\ f_{BOY}(x_2))$

$= \forall x [boy(x) \rightarrow ty . mother(y)(x)\ loves\ f_{BOY}(ty.mother(y)(x))]$

$= every\ boy's\ mother\ loves\ the\ boy\ or\ boys\ that\ are\ her\ children$

i.e., every boy is such that his mother loves him

In more general terms, indirect binding involves an anaphoric dependency between an indirect antecedent A (the possessor *every boy* in (23a)) and an anaphoric element B (the E-type pronoun *him₂* in (23a)) that goes through an independent binder C (the whole DP $[_{DP}\ Every\ boy's\ mom]_2$ in (23a)).¹⁶ Two features of this approach are important for our purposes. First, when a complex quantifier (i.e. a DP containing a quantified possessor) takes scope as in (24a) below, we can ignore (i.e. erase) the DP internal chain in the lower copy and use Trace Conversion for the remnant, as per the informal paraphrase in (23c), without any loss of information, because the DP-internal chain is interpreted in the upper copy:

(24) a. I admire every boy's mom

b. $[every\ boy_1[t_1's\ mom]]_2\ I\ admire\ [every\ boy_1[t_1's\ mom]]_2$

c. $[every\ boy_1[t_1's\ mom]]_2 \Rightarrow the\ mom\ identical\ to\ x_2$ [Trace Conversion]

The second key feature of this proposal is that indirect binding requires a formal link (i.e. coindexing) between an accessible antecedent and the E-type pronoun, as illustrated in (23a). It does *not* require any form of coindexing between the indirect antecedent and the pronoun: ellipsis

¹⁵ The interpretation in (22c) can be compositionally derived through a formally rather elegant pointwise extension of the rule for scope assignment to all types that 'end in t' (see, e.g., Rooth 1985).

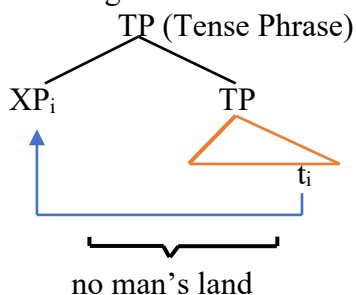
¹⁶ With respect to, e.g., (12a), the indirect antecedent is the definite description *his bonus*, the anaphoric element the E-type pronoun *it* and the independent binder is the subject *everyone else*.

handles agreement and feature transmission, as the case may be. Both of these two features turn out to be important in understanding SSCO effects, as we shall soon see.

2.2. The ‘bimodal’ approach to Crossover.

In general terms, crossover phenomena arise in the following configuration:

(25) The crossover configuration.



In (25) a phrase gets dislocated from its base position. No pronoun in the portion of the tree between the trace and XP’s landing site – ‘no man’s land’ - can be ‘caught’ (i.e., bound) as a result of this process, else ungrammaticality ensues. Reinhart (1983) famously proposed to capture this constraint through a C-command condition on binding: for a pronoun to be bound it must be C-commanded by an XP in an A-position. What is an A-position? The answer is controversial. Positions at which major θ -roles (Agent, Theme, Goal) are assigned constitute the most basic and agreed upon types of A-position. But the external subject position of raising verbs too constitutes an A-position. And, as we just saw, the landing site of local scrambling in Hindi and the site of CLLD in Italian display A-position properties in so far as anaphora is concerned.

The present paper is conservative in terms overall background assumptions on these matters. For SCO, we rely for the time being on the traditional take, based on Principle C of Chomsky’s (1981) binding theory. In due course we will consider a variant of the traditional approach under which Principle C is subsumed by (i.e. derived from) Minimality, i.e. the idea that movement to a position X should target the element (of the appropriate type/category) structurally closest to X (see Rizzi 1990). But our overall proposal doesn’t rest on this suggestion. The status of WCO is much more controversial in the literature. To make progress, we need a semantically explicit take and will adopt for this reason a version of the ‘bimodal’ approach proposed by Buring (2004) and further elaborated, in a different way, by Chierchia (2020). The rest of this section lays out the basics of this approach.

Buring (2004) develops the idea that there are two distinct operators devoted to binding, μ and β . The μ -operator is inserted at A’-positions and can only bind the trace of a moved constituent; the β -operator is inserted at A-positions and can only bind pronouns. Their semantics is as follows:

(26) Buring’s semantics for the μ/β -operators.

- a. μ : $\|\mu_i XP\|^g = \lambda u \|\|XP\|^{g[i/u]}$
- b. β : $\|\beta_i XP\|^g = \lambda u \|\|XP\|^{g[i/u]}(u)$

As is evident, the semantics of these two operators is very similar, both based on the λ -operator, with μ specialized in constituents of type t , and β specialized in predicates (i.e., functions that ‘end in t ’, of type $\langle a, t \rangle$, for any a). Let me illustrate how this accounts for WCO with an example:

- (27) a. i. Every cat is proud of its whiskers
 ii. Every cat β_3 is proud of its_3 whiskers
 b. i. Its whiskers ruined every cat
 ii. every cat μ_3 [its_3 whiskers ruined t_3]

In (27a) we can insert β next to the subject (say within vP), for that is an A-position and this correctly delivers pronoun binding. Per contrast, only the μ -operator can be inserted in (27b), because the QR landing site is your prototypical A'-position. And the μ -operator cannot ‘see’ indices on pronouns, even if the same digit is chosen. To emphasize this, and render visible the difference between traces and pronouns, I will henceforth underline indices on pronouns. The label ‘bimodal approach’ reflects the idea that trace vs. pronoun binding involve ‘separate channels’. WCO stems from the fact that the indices associated with scope assigning operations like QR can never directly bind those associated with pronouns: only the β -operator limited to A-positions can ‘link up’ two argument slots in a predicate.

Büring embeds his proposal within situation semantics, so that it can be extended to non C-command anaphora. Chierchia goes for the other main treatment of non C-command anaphora, namely dynamic semantics. But otherwise the ploy is very similar. For Chierchia (2020), there is an operator devoted to binding ‘ordinary’ Tarskian variables (= traces), and for that we might as well use λ . And one devoted to pronoun binding, where pronouns are viewed as ‘dynamic variables’. Dynamic variables are ‘introduced’ by an operator of type $\langle e, t \rangle$ of the form $\mathbb{N}(\)$. The formula $\mathbb{N}(u)$ declares that the individual u is stored at address $\underline{n}$, from which pronouns can pick it up. In fact, pronouns can *only* see individuals stored at underlined indices/addresses by some occurrence of the $\mathbb{N}(\)$ -operator. Assuming that λ is introduced at A'-positions and $\mathbb{N}(\)$ at A-positions, the treatment of crossover phenomena comes out just as on Büring’s proposal. Empirical differences between the two approaches derive from the fact that the two operators used by Chierchia obey different logical principles:

- (28) Tarskian variables
 $\lambda x_n [\alpha \dots x_n \dots] \dots x_n \dots$

The scope of the λ -operator in (28) is its C-command domain (i.e. the expression α , to which it is adjoined). Accordingly, the first occurrence of x_n in (28) is bound by λx_n , while the second is NOT and its value will be set independently of the value of the first. In contrast with this, the logic of dynamic variables is as follows:

- (29) a. [$\phi \ \mathbb{N}(u) \wedge \dots \omega_{\underline{n}} \dots \wedge \dots \omega_{\underline{n}} \dots \wedge \dots$]
 b. [$\alpha_{\text{TEST}} [\phi \dots \mathbb{N}(u) \wedge \dots \omega_{\underline{n}} \dots \wedge \dots \omega_{\underline{n}} \dots \wedge \dots]$]

In (29) I represent pronouns/dynamic variables as subscripted ω ’s; once u is stored at $\underline{n}$, each subsequent occurrence of $\omega_{\underline{n}}$ is going to refer to u . The scope of $\mathbb{N}(u)$ extends indefinitely across conjunctions as in (29). Its ‘frontier’ is constituted by the first ‘test’ α with which the sequence of conjunctions in ϕ combines (as indicated in (29b)). Simplifying somewhat, a test α is an operator

or quantifier that combines with an expression ϕ and checks the input context for whether it satisfies some property related to ϕ . If it does, it returns the input context unchanged, thereby wiping out all the info associated with the DRs in ϕ , which become inaccessible across α_{TEST} . Here is a standard exemplification of this.

- (30) a. i. [ϕ There was [a cat]₂³ on the mat; it₂ was black; I chased it₂ away; ...]
 ii. *? There [α wasn't] [ϕ a cat₂³ on the mat]. I saw it₂ on the kitchen table.
 b. i. [ϕ A friend of mine brought his guest₂³. She₂ was from Boston. ...]
 ii. *? [[α Every friend of mine] [ϕ brought his guest₂³]]. She₂ was from Boston.

The assumption here is that sequences of sentences are interpreted as conjoined (per asyndeton). And thus, sequences like (30a-b) exemplify the abstract structures in (29). In the first clause in (30a.i) or (30b.i), DRs associated with (in)definites are activated. The relevant info is passed on for an open-ended stretch, so long as the interpretation of the structure only involves conjunction. Pronouns can keep referring back to active DRs, adding incrementally information about them. When negation or a strong quantifier like *every friend of mine* intervenes, it ‘shuts off’ DRs in its scope and they become inaccessible for anaphoric pick up. It’s important to bear in mind that DRs introduced by $\lambda()$ can always bind across a conjunct, even in absence of C-command. Chierchia points out that this solves the mystery of how a DR in object position can antecede a pronoun in a high adjunct.

- (31) a. We’ll sell no wine before its time
 b. no wine λ_3 [[we’ll sell $\lambda_3(t_3)$] before its₃ time]
 c. $\neg \exists u \exists e$ [AG(e)(we) $\wedge$ [TH(e)]₃(u) $\wedge$ sell(e) $\wedge$ BEFORE(e)(ω_3 ’s time)]

The details of the implementation don’t matter much. The main point is that on a Davidsonian semantics, the modifier in (31a) applies to the event associated by the verb conjunctively, as in (31c). Thus, again, the semantics of (31b), namely (31c) exemplifies the structure in (29a), even if no conjunction is overtly present. A DR-introducing operator associated with the object will be able to bind a pronoun in the adjunct, in spite of the absence of C-command. As indicated by the LF in (31c), the QRed object will bind only its own trace, the semantic content of the trace is stored at index 3 by the DR-introducing operator, properly inserted at an A-position and this gets passed on across conjunction into the adjunct (see the Appendix for a formal execution of all this).¹⁷

Be that as it may, both Buring’s and Chierchia’s approaches rely on two operators, one for traces only, the other for pronouns only and yield a similar treatment of WCO. Both these approaches can and need to be extended along the following lines:

- (32) The Topic Condition.
 DR introducing operators are allowed at topic like positions

This is a stipulation. But a very natural one: Topics highlight constituents about which discourse unfolds and further information is added, and this requires topics to endow their targets with the

¹⁷ In fact, an anonymous referee suggests encouraging the reader who feels so inclined to take a look at the Appendix at this point, as that may significantly enhance understating the interplay between static and dynamic variables.

potential for antecedent pronouns. The main thesis of the present paper is that the Topic Condition, is the *only* stipulation we need in order to account for the complex behavior of the three kinds of fronting we are trying to understand.

3. The crosslinguistic reach of (S)SCO and (S)WCO.

In the present section we see how the approach outlined in Section 1 accounts for interaction of movement and crossover across three unrelated languages in a uniform manner. Going through these phenomena will allow us to flesh out more the proposals sketched in Section 2.

3.1. Italian CLLD and crossover.

With the background on indirect binding and crossover in place, let us start looking at CLLD in Italian. Its basics were sketched in Section 1: A constituent is moved to the left periphery, leaving behind a clitic obligatorily for direct arguments (subject and object), optionally for indirect arguments and adjuncts. The semantics of CLLD is that of topic formation. Some of its key properties are listed here, along with exemplifications thereof:

(33) Properties of CLLD

- i. If a DP is moved, it must be replaced in the original site obligatorily by a clitic; (33a).
- ii. When a subject is dislocated, the clitic is small *pro*; (33b).
- iii. If the moved element is a PP the clitic replacement is optional; (33c).
- iv. The dislocated XP and the corresponding clitic/gap must be in a local configuration. I.e. CLLD is sensitive to (strong) islands; (33d).
- v. The dislocated phrase is subject to Principle C effects; (33e).
- vi. Only definite and indefinite DPs can undergo CLLD. Strong quantifiers cannot; (29f) = (4c).

Examples

- a. **Maria**, Leo **la** incontra domani.
Maria Leo her meets tomorrow ‘Leo will meet Maria tomorrow’
- b. **Leo, Maria** [**pro la** incontra domani]
Leo, Maria [pro her meet tomorrow] ‘Leo will meet Maria tomorrow’
- c. **In palestra** Leo (**ci**) va volentieri.
To the gym Leo there goes with pleasure ‘To the gym Leo likes to go’
- d. ***Maria**, ho visto Leo prima che **la** incontrasse.
Maria (I) have seen Leo before that (he) her meet-SUBJ
Intended: ‘As for Maria, I’ve seen Leo before he met her’
- e. ***Il libro di [Leo]_i**, *pro_i* l’ha riletto volentieri.
The book of Leo_i (he_i) it has re-read with pleasure
Intended: ‘Leo reread his book with pleasure’
- f. *? Ogni studente Maria lo aiuta molto.
every student Maria him-helps a lot. Intended: ‘Maria helps every student a lot’

As can be gathered by this sketchy summary, CLLD is a rich and complex construction, to which we won’t be able to do justice within the limits of the present paper. Our focus is going to be on

Cecchetto, building on Torrego's (1994) work on clitic doubling, adopts a 'Big DP' analysis of CLLD, according to which the clitic and its double originate as sisters in a complex DP configuration such as the one in (34a). For a sentence like (34b), this gives us a base structure like (34c).

- The Big DP moves towards its final position as schematically represented in (34d-e). First, the Big DP, base generated in the canonical object position, A-moves as a whole to an intermediate Tense projection (say the host of the past participle *incontrata*), where it triggers Object Agreement, a general characteristic of object clitics in Italian (34d). Then, the Big DP is ‘pulled apart’: the clitic moves onto a higher T-position (where eventually the Aux is spelled out). The remnant of clitic movement, i.e. the double, is attracted to the Spec of the Top head. While the details of this derivation are not so crucial for us, we may want to observe that the syntactic attractor on the Top head is viewed here as some index (or indexed feature) which I am marking as a superscript on the Top head. On the semantic side, such a superscript is to be viewed as the $\mathbb{A}()$ -operator, which activates the address 3 and stores into it whatever is its Spec (cf. the Appendix for details). This essentially means that the Top head, besides activating whatever presuppositions/alternatives associate with ‘topichood’ (which we are not dealing with here), activates a fresh discourse referent (a dynamic variable) which can act as an antecedent not only for the clitic *la*, but for any accessible pronoun in its C-command domain, like for example:

- (35) Maria₃ Top³ suo₃ padre [la₃ t_{DP3}] ha incontrata t_{BIG DP3} ieri.
 Maria₃, her₃ father her₃ has met yesterday ‘Maria’s father met her yesterday’

In (35), Maria activates a DR that antecedes the pronoun *suo* in subject position. This is how WCO is overridden. Topics set up pivots about which novel information unfolds.

Further details of this specific implementation are the following. All DPs are indexed. In the spirit of the bimodal approach to crossover, the indices on non pronominal DPs (quantified or referential) and their traces correspond to (= are interpreted as) static Tarskian variables (marked as *non-underscored* subscripts). The indices on pronouns correspond to dynamic variables (marked as *underscored* indices). Activators of DRs are the Thematic Heads (like Agent, Theme, etc.) and heads like Top. A superscripted index on Top activates a DR and syntactically serves as an attractor for movement. Note that a dynamic index (*underscored*) may serve as an attractor for a static one, because they share a numerical value (i.e. ultimately, a syntactic feature), even if the semantics of underscored vs. non underscored indices is different. So in (35) Top³ attracts the DP with the same index hosted in the Big DP in object position. The static index on the DP *Maria* gets covalued in Top with the dynamic one, according to the following general semantics:

- (36) a. $[XP_n \text{ Top}^n \Phi] \Rightarrow [{}^n(XP) \wedge \lambda_n \Phi(XP)]$
 b. Example: Maria₃ TOP³ [her₃ father ... met t₃] $\Rightarrow$
 i. ${}^3(Maria) \wedge \lambda_{x_3} [\text{her}_{\bar{3}} \text{ father ... met } t_3](Maria) = {}^3(Maria) \wedge [\text{her}_{\bar{3}} \text{ father ... met Maria}] =$
 ii. ${}^3(Maria) \wedge \exists e[AG(e)(\omega_3\text{'s father}) \wedge TH(e)(Maria) \wedge \text{meet}(e)] =$
 iii. $\exists e[AG(e)(\text{Maria's father}) \wedge TH(e)(Maria) \wedge \text{meet}(e)] =$
 c. There is an event and an active DR 3 linked to Maria such that her father is the agent of e and Maria herself is the patient of e, e a meeting event.¹⁸

Semantically, Top takes two arguments, roughly the topicalized constituent (moved out of Top’s complement) and a property (built by λ -abstraction from the clause associated with Top’s complement). The superscript n is interpreted at the same time as the DR-introducing operator and as an abstractor on its sister. Notice the smooth interplay between static and dynamic components in example (36b), where we use English as a metalanguage for LF: The pronoun *her* embedded in subject position gets linked to the Topic in virtue of the introduction of DR associated with the topic position. The topic itself is then λ -ed back into the trace position. The λ -operator does not see or interfere with the pronoun.

I have omitted so far a detailed internal analysis of the big DP in the trace position. But the missing details are easy to fill in. The pronoun is a regular pronoun which is presupposed to be identical to the trace of the double. The doubling construction, in other words, triggers a presupposition that links the

¹⁸ While the LF and interpretive procedures are accurate, the reader will be able to check all the details of how it delivers only after having looked at the formal Appendix.

This approach can be readily extended to indefinites.

(a) Una mia amica suo padre la ha abbandonata a 10 anni
 A friend-FEM of mine her dad her has abandoned at 10 years (of age)
 ‘A (female) friend of mine was abandoned by her father when she was 10’

(b) $[XP_n \text{ Top}^n \Phi] \Rightarrow \exists u [XP(u) \wedge {}^n(u) \wedge \lambda_n \Phi(u)]$

I’m assuming in (b) that indefinites have (also) a predicative interpretation of type $\langle e, t \rangle$, which is being tapped in CLLD. There are interestingly intricate matters of scope in connection with CLLD of indefinites that we can’t get into within the limits of the present paper, we refer the interested reader to Cecchetto (2000).

pronoun to the double. Let me illustrate by focusing on certain details of our example. Zero in on the original site of the big DP in object position, represented in (37a):

- (37) a. [... met [_{DP} **her**₃ **Maria**₃] ...]
 where **her**₃ is the trace of the clitic pronoun, and **Maria**₃ the trace of the double *Maria*
 b. [_{DP} **her**₃ **Maria**₃] $\Rightarrow$: $\omega_3 = x_3$. ω_3
 in words: [_{DP} **her**₃ **t**₃] denotes whatever is stored at index 3, if that is identical to $g(t_3)$, i.e. Maria. The notational convention adopted here is the one standardly used in presupposition Theory. In $:\beta.\alpha$, β is α 's presupposition:
 c. $||: \omega_3 = x_3$. $\omega_3||^g = ||\omega_3||^g$, iff $||\omega_3||^g$ is the same as $g(x_3)$.

The approach just sketched derives without further ado the correct pattern for Principle C effects:

- (38) a. i. Gianni₁, Maria₂ [pro₁ l₂' ha incontrata ieri per la prima volta]
 Gianni₁, Maria₂ [pro₁ her₂ met yesterday for the first time]
 'Gianni met Maria yesterday for the first time'
 ii. Gianni₁ TOP¹ Maria₂ TOP² [[pro₁ t₁] met [her₂ t₂] yesterday] LF of (a.i)
 b. i. * Maria₁ [ieri lei₁ aveva detto che [[pro₁ t₁] sarebbe partita presto]]
 Maria₁ [yesterday lei₁ had-IMP said that [[pro₁ t₁] would have left early]
 'She had said that Maria would have left early'
 ii. * Maria₁ TOP¹ [she₁ had said [that [pro₁ t₁] would have left early]]¹⁹ LF of (b.i)

In (38a) we first dislocate the object, leaving behind the clitic *la*₂; then we dislocate the subject further to the left, leaving behind *pro*₁. The trace that is *pro*₁'s sister does not offend Principle C, because it is not A-bound: the Topic position is still an A'-position, albeit one at which a DR is introduced. Are we just playing with labels? We will come back to this point shortly. Notice meanwhile how the well formed (38a) contrasts with the ill formed (38b). Here the problem is the embedded trace *t*₁, which is A-bound by *lei*₁.²⁰ Notice that besides violating Principle C, the structure in (38b) also violate minimality: *lei*, when coindexed with Top, 'intervenes' between the probe (Top) and the goal *Maria*. So, structures like (38b) can be ruled out in two ways. This is significant and may pave the way to a general re-interpretation of Principle C in terms of Minimality, a line of inquiry we will explore in Section 4. So: no this is not a matter of playing with labels; even if we were to characterize Topic positions as A-positions, there would be an independent account of the ungrammaticality of (38b) in terms of minimality.

¹⁹ This sentence is deviant if *lei* is interpreted deictically (by pointing to Maria); however, as Cecchetto (p.c.) reminded me, it becomes acceptable on a 'hanging topic' kind of reading:

(b) (Quanto a) Maria, ieri lei ha detto che sarebbe partita presto

(As for) Maria, yesterday she said that (she) would leave

Perhaps hanging topics have a different structure than CLLD, one that doesn't involve movement. That would explain why on a hanging topic reading, Principle C effects do not arise.

In this connection, Cecchetto (2000) discusses how sentences of this sort are also improved by focus. For example:

(b) Maria, solo LEI può farcela alle prossime elezioni

Maria, only HER can make it at the next election.

Cecchetto puts forth reasonable proposals on this score, which are consistent with the general line adopted here.

²⁰ Sentence (38b) becomes grammatical if *lei* is understood as referring to some individual other than Maria.

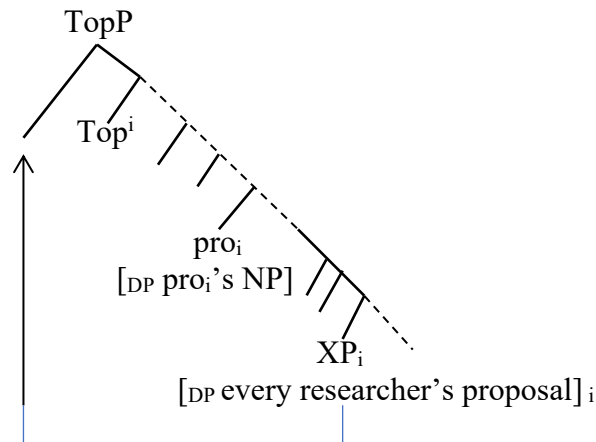
Let us now go back to the effects of CLLD on SSCO (which remains deviant under CLLD) vs. SWCO (which is ameliorated under CLLD). The key examples are repeated here:

- In both (39a) and (39b) the dislocated constituent is extracted from the object position. In (39a) the trace of the dislocated constituent is C-commanded by the subject pronoun and the possessor cannot co-vary with it. In (39b), the trace of the dislocated constituent is not C-commanded by the pronoun *suo* and the possessor can covary with it. Why? We've argued on independent grounds that possessor binding is a case of indirect binding. So, the LF for (39a) – again using English as a metalanguage – is (40a), where the subject *pro* is construed as an E-type pronoun. The only way to get the possessor to covary with the subject *pro* is by coindexing the whole DP [*every researcher_i's proposals*]₂ with it. But while there's nothing wrong with this on the semantic side of things, (40a) constitutes a Principle C violation, with the trace *t*₂ C-commanded by *pro*_{E-2}; or, equivalently, it yields a minimality offense with *pro*_{E-2} intervening between the attractor *Top*² and the Big DP in object position.

- Summing up, in CLLD, a Top head ‘attracts’ a DP coindexed with it, as in (41).

20

(41) a.



- b. XP_i is the closest target for TOP^i iff TOP^i C-commands XP_i and there is no bearer of index i ZP_i such that TOP^i C-commands ZP_i and ZP_i C-command XP_i .

For co-variation with a possessor to be possible, an (E-type) pro_i would have to be co-indexed with the trace of XP_i , when XP_i moves to Top^i . But then either Principle C is violated or Top^i would have to fail to attract the closest possible constituent with index i , a minimality violation (where the definition of ‘closest constituent’ is as in 41b). If, on the other hand, pro_i **does not** C-command the trace of the moved constituent, the syntactic obstacle is eliminated. By the Topic Condition, Top hosts DR-introduction. Amelioration of WCO follows from this. We thus arrive at a relatively simple and arguably principled account of a highly complex empirical paradigm that holds in parallel ways for basic and for secondary cases of crossover.

3.2. Hindi local scrambling and crossover.

There are some clear differences between CLLD in Italian and Local Scrambling in Hindi (henceforth, LS). To begin with the most obvious, Hindi lacks Romance style clitics. A second interesting difference is that long distance CLLD (which is subject to strong islands) behaves just like local CLLD with respect to anaphora; in contrast with this, according to B&K, local and long distance scrambling in Hindi do not pattern alike. Yet another difference concerns the dislocation of DPs headed by strong quantifiers like *har* ‘every’, which is marginal or impossible in CLLD, but unproblematic in Hindi (examples to follow shortly). Finally, Hindi local scrambling can happen sentence medially. While Italian too admits of sentence medial dislocations, they do not involve clitics and hence cannot be readily assimilated to CLLD.²² While keeping these differences well in mind, it remains plausible,

²² The following are examples of clause internal scrambling in Hindi, from Dayal Srivastav (1994, p. 237), who credits it to Gurtu (1992):

(a) Raam-ne Mohan-ko kitab dii [Subj IO DO V]

Ram-ERG Mohan-DAT book gave
Ram gave Mohan a book

(b) Raam-ne kitab Mohan-ko dii [Subj DO IO V]

Raam-erg book Mohan-DAT gave

Italian also allows similar alternations:

(c) Gianni ha dato un bel libro a Maria

in fact a priori desirable that CLLD and LS in Hindi have a syntax that at some level of abstraction is very similar. Broadly speaking, we are dealing in both cases with a Topic bound kind of movement.²³ So abstracting away from the differences we just mentioned (and whatever drives them), we will assume that LS in Hindi is also driven by a topic-like head Top^a that can be generated to the left of the clause (and, in Hindi, also at subclausal units) and syntactically attracts to its Spec the constituent with index n ; such a head introduces a novel DR linked to n . The differences between CLLD and LS can then be viewed as independently necessary (if poorly understood) parameters that endows Top with slightly different features across languages. Such features may impact, e.g., (i) how local topic movement is, (ii) how definite a topicalized constituent must be, (iii) where in the extended projection of V can Top occur, etc. Differences in locality, depth of embedding, or definiteness are familiar loci of microvariation. For example, English only allows strictly referential DPs to undergo topicalization (see., e.g., Lasnik and Stowell 1991), while Italian allows also indefinites of both the referential and the generic kind to do so. Hindi is a further step along this quasi continuum in allowing universally quantified DPs to be topics. Ultimately, these variations may well be related to the many slightly different notions of Topichood discussed in the literature, e.g.: Hanging topics (Cinque 1977), Contrastive topics (Büring 2003), Shifting topics (Zubizarreta 2006), all with slightly distinct semantics and syntax (Alexiadou 2006). We will not be saying anything here about these differences, happy to go along with whatever turns out to be the right typology of topic positions.

Let us go over the main Hindi data, some of which were previewed in the introduction. They are for the most part directly drawn from B&K, who credit especially Gurtu (1992) and Dayal Srivastav (1994) as direct antecedents.

(42) WCO effects and their amelioration under LS

- a. * [us-kii₁ behin-ne] [har larke-ko]₁ dāātaa
[he-GEN₁ sister-ERG] [every boy₁-ACC] scolded
His₁ sister scolded every boy₁ ²⁴
- b. [har larke-ko]₁ [us-kii₁ behin-ne] ___ dāātaa
[every boy-ACC] [he-GEN sister-ERG] scolded
His₁ sister scolded every boy₁
- c. * [us-kii₁ behin -ne] kis-ko₁ dāātaa?
[he-GEN₁ sister-ERG] who-ACC scolded ‘who₁ did his₁ sister scold t?’
- d. kis-ko₁ [us-kii₁ behin -ne] ___ dāātaa?
who₁ [he-GEN₁ sister-ERG] ___ scold ? ‘who_x is such that x’s sister scolded x?’
- SCO effects.
- e. * [har larke-ko]₁ [us-ne/apne aap-ne₁] dekhaa
[every boy-ACC] [he-ERG/self-ERG] saw. ‘He₁/himself₁ saw every boy₁’

-
- Gianni gave a good book to Maria
(d) Gianni ha dato a Maria un bel libro
Gianni gave to Maria a good book

However, in standard Italian, these alternations do not display clitic doubling. We won’t be able to delve more deeply into clause internal forms of dislocation, other than to note that it looks like it may be a further point of divergence between Hindi and Italian.

²³ Cecchetto (1994) is one of the first attempts at unifying CLLD with scrambling of the Germanic type, which in turn is not too dissimilar from scrambling of the Hindi type.

²⁴ In the English gloss I mark pronouns uniformly as masculine to enhance readability, but please bear in mind that pronouns in Hindi are unmarked for gender.

The ungrammaticality of (42a) is a typical WCO effect with quantified DPs in a low position being unable to bind the possessive pronoun in the subject DP; LS scrambling clearly ‘repairs’ it (42b). The pair in (42c-d) illustrates the same with respect to questions. The object wh-word cannot bind into the subject in (42c). If we scramble leftwards the wh-word *kis-ko* as in (42d), the question becomes well formed under the bound reading. This contrasts with (42e), which constitutes an attempt at repairing in a parallel manner a SCO violation: in (42e) we scramble an object quantifier over a subject pronoun (considered in both a reflexive and a non reflexive form). The result is ungrammatical. However, it must be added that such ungrammaticality could be accounted for in terms that are independent of SCO as such. Sentence (42e) could be bad on the intended reading because of a Principle B violation, the pronoun *us-ne* is ‘too close’ to its antecedent, or because the reflexive pronoun *apne-aap* is inherently subject oriented. However, the SSCO and SWCO cases from Section 1, repeated here, clinch the case for the relevance of crossover:

- (43) a. * [us-ne₁ [har larke-kii₁ behin-ko]₂ dāātaa] SSCO
 [he-ERG₁ [every boy₁-GEN sister-ACC]₂ scolded]
 For every boy x, x scolded x’s sister
- b. * [us-ke₁ dost-ne] [har larke-kii₁ behin-ko]₂ dāātaa SWCO
 [he-GEN₁ friend-ERG] [every boy₁-GEN sister-ACC]₂ scolded
 For every boy x, x’s friend scolded x’s sister
- c. * [har larke-kii₁ behin-ko]₂ [us-ne₁ ___₂ dāātaa]
 [every boy₁-GEN sister-ACC]₂ [he-ERG₁ ___₂ scolded]
 For every boy x, x scolded x’s sister
- d. [har larke-kii₁ behin-ko]₂ [us-ke₁ dost-ne] ___₂ dāātaa
 [every boy₁-GEN sister-ACC]₂ [he₁-GEN friend-ERG] ___₂ scolded
 For every boy x, x’s friend scolded x’s sister
- e. [Ram-kii₁ behin-ne] us-ko₁ dekhaa
 Ram-GEN sister-ERG him-ACC saw ‘Ram₁’s sister saw him₁’

The paradigm in (43) shows that Hindi displays both SSCO and SWCO effects: when no movement is involved a possessor within an object cannot bind into the subject position. Movement makes SWCO grammatical (43d), but not SSCO (43c). The ungrammaticality of (43c) cannot be attributed to the excess of closeness between *us-ne* and its antecedent, because in terms of antecedent-pronoun locality, (43c) is fully parallel to the grammatical (43e). The ungrammaticality of (43c) must thus be due to the presence of the trace of the dislocated constituent.

This paradigm is exactly what our proposal about Topic movement derives. Let us see how, by considering the LFs the scrambled SCO case (42e) and the scrambled WCO one (42b), according to our proposal. In displaying the relevant LFs, we will use English as a metalanguage and we will also follow Fox’s trace conversion.

- (44) a. * [every boy]₁ Top¹ [he₁ [~~every~~ boy]₁ scolded]] LF of (42e) SCO
 b. [every boy]₁ Top¹ [his₁ sister [~~every~~ boy]₁ scolded]] LF of (42b) WCO

The Top¹ head attracts the constituent with index 1 and (re-)introduces a DR linked to it in a position where it is accessible to the subject. Thus, in so far as semantics goes, anaphora from the

dislocated position is possible, as the grammaticality of (42b/44b) indeed shows. However, the structure in (44a), necessary to obtain binding of the subject, is syntactically ill-formed. It constitutes an open and shut Principle C violation. Alternatively, or in addition, it also constitutes a minimality violation. The semantics for Top^n in Hindi is identical to that for Top^n in Italian, modulo the necessary type theoretic adjustments to accommodate quantified DPs. Here is how such adjustments go:

- (45) a. $[\text{XP}_n \text{ Top}^n \Phi] \Rightarrow \text{XP} (\lambda u (\text{Top}^n(u) \wedge \lambda n \Phi(u))$
 b. Example: every boy₁ Top^1 $[[\text{her}_1 \text{ sister}] \text{ t}_1 \text{ scolded}] \Rightarrow$
 $\forall x \text{ boy}(x) \rightarrow \text{Top}^1(x) \wedge \lambda x_1 \exists e [\text{AG}(e)(\omega_1 \text{'s sister}) \wedge \text{TH}(e)(x_1) \wedge \text{scold}(e)](x) =$
 $\forall x \text{ boy}(x) \rightarrow \text{Top}^1(x) \wedge \exists e [\text{AG}(e)(\omega_1 \text{'s sister}) \wedge \text{TH}(e)(x) \wedge \text{scold}(e)] =$
 $\forall x \text{ boy}(x) \rightarrow \text{Top}^1(x) \wedge \exists e [\text{AG}(e)(x \text{'s sister}) \wedge \text{TH}(e)(x) \wedge \text{scold}(e)] =$
 $\forall x \text{ boy}(x) \rightarrow \exists e [\text{AG}(e)(x \text{'s sister}) \wedge \text{TH}(e)(x) \wedge \text{scold}(e)]$
 c. For every boy x, there is a scolding event of which x's sister is the agent and x the theme.

In (45b) we provide, for illustration, a step-by-step sample derivation of the WCO case, whose truth-conditional outcome is summarized in (45c). Questions (42c,d) work in a parallel way:

- (46) Questions and crossover
 a. i. Q^1 $[\text{his}_1 \text{ sister-SUBJ who}_1\text{-OBJ scolded}]$ Spell-out structure
 ii. $\text{who}_1 Q^1$ $[\text{his}_1 \text{ sister-SUBJ t}_1 \text{ scolded}]$ LF; WCO violation
 b. i. Q^1 $[\text{who}_1\text{-OBJ TOP}^1 [\text{his}_1 \text{ sister-SUBJ t}_1 \text{ scolded}]]$ Spell-out structure, WCO repair
 ii. $[\text{who}_1\text{-OBJ } Q^1 [\text{t}_1 \text{ TOP}^1 [\text{his}_1 \text{ sister-SUBJ t}_1 \text{ scolded}]]]$ LF

Structure (46a.i) represents the surface form of the unscrambled question. Following Dayal's (1996) groundbreaking work, we assume that the WH-word in (46a.i) is moved to the interrogative Comp at LF, as in (46a.ii), thereby creating a question meaning. Since WH-movement does NOT activate DRs, the pronoun embedded in the subject position has no sentence internal antecedent, and a WCO violation ensues. In (46b.i), on the other hand, first the WH-word undergoes movement to TOP, and gets scrambled sentence initially. Movement to TOP does create a DR, associated with the WH-word, thereby allowing the pronoun embedded within the subject to find a suitable antecedent. Then the WH-word undergoes further movement to Comp at LF and a question meaning is created. This eliminates the WCO effect.

Assuming that indirect binding works identically in Hindi and Italian, our results extend automatically to cases of SSCO and SWCO. First let us look at the LFs corresponding to (43c) and (43d) using again English as semantic metalanguage:

- (47) a. * $[\text{every boy}_1 \text{'s sister-ACC}]_2 \text{ Top}^2 [\text{he-ERG}_{E-2} \text{ boy}] [\text{every boy's sister-ACC}]_2 \text{ scolded}$
 b. $[\text{every boy}_1 \text{'s sister-ACC}]_2 \text{ Top}^2 [\text{he-GEN}_{E-2} \text{ boy friend-ERG}] [\text{every boy's sister-ACC}]_2 \text{ scolded}$

The only way to obtain co-variation in the SSCO sentence is to insert an E-type pronoun in subject

position and coindex it with the object as in (47a). While E-type pronouns have a different semantics from non E-type pronouns, their indices behave alike with respect to the Principle C and so (47a) constitutes a Principle C or a Minimality violation. In contrast with this, (47b) is syntactically well-formed. No trace is C-commanded by a coindexed DP and minimality is respected. The semantics of (47b) works like that of (46b), modulo the difference in meaning of the dislocated DP, which is more complex, namely as in (48a). The step-by-step derivation is shown in (48b). If you trust my computations (and you shouldn't), you can skip directly to the last line:

- (48) a. $\lambda P. \forall x \text{ boy}(x) \rightarrow P(\text{ty sister}(y)(x))$
 b. $\lambda P. \forall x \text{ boy}(x) \rightarrow P(\text{ty sister}(y)(x)(\lambda u. {}^2(u) \wedge \lambda x_2 \exists e[AG(e)(f_{\text{boy}}(\omega_2))'s \text{ friend}] \wedge TH(e)(x_2) \wedge \text{scold}(e)](u)) =$
 $\lambda P. \forall x \text{ boy}(x) \rightarrow P(\text{ty sister}(y)(x)(\lambda u. \exists e[AG(e)(f_{\text{boy}}(u))'s \text{ friend}] \wedge TH(e)(u) \wedge \text{scold}(e)] =$
 $\forall x \text{ boy}(x) \rightarrow \lambda u[\exists e[AG(e)(f_{\text{boy}}(u))'s \text{ friend}] \wedge TH(e)(u) \wedge \text{scold}(e)](\text{ty sister}(y)(x)) =$
 $\forall x \text{ boy}(x) \rightarrow \neg \exists e[AG(e)(f_{\text{boy}}(\text{ty sister}(y)(x)))'s \text{ friend}] \wedge TH(e)(\text{ty sister}(y)(x)) \wedge \text{scold}(e)]$

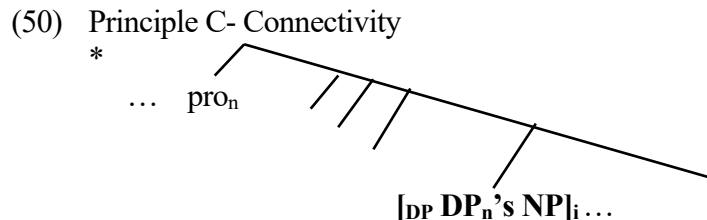
The E-type pronoun f_{boy} maps any u into u 's brothers. So the last line in (48b) winds up saying that for any boy x , the friend of the brother of x 's sister (= the friend of x himself), scolded x 's sister; f_{boy} 'undoes' the sister-function: for any x , the brother of x 's sister is x . This semantic composition may well be convoluted, but it follows from general principles. And it works.

If we are on the right track, CLLD and LS are virtually the same construction with respect to how dislocation and crossover interact. It is worth underscoring that core cases of Principle C are determined solely by the 'outer' index of a complex DP (cf. 49.b.ii). The internal structure of the trace doesn't play any role. Uninterpretable material in the base position is elided in simple cases as in complex ones:

- (49) a. i. he_1 saw everyone₁
 ii. everyone₁ he_1 saw **everyone**₁
 b. i. He saw everyone's father
 ii. [everyone₁ t_1 's father]₂ he_{E-2} saw **[everyone₁ t_1 's father]**₂

In (49b) the DP internal chain is interpretable only in the upper copy (where it makes the DP quantificational). In the lower copy it cannot be interpreted, just like the D cannot be interpreted in its base position in (49a) (whence its erasure). So (49b), i.e. SSCO, gets assimilated on the present approach to a *plain* Principle C violation. All that matters to get SSCO is just the index of the containing DP.

The present take goes in a different direction from B&K's proposal. As we know, they see in the Hindi pattern a case in favor of 'Principle C Connectivity'. In so far as I can see, the latter is to be conceived as a syntactic ban against a representation of the following sort:



The part in bold is a (complex) trace of a dislocated constituent. Structure (50) contains, B&K argue, a Principle C violation involving the pro_n and the (copy of the) possessor DP_n . As B&K insist that possessors cannot be extracted from the containing DP, possessor binding must be a form of indirect

binding. But if possessor binding is indirect, the index borne by the possessor DP has no bearing on how the pronoun gets to co-vary with the possessor. For it is not the possessor DP that does the binding: it is the whole containing DP that does. Coindexing or contraindexing the pronoun with the possessor has no semantic import. This makes holding on to complex traces kind of useless in the case at hand. Notice that this is not a matter of the eternal debate between syntax heavy/semantics light accounts vs. syntax light/semantics heavy ones. If possessor binding is indirect binding, a treatment like the one advocated in the present paper is the only game in town, in so far as I can see.

Things are different of course, if possessor binding and related phenomena are cases of direct binding after all. But then the problem becomes that a direct binding approach calls for a different syntax and/or a different semantics from either the one adopted here or the one adopted by B&K. One way to go direct, for example, involves pulling the possessor out of the DP that contains it and assigning the possessor scope independently of that of the containing DP. Under this view, possessor raising which is impossible in overt syntax in Hindi, would have to be possible at LF. This approach would warrant keeping track of complex traces in order to account for the observed pattern. Another way to go would involve treating a DP containing a possessor as a complex binary quantifier with two indices that bind into its trace sort of simultaneously. This requires working out a novel semantics and it remains to be seen what consequences that would have on syntax and the binding theory.

3.3. PP fronting in English and crossover.

Can the account we've been developing for Italian and Hindi be maintained for English? Broadly speaking, the answer is the same as for the Hindi case: it can, to the extent that the syntax of PP-fronting is sufficiently similar to that of CLLD. Addressing this question in the specific case of PP-fronting forces us to face the status of PP - as opposed to DP - dislocation, which makes things interestingly more complicated. The key minimal pair we want to deal with is:

- (51) a. * At [every student₁'s defense]₂, he₁ burst into tears
 b. At [every student₁'s defense]₂, his₁ advisor burst into tears

By now, we should be convinced that a base generation approach to PP fronting stands no chance at accounting for the contrast at hand. We must assume that the PP is fronted from a vP-internal position. Thus, we will have to have for, say, (51a) a surface structure such as (52a), where the fronted PP has a vP internal copy; QR out of the PP (i.e. assigning scope to the prepositional object) should then apply to (52a), yielding (52b):

- (52) S-structure and LF of the SSCO case (51a).
 a. * [PP at [DP every student₁'s defense]₂]₃ TOP³ [[DP he_{E-3} ~~student~~] burst into tears [PP at [DP **every student₁'s defense**]₂]₃]
 b. * [DP every student's defense]₂ [PP at [DP ~~every student₁'s defense~~]₂]₃ TOP³
 [[DP he_{E-3} ~~student~~] burst into tears [PP at [DP **every student₁'s defense**]₂]₃]

Structures like (52a-b) have several features we have to consider carefully. Two key ones are the following. First, PP fronting must be movement to a topic-like head TOPⁿ, that attracts the constituent with index n. In contrast with Italian or Hindi, Topⁿ in English can only attract strictly

referential arguments²⁵ and PP-adjuncts, maybe because PP-adjuncts too can be regarded as ‘referential’ – see below from example (55) on. By the Topic Condition, DRs can be introduced and semantically associated with the fronted PP; these DRs must be then formally linked to E-type pronouns in subject position, if the latter are to covary with the topicalized PP. The index on the topic-like head that syntactically attracts the PP is interpreted semantically in terms of the DR-introducing operator λ (). I spell out in (52a) for extra clarity also the internal structure of the pronoun: the ellipsis antecedent for it is the NP *student*. But the direct binder of the anaphoric index is the PP *at every student’s defense* itself, which does contain a quantified DP and hence is to be regarded as quantificational. What we have to determine is in which way the PP provides a direct antecedent for the E-type pronoun. We will see how that goes shortly. Meanwhile, the prediction/expectation is that structures like (52a-b) are syntactically ill-formed: the trace of the PP with index 3 winds up coindexed with a C-commanding pronoun, a Principle C violation. Or we can view it as a minimality violation: TOP³ fails to attract the structurally closest constituent with index 3. The WCO cousin of (51a), namely (51b), has no problem with either Principle C or minimality. Its LF is as follows:

(53) LF of the SWCO case (51b).

[_{DP} every student₂’s defense]₂ [_{PP} at [_{DP} ~~every student₁’s defense~~]₂]₃ TOP³
 [[_{DP} [_{he_{E-3}} student]₁’s advisor] burst into tears [_{PP} at [_{DP} ~~every student₁’s defense~~]₂]₃]

In (53) the PP trace is not C-commanded by a coindexed pronoun. To be duly noted: in (51)-(53) I am assuming that a PP like *at every student’s defense* can be coindexed with a DP like *he/his*. This is a necessary assumption for indirect binding: the index associated with the PP must connect with the pronoun. I don’t see any particular problem with this assumption.²⁶ The novelty of our proposal is that coindexing for indirect binding purposes IS regarded as subject to binding theoretic conditions. The line strokes in (49) eliminate the parts of each trace that cannot be interpreted in that position. Thus, in the dislocated PP, the DP-internal chain associated with the possessor is erased and the rest of the trace is interpreted according to Fox’s trace conversion, roughly as “the defense identical to x_2 ”. Moreover the lower trace of the whole PP is simply interpreted as a variable ranging over PP-meanings: the meaning of the fronted PP is just ‘ λ -ed’ back. I will shortly give full details of how it works out. Conceptually, each assumption is simple and independently motivated. But they pile up in ways that can be confusing.

Let us start with the PP as such. In a Davidsonian semantics, they are generally treated as event-modifiers and in adverbial position, they modify the main verb. A sentence like (54a) has the semantics in (54b).

(54) a. John cried at Mary’s defense

b. $\exists e [AG(e)(j) \wedge cry(e) \wedge at(e_{MD})(e)]$, where e_{MD} = Mary’s defense (an event)

²⁵ As in:

(a) This book I expect its author to disavow (Lasnik and Stowell 1991)

²⁶ It is independently necessary to assume that coindexing for binding purposes must tolerate mismatches of syntactic categories. For example, in Italian the possessive pronoun *suo* is morpho-syntactically adjective-like, not DP-like. And yet of course covariance with DPs works out as in English.

(a) [Ogni ragazzo]₃ litiga con i suoi₃ genitori
 [_{DP}Every boy-SG]₁ quarrels with the [_{AP} his-PL]₁ parents

The position and agreement pattern of the possessive *suo* in (a) is that of an adjective.

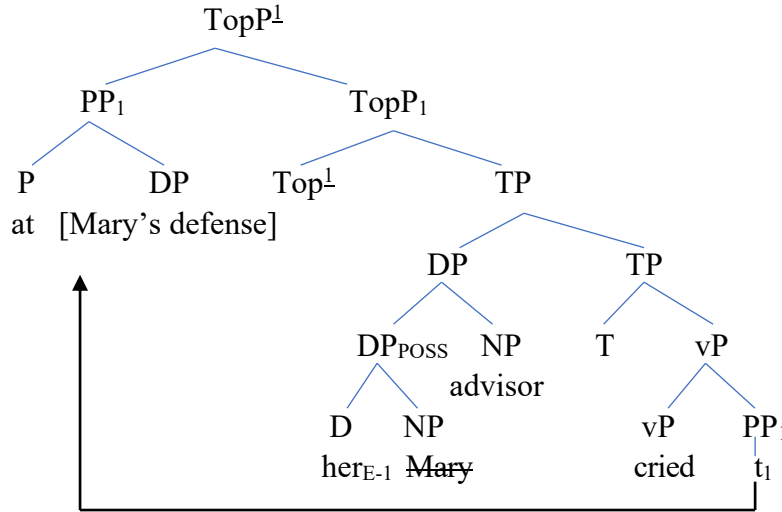
- c. $at(e_{MD})$ is a property of eventualities of type $\langle e, t \rangle$, true of the sets of events that take place at Mary's defense.²⁷

What happens when the PP is fronted? We expect that a DR gets introduced. It stands to reason that such DR is itself something contextually relevant of which $at(e_{MD})$ holds true. So, in simple cases the semantics associated with fronting of the PP comes to the following:

- (55) a. $[At\ Mary's\ defense]_1\ TOP^1\ John\ cried\ [PP\ at\ Mary's\ defense]_1$
 b. i. $\exists e' \cdot [at(e_{MD})(e') \wedge \lambda(e') \wedge \lambda P_1 \exists e [AG(e)(j) \wedge cry(e) \wedge P_1(e)] (at(e_{MD}))] =$
 ii. $\exists e' \cdot [at(e_{MD})(e') \wedge \lambda(e') \wedge \exists e [AG(e)(j) \wedge cry(e) \wedge at(e_{MD})(e)]] =$
 iii. $\exists e [AG(e)(j) \wedge cry(e) \wedge at(e_{MD})(e)]$

What the semantics in (55b.i) delivers is truth conditionally equivalent to (55b.iii), which tantamounts to semantically reconstructing the PP in its original site.²⁸ Things change if we introduce a (syntactically legit) pronoun in subject position:

- (56) a. i. At Mary's defense, her advisor cried
 ii. $[PP\ at\ [DP\ Mary's\ defense]]_1\ TOP^1\ [[DP\ [her_{E-1}\ MARY]'s\ advisor]\ cried\ t_1]$ ²⁹
 iii.



- b. $\exists e' \cdot [at(e_{MD})(e') \wedge \lambda(e') \wedge \exists e [AG(e)(f_{MARY}(\omega_1)'s\ advisor) \wedge cry(e) \wedge at(e_{MD})(e)]] =$
 $\exists e' \cdot [at(e_{MD})(e') \wedge \lambda(e') \wedge \exists e [AG(e)(f_{MARY}(e')'s\ advisor) \wedge cry(e) \wedge at(e_{MD})(e)]] =$
 $\exists e' \cdot [at(e_{MD})(e') \wedge \exists e [AG(e)(mary's\ advisor) \wedge cry(e) \wedge at(e_{MD})(e)]] =$
 $\exists e [AG(e)(mary's\ advisor) \wedge cry(e) \wedge at(e_{MD})(e)]$
 c. Semantics for PP-topicalization:³⁰
 $[PP_i\ TOP^i\ \Phi] \Rightarrow \exists e' PP(e') \wedge \lambda(e') \wedge \lambda_i \Phi (PP)$

²⁷ In fact, $at(e_{MD})$ would have to be 'multisorted'. In a phrase like *the people at Mary's defense* holds true of ordinary individuals, not events.

²⁸ Fronting may well have consequences for the discourse status of the PP (i.e. what kind of topic it may be); but we are not investigating that aspect of the construction.

²⁹ This is not the place to plunge into the syntax and semantics of Proper Names (PNs). For our purposes, we can assume that they are base generated in N (with a singleton set semantics, e.g., $[N\ John] \Rightarrow \lambda x.x = John$). This provides the basis for ellipsis in (52a.ii). They combine in D with a null ι and maybe undergo N to D raising a la Longobardi (1994).

³⁰ This semantics is identical to that for (specific) indefinites in Italian CLLD: cf. fn. 18 above.

We are now ready for our lead example (51b), repeated below as (57a). The only key difference between (57b) and (56a) is the presence of a quantified complex DP like *every student's defense*, in place of the referential one *Mary's defense*. The interpretation of the complex DP in (57a) is the familiar quantificational one in (57b):

- (57) a. At every student's defense, her advisor cried
 b. $\text{every student's defense} \Rightarrow \lambda P \forall x [\text{student}(x) \rightarrow P(\text{1y defense}(y)(x))]$
 where $\text{defense}(y)(x) = x$ is y 's defense.

So, sentence (57a) gets the LF (58a) and interpretation (58b):

- (58) a. $\text{every student's defense}_2 [[\text{PP at } t_2]_1 \text{ TOP}^1 [\text{her}_1 \text{ advisor cried } t_1]]$
 b. $\lambda P \forall x [\text{student}(x) \rightarrow P(\text{1y defense}(y)(x))]$
 $(\lambda e_2 \exists e' [\text{at}(e_2)(e') \wedge \text{1}(e') \wedge \lambda X_1 \exists e [\text{AG}(e)(f_{\text{STUDENT}}(\omega_1)'s \text{ advisor}) \wedge \text{cry}(e) \wedge X_1(e)](\text{at}(e_2))) =$
 i. $\lambda P \forall x [\text{student}(x) \rightarrow P(\text{1y defense}(y)(x))]$
 $(\lambda e_2 \exists e' [\text{at}(e_2)(e') \wedge \text{1}(e') \wedge \exists e [\text{AG}(e)(f_{\text{STUDENT}}(\omega_1)'s \text{ advisor}) \wedge \text{cry}(e) \wedge \text{at}(e_2)(e)]) =$
 ii. $\forall x [\text{student}(x) \rightarrow \lambda e_2 \exists e' [\text{at}(e_2)(e') \wedge \text{1}(e') \wedge \exists e [\text{AG}(e)(f_{\text{STUDENT}}(\omega_1)'s \text{ advisor}) \wedge \text{cry}(e) \wedge$
 $\text{at}(e_2)(e)] (\text{1y defense}(y)(x))] =$
 iii. $\forall x [\text{student}(x) \rightarrow \exists e' [\text{at}(\text{1y defense}(y)(x))(e') \wedge \text{1}(e') \wedge \exists e [\text{AG}(e)(f_{\text{STUDENT}}(\omega_1)'s \text{ advisor}) \wedge$
 $\text{cry}(e) \wedge \text{at}(\text{1y defense}(y)(x))(e)] =$
 iv. $\forall x [\text{student}(x) \rightarrow \exists e' [\text{at}(\text{1y defense}(y)(x))(e') \wedge \text{1}(e') \wedge \exists e [\text{AG}(e)(f_{\text{STUDENT}}(e')'s \text{ advisor}) \wedge$
 $\text{cry}(e) \wedge \text{at}(\text{1y defense}(y)(x))(e)]$
 Here the pragmatics kicks in: e' is most naturally construed as the defense of each student; and $f_{\text{student}}(e')$ returns the student protagonist of the defense. Hence, as desired, (iv) boils down to:
 v. $\forall x [\text{student}(x) \rightarrow \exists e [\text{AG}(e)(x's \text{ advisor}) \wedge \text{cry}(e) \wedge \text{at}(\text{1y DF}(y)(x))(e)]$

The steps of the derivation are pedantically laid out and the final truth conditional outcome is the last line in (58b).

Summing up, the properties of English PP-fronting follow under the assumption that it is movement to a designated left peripheral position at which a DR is introduced: PP-movement must be driven by some topic-like head which attracts the bearer of a designated index. Such index is relevant for binding theoretic (or Minimality) purposes. The unwanted reading of (51a), repeated below as (59a) is ruled out as a case of Principle C/Minimality on the basis of the LF in (59b), the only one that could in principle derive the intended reading of (59a):

- (59) a. * At every student₁'s defense, he₁ burst into tears
 b. $[\text{DP every student's defense}]_2 [\text{PP at } t_2]_1 \text{ TOP}^1 [[\text{DP he}_{E-1} \text{ student}] \text{ burst into tears } t_1]$

4. Comparisons and perspectives.

The relation between overt dislocation and crossover displays a striking crosslinguistic stability: movement to topic positions alleviates WCO of the primary or secondary kind, but not strong

crossover.³¹ This can be made sense of, if we assume that such movement has the following simple ‘triangular’ shape:

(60) Topic movement

$XP_n \text{ Top}^n \dots ([\text{pro}_n]) \dots t_{XP_n} \dots$

Where t_{XP_n} is the trace/copy of XP_n and Top^n is a cover for any number of slightly different topic-oriented heads.

Top^n attracts XP_n and, by the Topic Condition, introduces a DR linked to n . If there is a pro_n between Top^n and the trace t_{XP_n} , Top^n provides an antecedent for it. But if the pro_n C-commands t_{XP_n} , a Principle C/Minimality violation ensues. This contrasts with what happens with wh-dependencies:

(61) WH-movement

$XP_n \text{ Q}^n \dots ([\text{pro}_n]) \dots t_{XP_n} \dots$

The syntax of (60) and (61) is basically the same. Q is an interrogative morpheme or feature that attracts wh-words to the comp region and drives question formation. The intervention of a pro_n on the C-command line between Q and the trace is equally disruptive for Topic- and for WH-movement. The difference is that the Q-morpheme *disallows* DR-introduction. Hence, the moved constituent will not be able to antecede a pro_n , even when the latter does not ‘intervene’. That is why WH-movement displays both SCO and WCO effects. A’-dependencies, as we’ve known for a long time, fall into two categories, which we propose to conceptualize as DR-introducing vs. non DR-introducing.

In the rest of the present section, we will first carry on some more comparisons with other approaches to crossover, which will allow us to formulate a hypothesis on why two types of A’ - movement (DR-introducing vs. non DR-introducing) exist. Then we will look at further consequences of the present proposal for the binding theory and conclude.

4.1. Where do two types of A’- dependencies come from?

Discussions on the role of ‘leftness’ in pronoun interpretation have a long history of which, e.g., Postal (1968), Chomsky (1976) and Jacobson (1977) are early important voices; they carry right into modern approaches which are directly inspired by leftness, like Shan and Barker (2006) or Bumford and Charlow (2023). The basic idea behind leftness based approaches is that crossover stems ultimately from a left-to-right asymmetry. Essentially, a quantificational element cannot bind something to the left of its base position, perhaps a formal reflex of a processing constraint. This is at the basis of the deviance of sentences like (62a,b), in contrast with (62c). The base position coincides with the surface (i.e., spellout) position in (62a), but not in (62b).

- (62) a. * His₁ advisor failed to help every student₁
 b. * Who₁ did his₁ advisor fail to help t₁?
 c. If his₁ advisor helps him₁ in a timely fashion, a student₁ rarely fails

³¹ In the final revision stages for this paper, I am found out from Wu (2024) that the voice system of Seediq (Austronesian, Northern Taiwan) displays a form of ‘pivot’ selection with rightward dislocation that is subject to the same basic constraint studied in the present paper: it repairs weak crossover violations, but not strong crossover ones.

This is a simple idea, and one can find it quite appealing. Yet, the data studied in this paper strike me as a definitive refutation of the viability of leftness based approaches to crossover (and I am happy, of course, to be proven wrong). Consider any instance of the dislocation paradigm we have investigated here, like:

- (63) a. * [At every student_i's defense], he_i cried t_i
 b. [At every student_i's defense] his_i advisor cried t_i

First, it seems outright impossible to force a movement analysis for one of the two sentences in (63) and block it for the other: the type of constituents involved, type of locality, and every other imaginable condition are the same across the two sentences. Second, if we go for a base generation analysis and assume that both PPs in (63) are base generated to the left, what could possibly block (63a)? We know that possessor binding out of a DP is possible. If, instead, a movement analysis is adopted for both (63a-b), leftness would have to apply with respect to the position of the trace in (63a) (just like WH-movement in English according to Shan and Barker); but it would have to apply with respect to the surface position in (63b) (just like QR according to Shan and Barker). What on earth could justify this? The crosslinguistic stability of this pattern through differences in morphosyntax, locality conditions, etc. makes it impossible to dismiss the reality of the generalization under discussion as a quirk.

In so far as I can see, bimodal theories of WCO stand a better chance at making sense of what is going on. In the remainder of this subsection, I want to draw attention to certain internal differences between Büring's and Chierchia's systems, in the hope of gaining insight on why there are two types of A'-dependencies. The chief differences between Büring's and Chierchia's approaches are three. The first is that on Chierchia's approach the two operators differ semantically (i.e. model theoretically), while on Büring's they differ syntactically (i.e. in the syntax of LF). The second difference has to do with general life-style choices between the dynamic vs. the situation based approaches to non C-command anaphora. The third, a corollary of the second, concerns binding from objects into adjuncts, already discussed briefly in Section 2.2. but worthy of a further glance. Let me briefly review these differences, in turn.

Büring's μ and β operators are quite similar semantically to each other: μ is abstraction restricted to expressions of type t , β abstraction restricted to constituents of type $\langle a, t \rangle$ (for any type a). This is a difference in (logical) syntax. And one that will have to be abandoned for topic movement. Topics of the kind considered here are created via movement out of full blown clauses, i.e. things of type t . Hence, if we are to extend use of β to them, we will have to let β apply to things of type t as well. This makes μ and β all the more alike. The situation is very different on Chierchia's dynamic take. On the one hand, for non DR introducing movement, we have good old λ -abstraction that targets static variables. For DR introducing movement, we have the λ^{D} -operator with a scope potential which spans unboundedly through conjunctively interpreted constituents. So on Büring's approach the difference between the operators appears somewhat arbitrary,³² while on Chierchia's it is grounded in a scope behavior that arguably has a semantically natural characterization. This leads to the second difference. As is well known, situation based theories vs dynamic theories diverge empirically on several aspects of anaphora. We cannot review these divergences here in detail, but we can report a few classic examples.

³² As Büring (2004, p. 27, fn.3) himself admits.

- (64) a. A glass broke. John was very upset. **It** was a family heirloom.
 b. In fact two glasses broke, But John didn't care about one of the two.
 c. If a glass broke and it was a family heirloom, John will be very upset.

How does the pronoun in the third sentence in (64a), marked in boldface, get its antecedent? On the dynamic approach, it is a case of *direct* binding (in spite of the lack of C-command): A DR is introduced in the first sentence, and discourse sequencing is just conjunction (i.e. composition of sentence values) which passes onwards active DRs. There is no semantic difference between (64) and sentence internal, C-command based anaphora like (65):

- (65) a. John introduced Mary to his father
 b. $\exists e [AG(e)^1(john) \wedge [TH(e)]^2(mary) \wedge [GO(e)]^5(\omega_3's \text{ father}) \wedge introduce(e)]^{33}$
 ω_3 picks out the individual stored at the coordinate 3, i.e. Mary.

On an event based analysis, clauses are interpreted as a structurally determined sequences of conjuncts, as illustrated in (65b). And thus C-command anaphora is subsumed under the same mechanism as discourse sequencing. In contrast with this, on any standard situation theoretic approach, quantifiers are limited in their scope to their C-command domain. So no quantifier from the first sentence in (64a) can reach into the second sentence, not even the quantifier over events or situation. This entails that the pronoun *it* in (64a) must be an E-type pronoun, i.e. a description and descriptions require some kind of uniqueness/maximality. E.g., the pronoun *it* in (64a) could be understood as something like ‘the glass that broke’. But the uniqueness presupposition that this resolution of the E-type pronoun requires is at best an implicature that can be readily cancelled (as in 64b) or disappears *sua sponte* in e.g. the antecedents of conditionals (as in 64c). So what kind of description might do the trick and provide a good candidate for the interpretation of the pronoun *it* in (64)? I don't know of a satisfactory answer to this question. For my money, this is a case where the dynamic approach gets things right and situation-based approaches get it wrong. But people disagree and the matter certainly can't be considered closed.

This second background difference between Buring's and Chierchia's theories brings about the third, crossover specific difference, having to do with binding into adjuncts. The latter is a long standing issue, which we reviewed in (31) above, repeated here:

- (66) a. We'll sell no wine before its time
 b. no wine₃ [we'll sell TH³ t₃ before its₃ time] LF
 c. $\neg \exists u \exists e [AG(e)^1(we) \wedge TH(e)^3(u) \wedge sell(e) \wedge BEFORE(e)(\omega_3's \text{ time})]$ Interpretation

There is just no A-position where we can attach Buring's β -operator that would yield the desired results: if we attach it to the object, β could not bind into the adjunct, which is higher up. If we attach it to the subject, the anaphoric link will be to the subject, which is not what we want. And that's it: there are no other A-positions to play with. The dynamic approach, on the other hand, predicts that this binding should be just like C-command binding, because temporal modifiers are linked to the clause they modify conjunctively as in (66c), and DRs are passed on across conjuncts, regardless of C-command.

Having clarified the trade-offs between the two bimodal ‘two-track’ approaches under discussion, we can speculate as to what may drive DR-introduction vs. non DR-introduction at sites derived via

³³ The superscripts mark active DRs; see (67) below and (2c) in the Appendix for a formal explanation of this notation.

movement. The speculations I am about to present flows very naturally, I think, only from the dynamic take. In the tradition of dynamic semantics (and also of the related Discourse Representation Theory), it is DPs (i.e. arguments) that are taken to introduce DRs. Chierchia (2020) makes an alternative proposal, which he dubs the Dynamic Predication Principle that turns this tradition around:

(67) Dynamic Predication Principle (DPP):

Only Predicates can introduce DRs. DPs as such never do.

i.e. the λ -operator must associate with predicates and cannot associate with DPs.

Example:

a. John snored

b. John [λv snored]

c. $\exists e$ [[AG(e)] λ (john) $\wedge$ snore(e)]

where: [AG(e)] λ = λu [λ (u) $\wedge$ AG(e)(u)] and [λv snored] $\Rightarrow \lambda u$ [[AG(e)] λ (u) $\wedge$ snore(e)]

DPP makes thematic heads the privileged locus for DR introduction, as in the example in (67a-c). In (67a) it is the agent introducing head (say, little *v*) that brings about a DR, linked to the subject, as indicated in the LF (67b), and its interpretation (67c). Now, when movement comes in, it creates *derived* predicates by ‘punching holes’ into the clause. Can derived predicates introduce DRs? The DPP offers a partial answer to this question. Predicates created by a type mismatch, for the sole purpose of discharging scope do NOT introduce DR. For otherwise, the DPP would be vacuous. All quantificational DPs *must* be assigned scope, and referential DPs *may* also be scope taking in some circumstances (see e.g. Fox 2000). If DR introduction was legit and automatic at scope sites, it would entail that all DPs create DRs, after all. And the λ -operator would end up associating with DPs, contra the DPP. It follows that the main scope assigning device, namely QR, and *wh*-movement (at least on Karttunen’s 1977 classic analysis) *cannot* be DR-introducing types of movement. They thus will give rise to WCO effects: A QRed DP will not be able to antecede a pronoun in ‘no man’s land,’ i.e. between its base position and its landing site.

In conclusion, the DPP offers a partial, but principled answer to *why* certain types of movement form a natural class in disallowing DR-introduction, thereby yielding full-ledged crossover effects. For focus or topic driven movement, we’ll have to figure things out on a case-by-case basis, using WCO obviation as a criterion. That topics (and, e.g., the EPP subject position, which is topic-like – see Rizzi 2005) should allow DR activation is a fairly natural hypothesis, spelled out here as the Topic Condition.

4.2. Further binding theoretic consequences.

In dealing with SCO, we have tried to keep our options open between a traditional Principle C based account and one in terms of Minimality. But all the phenomenology discussed in the present paper is amenable to a Minimality account, as we saw throughout. This raises a daunting question which can no longer be put off: is Principle C necessary? Do we still need it?

Consider some of the central cases typically attributed to Principle C, starting with when a lower quantifier takes scope over an upper one.

- (68) a. i. A doctor visited every patient
 ii. Every patient₁ [a doctor inspected t₁]
 b. i. * he₁ visited every patient₁
 ii. * every patient₁ [he₁ visited t₁]

While (68a.i) is perfectly natural on the reading where the object has scope over the subject, reading that can readily be obtained via QR as in (68a.ii), binding from a lower quantifier into a subject is totally out. This is usually attributed to Principle C. The same holds for cases where QR is slightly less local as with ECM, Raising, and Restructuring constructions. The following paradigm is from Cecchetto (2004):

- (69) $\forall$ -over- $\exists$ readings possible
- | | |
|--|---------------|
| a. Someone expected every republican to win | ECM |
| b. Someone seemed to attend every class | Raising |
| c. Someone wants to dance with every woman | Restructuring |
| Binding impossible | |
| d. * He_1 expected every republican to win. | |
| e. * He_1 seemed to attend to every patient _i | |
| f. * He_1 wants to dance with every man _i | |

As it turns out, these classic facts follow from the present account without any need for Principle C. They follow because (i) Quantifiers cannot ever directly bind a pronoun and (ii) QR does not introduce DRs. These two points are the essence of the bimodal approach to crossover. In dynamic parlance, traces are Tarskian variables and pronouns are dynamic variables. Even if the numeric exponent is the same, they are of different semantic types and never directly ‘talk’ to each other. One needs an explicit occurrence of a DR-introducing operator $\underline{u}(u_n)$ to link the value of the static variable u_n to the dynamic address $\underline{n}$.

- (70) a. i. every man_i $he_{\underline{i}}$ wants to dance with t_i
 ii. every(man) ($\lambda x_1. \omega_1$ want to dance with x_1)
 b. * every man_i Top_i he_i wants to dance with t_i
 every (man)($\lambda x_1. \underline{1}(x_1) \wedge \omega_1$ wants to dance with t_i)

In (70a) the lambda operator cannot directly bind the dynamic variable t_i . If we had movement to a DR-introducing site, as we do have with Hindi scrambling repeated here in (70b), DR-introduction would introduce an antecedent for the pronoun. But then minimality would kick in and the Top-head would be unable to target directly the object, by-passing the subject. These are necessary consequences of our take. Sentences like those in (68b) or (69d-f) all fail to have sentence internal antecedents, because QR is NOT movement to a position where DRs are introduced. We do not need to say more. We do not need to ban, e.g., (70a) from being well formed. This configuration fails to achieve binding of the pronoun in subject position by the object. And that suffices.

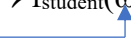
These basic cases extend to more complex ones like:

- (71) a. * who_i does he_i think Mary will marry t_i
 b. who_i Qⁱ does $he_{\underline{i}}$ think Mary will marry t_i

The deviance of a sentence like (71a), on the interpretation associated with the choice of indices indicated there, is generally imputed to Principle C: in (68a) the trace t_i is A-bound by the pronoun he_i . We could extend the minimality based logic used in connection with Topicalization to the

cases in (71). But there is a more direct way of ruling (71a) out. On the present approach, pronouns can only get their antecedents from duly activated DRs. But the subject pronoun *he_I* has no DR accessible to it and wh-movement, unlike topicalization, cannot create one. Hence, *he_I* lacks a sentence internal potential binder.

Let us turn, finally, to indirect binding/SSCO.

- (72) a. *He_I* is grateful to every student₁'s advisor
 b. every student's advisor $\Rightarrow \lambda P \forall x [\text{student}(x) \rightarrow P(\iota y. \text{advisor-of}(x)(y))]$
 c. [every student's advisor]₁ [*he_I* is grateful to t₁]
 d. i. [every student₁'s advisor]₂ [*he_{E-2}* ~~student~~ is grateful to t₂]
 ii. $\lambda P \forall x [\text{student}(x) \rightarrow P(\iota y. \text{advisor-of}(x)(y))]$ ($\lambda x_2 . f(\omega_2)$ is grateful to x_2) =
 iii. $\forall x [\text{student}(x) \rightarrow f_{\text{student}}(\omega_2)$ is grateful to $\iota y. \text{advisor-of}(x)(y))$
 unbound 

There is no indexing that can get the pronoun *he_I* to co-vary with the student or with the advisor, regardless of whether *he_I* is interpreted as E-type pronoun or as a standard one. Recall that the object DP is interpreted as in (72b). We already know from (71) that *he* cannot be directly bound by applying QR to (72a) as in (72c). This configuration fails to yield binding of the pronoun in subject position for the same reasons why (71a) does. If we try indirect binding with *every student*, we run into a similar problem. The E-type pronoun is a function that contains a standard pronoun that must find an active DR as its anaphoric antecedent. But QR does not introduce DRs. So the E-type pronoun is doomed not to have a suitable antecedent.

One must be cautious with something that has so many ramifications like Principle C. But I am tentatively led to conclude that, in so far as I can see, the effects of Principle C are independently derivable. If correct, this is exciting. It reduces the complexity of the Binding Theory in favor of a highly natural locality condition on movement that is independently necessary. And this reduction eliminates the most ‘cumbersome’ of the binding theoretic principles, the one whose functional grounding is hardest to see.

The present approach, however, says nothing about coreference.

- (73) a. *He_I* likes John₁
 b. *He_I* likes John₁'s father
 c. *He_I* thinks that John₁'s father is boring.
 d. John₁ [*he_I* likes t₁]

As on the classic Binding Theory, the present approach has an account for why *John* in (73a-c) cannot bind the pronoun in subject position. But to the extent that pronouns can be referential, it cannot prevent them from co-referring with referential DPs, like proper names. To the best of my understanding, the same holds for Principle C. The latter can block, e.g., (73d), but it cannot block co-referential interpretations for (73a-c). Regulating (non)coreference falls under the purview or Rule I and its derivatives.

Interestingly, even if we drop Principle C, the overall architecture of the Binding Theory doesn't change by much: Principle A/B and Minimality rule out certain syntactic configurations and in so doing they prevent certain interpretations from coming about. (Non) Coreference is

handled by Rule I/Rule H. Of course, the role of Rule I/H would come out considerably enhanced, as all of the phenomena imputed to Principle C should now fall under Rule I's umbrella.

Be that as it may, let me spend a few final words to show the underlying semantics of binding makes a difference also to the characterization of Non Coreference. On the classical approach, the only form of direct binding is C-command binding. Everything else has to go via indirect binding and E-type pronouns. On a dynamic approach, this is not so. Direct binding takes place from active DRs across anything interpreted conjunctively. C-command binding gets subsumed under this view because the clause is made up of conjunctions, with each theta-role adding one, in a structurally determined order. So, differences in the underlying logic have consequences also for Rule I- type phenomena. Consider the triplet in (74a-c):

- (74) a. * He_n likes John_n's father.
 b. * He_n is home and John_n watches TV.
 c. If he_n is home, John_n watches TV.
 d. ($\phi \Rightarrow_{\text{ADV}_n} \psi$)

Sentences (74a-b) are deviant on the coreferential interpretation. The classical binding theory account for the deviance of (74a) in terms of Rule I runs roughly as follows: replacing *John* in (74a) with something bindable would yield a structure where C-command binding by the subject is possible. Hence coreference here is typically not intended/ruled out. However, this account does not extend to (74b), where C-command binding between the first and the second clause is out of the question, even though (74a) and (74b) feel very similar in their deviance. On the dynamic approach, the links between (74a) and (74b) are, in fact, identical and if the first counts as direct binding, then so does the second. Hence Rule I would have to apply to both (74a) and (74b) in the same way; but not so in (74c). In dynamic semantics, the binding relationship between the antecedent and the consequent in conditionals and related structures is mediated by a (possibly null) adverb of quantification, as per the schematic logical form in (74d). Hence, nothing in the antecedent of a conditional ever *directly* binds anything in the consequent. Consequently, Rule I is irrelevant to (74c). This illustrates further the potential of dynamic semantics for the treatment of recalcitrant aspects of anaphora.³⁴

Be that as it may, the elimination of Principle C in favor of an expanded role for Minimality seems feasible and might well be a worthy pursuit.

³⁴ An anonymous referee points out the following sentence, as a possible counterexample to the claim that cataphora across conjunction (74b) should have the feel of a violation of non co-reference as severe that those that arise in Principle C configurations like (74a):

(a) He₁ got to be home on a week end after a very long time and so now finally John₁ can watch the football game live.

Two quick points in response to this. First, the discourse structure in (a) improves also 'Principle C configurations' like:

(b) He₁ was happy to be home on a week end after a very long time, because so John₁ could now finally watch the football game live.

vs.

(c) * He₁ was happy because John₁ was home

Second, Rule I is not a well-formedness constraint like Principle C. It's an interpretive procedure meant to accommodate the many different cases that a mechanism governing non coreference need to take care of (like, e.g., *He put on John's coat: He must be John*, etc.). How this will eventually happen is still very much up for grabs.

4.3. Concluding remarks.

Topic oriented movements ameliorate WCO, but not SCO, systematically across languages. Our proposal on this is perhaps too simple to be true: Topic heads introduce DRs linked to the moved constituents. Such DRs can act as antecedents for pronouns occurring in no one's land between the trace and the landing site. If Principle C gets subsumed under Minimality, a general (if parametrized) constraint on movement, then that is the end of the story. If not, then we may simply observe that Topic movement still constitutes, in traditional terms, a kind of A'-movement, and hence is expected to be Principle C compliant. Topic movement contrasts with movement created to discharge or mark scope, which hasn't the power to introduce DRs, and hence displays WCO effects, along with SCO ones. If one buys into the Dynamic Predication Principle ('DPs never introduce DRs, only predicates do') the latter property of scope discharging is to be expected. Embedding this line of analysis within a dynamic set up gives it extra bite, because dynamic semantics IS the logic of DR introduction (besides solving empirical issues like e.g. how pronoun binding from the object position into high adjuncts is possible in spite of the lack of C-command).

We have adopted, in line with Büring and many others, an indirect approach to possessor binding and related phenomena. In, e.g., *every Englishman's mother loves him*, the pronoun *him* is an E-type pronoun formally coindexed with the subject DP, which acts as its direct binder; co-variation with the possessor comes about thanks to the functional interpretation of pronouns. We have offered a rather simple, fully explicit model for this kind of link. Under this view, there is nothing to say about SWCO vs. SSCO. They are simply reduced to the more basic crossover patterns (without having to keep track of complex traces).

These results constitute evidence that more and more aspects of the interaction of anaphora and movement are yielding and that the dynamic perspective keeps being a fruitful tool for the investigation of anaphoric dependencies across languages.

Appendix: A formal existence proof.

In this Appendix we sketch a formal system that sustains a bimodal approach to crossover and dislocation, in the sense that provides a logic into which Logical Forms of the kind discussed in this paper can be mapped by simple, compositional, general procedures. For the reasons laid out in Section 4, we choose a dynamic system. This does not mean that, modulo the considerations in Section 4, a Büring style situation semantics could not also do the job. The formal system sketched here is a minor variant of the one in Chierchia (2020). We only outline the extensional part of the system.

(1) Types.

a. Basic Types: e , t , n ; where:

$D_e = U$ (domain of individuals, *including events*)

$D_t = \{0, 1\}$,

$D_n = N$ (the set of positive integers).³⁵

We thus have in the representation language constants of the form $\underline{1}$, $\underline{2}$, ..., $\underline{n}$,... ranging over numbers. They will be used to model discourse referents.

³⁵ One referee makes the following (cute/friendly) "nitpicky point": The number 1 belongs to both D_t and D_n . Indeed it does. But this has no dire consequences.

b. Derived types:

$D_{\langle a, b \rangle} = [D_a \Rightarrow D_b]$ (the set of all total or partial functions from D_a into D_b)

The set of meaningful expressions, models, etc. for a formal language with these types is defined as for any multisorted type theory like Gallin's (1975) TY2.

c. Some key derived types:

- i. $\omega = \langle n, e \rangle$: assignments in the representation language
- ii. $T = \langle \omega, \langle \omega, t \rangle \rangle$: Context Change Potentials (= relations between assignments)
T is the type for the value of clauses; it plays the role t does in a static set up.

d. Dynamic Predicate Lift:

- i. P is of type $\langle e, t \rangle$, $[P] = \lambda u \lambda \omega \lambda \omega'. \omega = \omega' \wedge P(u)$
[] maps static predicate values (of type $\langle e, t \rangle$) into dynamic ones (of type $\langle e, T \rangle$), which are functions from individuals into CCPs. In the examples that follows, I use e, e', e_n, ... as variables ranging over events (which are a sort of individuals, included in D_e). E.g.:
- ii. $\text{run}_{\langle e, T \rangle} = [\lambda e. \text{run}_{\langle e, t \rangle}(e)] = \lambda e \lambda \omega \lambda \omega'. \omega = \omega' \wedge \text{run}(e)$
 $\text{run}_{\langle e, T \rangle}$ is a function that maps any event e into the CCP associated with 'e is a running event'.

(2) DR-introduction.

The DR-introducing operator is of type $\langle e, T \rangle$. It takes an individual and places it at a designated address n.

$$a. \text{ } \overset{n}{\lambda} () = \lambda u \overset{n}{\lambda} (u) = \lambda u [\lambda \omega \lambda \omega'. \omega =_{n/u} \omega']$$

where $\omega =_{n/u} \omega'$ is an abbreviation of $\omega' = \omega \cup \langle n, u \rangle$, defined only if the input assignment is *undefined* for the nth-coordinate.

The 'undefinedness' condition captures Heim's 'novelty condition' on DR introduction.

When a DR is introduced, a fresh address must be used.

It is also useful to have a predicate level version of DR introduction:

$$b. \text{ If P is of type } \langle e, T \rangle, [P]^{\overset{n}{\lambda}} = \lambda u. \overset{n}{\lambda} (u) \wedge P(u)$$

(3) Pronouns/Discourse Referents.

Pronouns are of type $\langle \omega, e \rangle$. E.g.:

$$a. \text{he}_n = \lambda \omega. \omega_{\overset{n}{\lambda}}, \text{ where } \omega_{\overset{n}{\lambda}} \text{ is an abbreviation for } \omega(\overset{n}{\lambda})$$

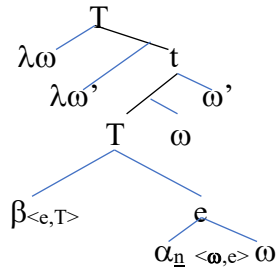
Pronouns combine with dynamic predicates (of type $\langle e, T \rangle$) via function composition:

$$b. \text{ If } \beta \text{ is of type } \langle e, T \rangle, \alpha_{\overset{n}{\lambda}} \text{ is of type } \langle \omega, e \rangle, \text{ then}$$

$$\beta(\alpha_{\overset{n}{\lambda}}) = \lambda \omega \lambda \omega'. \beta(\alpha_{\overset{n}{\lambda}}(\omega))(\omega)(\omega').$$

Since it is difficult to process, I provide a type-theoretic tree for this definition:

c.



d.Example: $[AG(e_3)](\lambda\omega. \omega_1) = \lambda\omega\lambda\omega'. \omega = \omega \wedge AG(e_3)(\omega_1)$.

‘He₁ is the agent of e₃’

We will write $[AG(e_3)](\omega_1)$ for $[AG(e_3)](\lambda\omega. \omega_1)$.

(4) The logic.

For any ψ, ϕ of type T, any variable ω, ω' of type ω and any variable α_a of type a:

a. $\downarrow_{\omega} \phi = \exists\omega' [\phi(\omega)(\omega')]$ ‘ ϕ is true relative to ω' ’; $\downarrow$ is of type $\langle\omega, \langle T, t \rangle\rangle$

b. $\neg \phi = \lambda\omega\lambda\omega'. \omega = \omega' \wedge \neg\downarrow_{\omega} \phi$

c. $(\phi \wedge \psi) = \lambda\omega\lambda\omega'. \exists\omega'' [\phi(\omega)(\omega'') \wedge \psi(\omega'')(\omega')]$ Composition of relations.

d. $\phi \rightarrow \forall \psi = \neg(\phi \wedge \neg \psi)$

e. $\phi \vee \psi = \neg(\neg \phi \wedge \neg \psi)$

f. $\exists\alpha_a \phi = \lambda\omega\lambda\omega'. \exists\alpha_a [\phi(\omega)(\omega')]$

g. $\forall\alpha_a \phi = \neg\exists\alpha_a \neg \phi$

This logic encodes a standard characterization of accessibility.

(Cf., e.g., Groenendijk and Stokhof 1991).

(5) Example.

We analyze:

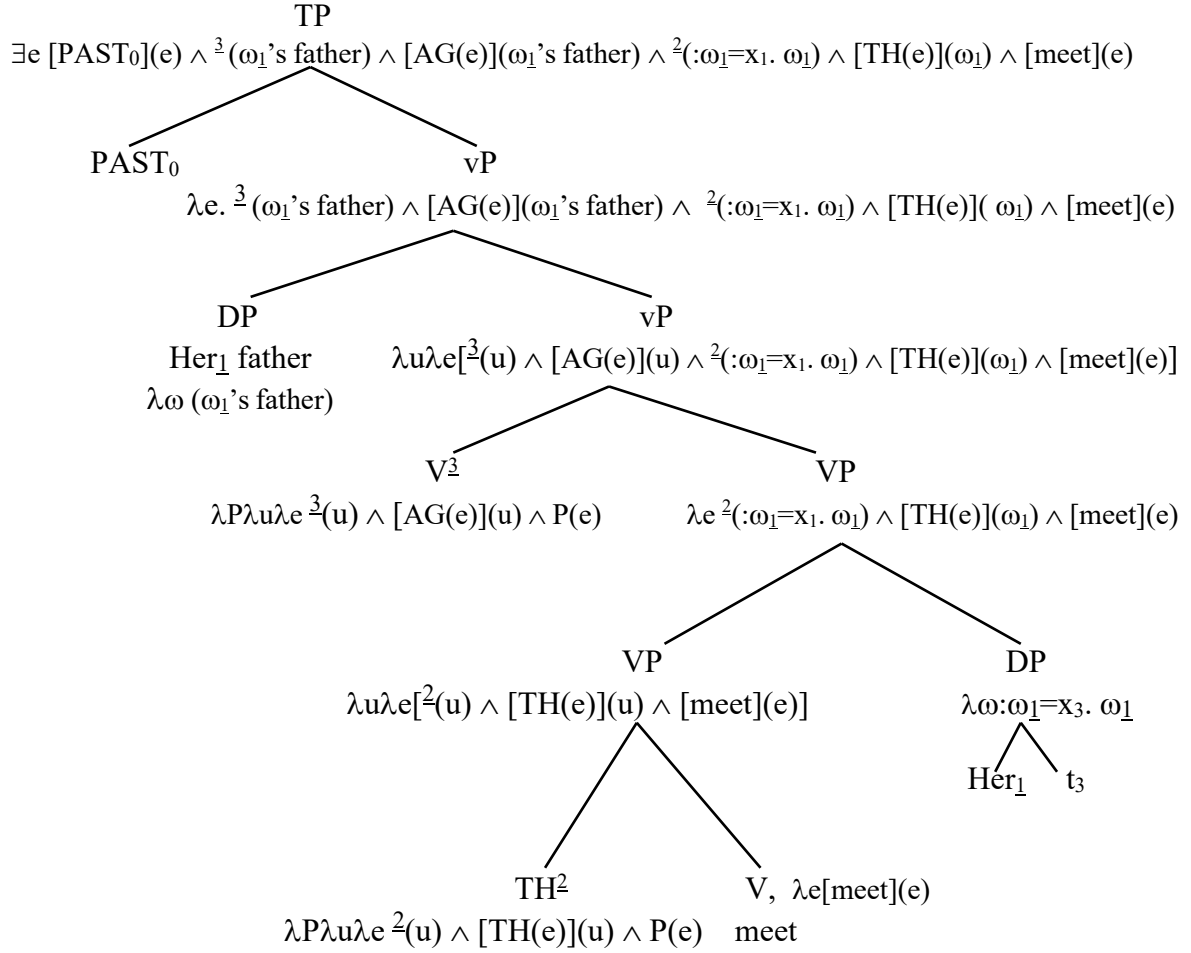
- a. Maria₁ Top¹ [_{TP} suo₁ padre [_{la}₁ t₁] ha incontrata t₁ ieri]
 Maria, her father her has met yesterday
 ‘Maria’s father met her yesterday’

As usual, dynamic variables are underlined. We will go through the TP first. To simplify things, we will analyze the clitic as if it was in its base position. I.e. we will work with the following structure:

- b. suo₁ padre ha incontrata [_{la}₃ t₃] ieri
 her father has met her yesterday

We skip the treatment of the temporal adverb *yesterday*. In the LF below, DR-introducing operators are marked as superscripts on DR-introducing heads. Each node has its logical translation below or next to itself. Take a first global look at the whole tree, then work your way up from the leaves to the root node. In this example, two DRs are introduced, one for the theme, linked to the trace + object pronoun complex; the other for the agent, linked to the *her₂ father* constituent. The topic will eventually provide an extra DR and an antecedent for both the object pronoun and for the one in subject position (as they are co-indexed).

c.



θ -introducing heads (heads that introduce thematic roles) are: v, TH; v introduces the external theta role (Agent), TH the internal one (Theme). The semantics of any θ -introducing head is as in standard event semantics, with the addition of DR-introduction:

d. $\theta^a \Rightarrow \lambda P \lambda u \lambda e. {}^a(u) \wedge [\theta(e)](u) \wedge P_w(e)$

PAST in (c) combines with a property of events, places the event's running time relative to the reference time and closes existentially the event argument:

e. $PAST_0 = \lambda P \exists e [PAST_0](e) \wedge P(e)$, where $[PAST_0] = [\lambda e. \tau(e) < t_0]$ and τ is the temporal trace function that maps events into their running time.

We now consider the whole of the topic phrase.

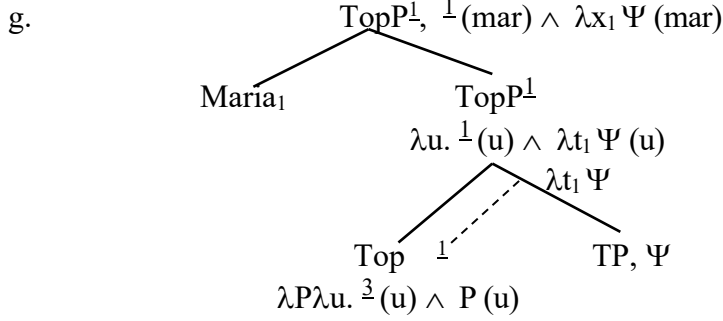
e. Lexical entries:

i. $Top^a \Rightarrow \lambda P \lambda u. {}^a(u) \wedge P(u)$

ii. $Maria_n \Rightarrow mar$

f. Top-schema: $\|Top^a \Psi\|^g = Top^a(\lambda t_n. \|\Psi\|^g)$

The superscript on *Top* is interpreted as an abstractor over *Top*'s sister, by analogy with the interpretation of QR in Heim and Kratzer (1998).



Where $\Psi = \exists e [\text{PAST}_0](e) \wedge \text{}^3(\omega_1 \text{'s father}) \wedge [\text{AG}(e)](\omega_1 \text{'s father}) \wedge \text{}^2(\omega_1 = x_1. \omega_1) \wedge [\text{TH}(e)](\omega_1) \wedge [\text{meet}](e)$

g. Reductions:

i. $\text{}^1(\text{mar}) \wedge \lambda x_1 \exists e [\text{PAST}](e) \wedge \text{}^3(\omega_1 \text{'s father}) \wedge [\text{AG}(e)](\omega_1 \text{'s father}) \wedge \text{}^2(\omega_1 = x_1. \omega_1) \wedge [\text{TH}(e)](\omega_1) \wedge [\text{meet}](e)](\text{mar})$

ii. $\text{}^1(\text{mar}) \wedge \exists e [\text{PAST}](e) \wedge \text{}^3(\omega_1 \text{'s father}) \wedge [\text{AG}(e)](\omega_1 \text{'s father}) \wedge \text{}^2(\text{mar}) \wedge [\text{TH}(e)](\omega_1) \wedge [\text{meet}](e)]$

In going from (ii) to (iii), I eliminate for ease of readability one of the two active DRs to which *Maria* winds up linked (where the outermost 1 comes from Top, the other 2 from TH).

iii. $\text{}^1(\text{mar}) \wedge \exists e [\text{PAST}](e) \wedge \text{}^2(\omega_1 \text{'s father}) \wedge [\text{AG}(e)](\omega_1 \text{'s father}) \wedge [\text{TH}(e)](\omega_1) \wedge [\text{meet}](e)]$

In primitive notation, i.e. applying the definitions in (1)-(4) and performing tautological transformations, the CCP associated with (iii) comes to:

iv. $\lambda \omega \lambda \omega'. \lambda \omega = \langle 1/\text{mar}, 2/\text{mar's father} \rangle \omega' \wedge \exists e [\text{PAST}(e) \wedge \text{AG}(e)(\text{mar's father}) \wedge \text{TH}(e)(\text{mar}) \wedge \text{meet}(e)]$

The first conjunct in (iv) codes the active DRs; the second conjunct transparently gives the truth conditions, viz:

v. $\exists e [\text{PAST}(e) \wedge \text{AG}(e)(\text{mar's father}) \wedge \text{TH}(e)(\text{mar}) \wedge \text{meet}(e)]$

Acknowledgements. I am grateful to three anonymous referees for thoughtful and useful critical advice. I am furthermore grateful to Rajesh Bhatt, Ivano Caponigro, Carlo Cecchetto, Veneeta Dayal, Kate Davidson, Danny Fox, Stephan Keine, the audience of my 2023 LSA Institute class, and the audience of my Spring 2024 Harvard Seminar (Co-taught with Kate Davidson). My debt to D. Buring's work is evident throughout.

References.

- Alexiadou, A. (2006) "Left dislocation (including CLLD)." In M. Everaert, H. van Riemsdijk, R. Goedemans and B. Hollebrandse (eds) *The Blackwell companion to syntax*, Blackwell, Oxford, pp. 668-699.
- Angelopoulos, N and D. Sportiche (2017) "Scrambling(s) + Clitics = CLLD," paper presented at Going Romance, 2016, Goethe University, Frankfurt.
- Anvari, A. (2024) "Revisiting the E-type theory" talk presentend at LingLunch, MIT.
- Bhatt, R. and S. Keine (2019) 'Secondary strong crossover in Hindi and the typology of movement' in M. Baird and J. Pesetsky (eds) *Proceedings of NELS 49*, Amherst: GLSA, p. 125–134.
- Bhatt, R. and S. Keine (2023) "Crossover asymmetries," Manuscript, University of Massachusetts at Amherst and UCLA.

- Bumford, D. and S. Charlow (2023) ‘Dynamic semantics with static types,’ Manuscript, Rutgers University and UCLA.
- Büring, D. (2003) “On D-trees, Beans, and B-accents,” *Linguistics and Philosophy*, 26.):511–545.
- Büring, D. (2004) ‘Crossover situations’, *Natural Language Semantics*, 12.1: 23-62.
- Büring, D. (2005) ‘Bound to bind’, *Linguistic Inquiry*, 36 .2: 259–274.
- Cecchetto, C. (1994) “A semantic trigger for scrambling.” *Rivista di Grammatica Generativa*, 19: 33-69.
- Cecchetto, C. (2000) ‘Doubling Structures and Reconstruction,’ *Probus*, 12:1: 1-34.
- Cecchetto, C. (2004) ‘Explaining the locality conditions of QR. Consequences for the Theory of Phases,’ *Natural Language Semantics* 12.4: 345-397.
- Chierchia, G. (2020) ‘Origins of weak crossover: when dynamic semantics meets event semantics,’ *Natural Language Semantics*, 28.1: 23–76.
- Chierchia, G. (2022) ‘People are fed up; don’t mess with them: Non quantificational arguments and polarity reversals,’ *Journal of Semantics*, 39.3: 475-521.
- Chomsky, N. (1976) “Conditions on rules of grammar”, *Linguistic Analysis*, 2.4: 303-351.
- Chomsky, N. (1981) *Lectures on Government and Binding*, Foris, Dordrecht.
- Cinque, G. (1977) ‘The movement nature of left dislocation,’ in *Linguistic Inquiry*, 8.2: 397-412.
- Cinque, G. (1990) *Types of A’-dependencies*, MIT Press, Cambridge, Mass.
- Cooper, R. (1979) “The Interpretation of Pronouns,” In F. Heny (ed) *Syntax and Semantics*. Vol 10, Brill, Leiden.
- Dayal, V. (1996) *Locality in WH-quantification*, Kluwer, Dordrecht.
- Dayal, V. (2016) *Questions*, Oxford University Press, Oxford, U.K.
- Dayal Srivastav, V. (1994) ‘Binding facts in Hindi and the scrambling phenomenon,’ In in M. Butt, T. Holloway and G. Ramchand (eds) *Theoretical Perspectives on Word Order in South Asian Languages*, CSLI, Stanford, pp. 237-262.
- De Cat, C. (2007) “French dislocation without movement,” *Natural Language & Linguistic Theory*, 25.3: 485-534.
- Elbourne, P. (2005) *Situations and Individuals*, MIT Press, Cambridge, Ma.
- Fox, D. (2000) *Economy and semantic interpretation*, MIT Press, Cambridge Ma.
- Fox, D. (2002) “Antecedent contained deletion and the copy-theory of traces”, *Linguistic Inquiry*, 33.1: 63-96.
- Fox, D. (2003) “On Logical Form”, in R. Hendricks (ed) *Minimalist Syntax*, Wiley On Line Library, Blackwell, Oxford, U.K.
- Gallin, D. (1975) *Intensional and higher order modal logic*, Elsevier, New York
- Grodzinsky, Y. and T. Reinhart (1993) ‘The innateness of binding and coreference.’ *Linguistic Inquiry* 24.1: 69–101.
- Groenendijk, J. and M. Stokhof (1991) ‘Dynamic Predicate Logic,’ *Linguistics and Philosophy*, 14.1: 39-100.
- Gurtu, M. (1992). *Anaphoric Relations in Hindi and English*. Munshiram Manoharlal, New Delhi.
- Heim, I. (1990) “E-Type Pronouns and Donkey Anaphora,” *Linguistics and Philosophy* 13.2: 137-177.
- Heim, I. (2007) ‘Anaphora and semantic interpretation: A reinterpretation of Reinhart’s approach,’ in U. Sauerland and O. Percus (eds) *The interpretive tract*, MIT Working Papers in Linguistics, Cambridge, Ma.

- Heim, I. and A. Kratzer (1998) *Semantics in generative grammar*, Blackwell, London.
- Jacobson, P. (1977) *The syntax of crossing coreference*, Ph. D. Dissertation, University of California at Berkely. Re-published in 1980, by Garland, Inc., New York.
- Jacobson, P. (2000) "Paycheck pronouns, Bach-Peters sentences, and variable free semantics", *Natural Language Semantics*, 8.2: 77-155.
- Kanazawa, M. (2001) "Singular donkey pronouns are semantically singular," *Linguistics and Philosophy*, 24.3: 383-403.
- Karttunen, L. (1977) 'Syntax and semantics of questions,' *Linguistics and Philosophy*, 1.1: 3-44.
- Kayne, R. (1994), *The antisymmetry of syntax*, MIT Press, Cambridge, Ma.
- Lasnik, H. and T. Stowell (1991) "Weakest Crossover," *Linguistic Inquiry*, 22.4: 687-720.
- Lebeaux, D. (1988) *Language Acquisition and the Form of the Grammar*, Ph.D. Dissertation, University of Massachusetts, Amherst, MA.
- Longobardi G. (1994) 'Reference and proper names: A theory of N-movement in syntax and logical form,' *Linguistic Inquiry*, 25.4: 609-665.
- Marty, P. (2017) *Implicatures in the DP domain*, Ph. D. Dissertation, MIT, Cambridge, Mass.
- May, R. (1985) *Logical Form. Its structure and derivation*, MIT Press, Cambridge, Ma.
- Neale, S. (1990) *Descriptions*, MIT Press, Cambridge, Ma.
- Postal, P. (1968) *Crossover phenomena*, United States Air Force report, contract no. AF 19(628)-5127, Scientific report no. 3, IBM Thomas J. Watson Research Center. Re-published in 1971 by Holt, Rinehart and Winston, New York.
- Postal, P. (1969) "On the so-called 'pronouns' in English", in D. Reibel and S. Shane (eds) *Modern Studies in English*, Prentice-Hall, Englewood Cliffs.
- Postal, P. (1993) 'Remarks on weak crossover effects,' *Linguistic Inquiry* 24.3: 539-556.
- Reinhart, T. (1983) "Coreference and Bound Anaphora: A Restatement of the Anaphora Question," *Linguistics and Philosophy*, 6.1: 47-88.
- Rizzi, L. (1990) *Relativized minimality*, MIT Press, Cambridge, Ma.
- Rizzi, L. (2005) "On Some Properties of Subjects and Topics." *Proceedings of the XXX Incontro di Grammatica Generativa*. Venezia, Cafoscarina.
- Rooth, M. (1985) *Association with focus*, Ph. D. Dissertation, University of Massachusetts, Amherst, Ma.
- Ross, H., G. Chierchia and K. Davidson (2023) 'Quantifying weak and strong crossover for *wh*-crossover and proper names,' in M. Onoeva, A. Staňková and R. Šimík (eds) *Proceedings of Sinn und Bedeutung 27*, Prague, Czech Republic.
- Shan, C. and C. Barker (2006) 'Explaining crossover and superiority as left-to-right evaluation,' *Linguistics and Philosophy* 29.1: 91-134
- Torrego, E. (1994) 'On the Nature of Clitic Doubling,' *Anuario del Seminario de Filologia Vasca Julio de Urquijo; International Journal of Basque Linguistics and Philology*, 28.2: 199-214.
- Wu, Y.-C. (2024) 'The Seediq pivot and the A/A' distinction', Manuscript, Harvard University.
- Zubizarreta, M. L. (1998) *Prosody, focus and word order*, MIT Press, Cambridge, Ma.
- Zubizarreta, M. L. (2006) "Las functionas informativas: Tema y foco." In I. Bosque and V. Demonte (eds) *Gramatica descriptiva de la lengua Española*, Espasa Calpe, Madrid.