Chapter 10

Long-distance agreement and locality: A reprojection approach

Kristin Börjesson
Universität Leipzig

Gereon Müller
Universität Leipzig

Based on the assumption that all cases of long-distance agreement should be analyzed as involving only local agreement, and based on the observation that all existing approaches to long-distance agreement in terms of local agreement face substantial empirical and/or conceptual problems, the present paper sets out to develop a radically new approach to long-distance agreement. We suggest that long-distance agreement can and should be analyzed as a strictly local operation taking place early in the derivation, and giving rise to a counter-bleeding effect (i.e., apparent non-locality) later in the derivation as a consequence of regular syntactic structure building. More specifically, we argue that long-distance agreement involves (a) (what will become) a matrix verb V_1 which enters the syntactic derivation as part of a complex predicate V_1-V_2 that is merged with the embedded internal argument, agreeing with it locally early in the derivation; and (b) subsequent reprojection movement of V_1 out of V_2's clause, which eventually produces a biclausal structure (and thereby leads to a counter-bleeding effect with agreement). Empirical evidence for the new approach mainly comes from Nakh-Daghestanian languages, among them Hinuq, Khwarshi, and Tsakhur.

1 Introduction

Long-distance agreement is a phenomenon where agreement seems to take place in a non-local configuration, that is, across a clause boundary. More specifically, in cases of long-distance agreement, the verb in the matrix sentence agrees with

Kristin Börjesson & Gereon Müller. 2020. Long-distance agreement and locality: A reprojection approach. In Peter W. Smith, Johannes Mursell & Katharina Hartmann (eds.), *Agree to Agree: Agreement in the Minimalist Programme*, 307–346. Berlin: Language Science Press. DOI:10.5281/zenodo.3541761

respect to ϕ-features with an argument of the verb in an embedded sentence. A language that shows this phenomenon is Hindi-Urdu (see Mahajan 1990, Butt 1995; 2008, Bhatt 2005, and Chandra 2005, among others). A relevant pair of examples is given in (1). In Hindi, a DP qualifies as an agreement controller if it is not overtly case-marked (if both an external and an internal argument fail to be overtly case-marked, agreement is with the subject). Long-distance agreement is optional here: either there is agreement of both the matrix verb and the embedded verb with the embedded absolutive object DP (as in 1b), or the verbs show default agreement (as in 1a).

(1) a. Raam-ne [$_\alpha$ rotii khaanaa] chaahaa
Ram.MASC.-ERG bread.FEM eat.INF.MASC want.PERF.PST.MASC
'Ram wanted to eat bread.' (Mahajan 1990: 90)

b. Raam-ne [$_\alpha$ rotii khaanii] caahii
Ram.MASC-ERG bread.FEM eat.INF.FEM want.PERF.PST.FEM
'Ram wanted to eat bread.' (Mahajan 1990: 91)

Long-distance agreement is also widespread in Nakh-Daghestanian languages. A relevant pair of examples from Tsez is given in (2) (see Polinsky & Potsdam 2001). Agreement with respect to gender (III, in the case at hand) is controlled by absolutive DPs in Tsez. It always shows up on the embedded verb (if that verb can host overt agreement morphology in principle), and may then optionally also show up on the matrix verb (as in 2b); alternatively, there is no long-distance agreement, and the matrix verb exhibits default (IV) agreement marking (as in 2a).[1]

(2) a. Eni-r [$_\alpha$ už-ā magalu b-āc'-ru-łi]
mother-DAT boy-ERG bread.III.ABS III-eat-PSTPRT-NMLZ
r-iy-xo
IV-know-PRS
'The mother knows that the boy ate the bread.' (Polinsky & Potsdam 2001: 584)

[1]As noted by Bhatt (2005), some varieties of Hindi behave similarly in that they also exhibit an asymmetry between agreement with matrix as opposed to embedded verbs (such that embedded verbs can agree while matrix verbs do not have to), whereas other varieties of Hindi show a strict one-to-one correspondence (such that embedded verb agreement implies matrix verb agreement), as presupposed in the main text above.

b. Eni-r [$_{\alpha}$ už-ā magalu b-āc'-ru-ɬi]
mother-DAT boy-ERG bread.III.ABS III-eat-PSTPRT-NMLZ
b-iy-xo
III-know-PRS
'The mother knows that the boy ate the bread.' (Polinsky & Potsdam 2001: 584)

Another example for long-distance agreement in Nakh-Daghestanian languages comes from Hinuq; see (3a and 3b). Again, gender agreement is controlled by an absolutive DP, and gender-based long-distance agreement is optional.

(3) a. Saʔ ida-z r-eq'i-yo [$_{\alpha}$ Madina-y ɣi
Saida-DAT V-know.PRS Madina-ERG milk(IV).ABS
ga:-s-ɬi]$_{V}$
drink-RES-ABST
'Saida knows that Madina drank milk.' (Forker 2012: 468)

b. Saʔ ida-z y-eq'i-yo [$_{\alpha}$ Madina-y ɣi
Saida-DAT IV-know.PRS Madina-ERG milk(IV).ABS
ga:-s-ɬi]$_{V}$
drink-RES-ABST
'Saida knows that Madina drank milk.' (Forker 2012: 468)

Other languages exhibiting long-distance agreement are Itelmen (see Bobaljik & Wurmbrand 2005; a relevant pair of examples is given in 4), Innu-aimûn (see Branigan & MacKenzie 2002, with relevant examples in 5), Passamaquoddy (see Bruening 2001), Chukchee (see Bošković 2007) and Blackfoot (see Bliss 2009).

(4) a. Na netxa-in [$_{\alpha}$ kma jeβna-s]
he forget-3SG.SUBJ(INTRANS) me meet-INF
'He forgot to meet me.' (Bobaljik & Wurmbrand 2005: 50)

b. Na əntxa-βum=nm [$_{\alpha}$ kma jeβna-s]
he forget-1SG.OBJ=3.CL me meet-INF
'He forgot to meet me.' (Bobaljik & Wurmbrand 2005: 50)

(5) a. Ni-tshissenit-en [$_{\alpha}$ Pûn kâ-mûpisht-âshk]
1-know-TI Paul PRT-visited-2/INV
'I know that Paul visited you.' (Bobaljik & Wurmbrand 2005: 50)

b. Ni-tshissenim-âu [$_{\alpha}$ Pûn kâ-mûpisht-âshk]
1-know-3 Paul PRT-visited-2/INV
'I know that Paul visited you.' (Branigan & MacKenzie 2002: 389)

While long-distance agreement is optional in all these languages, there are also some languages where long-distance agreement is obligatory in certain environments; this holds, e.g., for Icelandic, Kutchi Gujarati and Chamorro. In this paper, we will concentrate on those languages in which long-distance agreement seems optional. Still, as with many other cases of syntactic optionality, it turns out that in all these cases, the choice of long-distance agreement in a sentence goes along with an interpretation of the controller of long-distance agreement as having a particular information structural status, i.e., an interpretation that the other member of the sentence pair – that with (only) local agreement – lacks.

The central challenge posed by long-distance agreement for syntactic theory is, of course, the apparent non-locality of the operation. More specifically, the embedded DP that controls the agreement would seem to be separated by a clause-like constituent (α in the examples above) from the matrix verb, and therefore be too far away to permit establishing a local relation. This is potentially problematic because most current syntactic theories do indeed postulate that syntactic operations (like agreement) are highly local (e.g., this holds for the Minimalist Program, HPSG, Categorial Grammar, and Optimality Theory); in line with this, in all these theories, apparently non-local dependencies like long-distance movement, long-distance reflexivization, long-distance case assignment, sequence of tense, and switch reference have successfully been reanalyzed as involving only fairly local operations (see Alexiadou et al. 2012 for an overview).

Against the background of the Minimalist Program, the question raised by long-distance agreement is how it can be ensured that the basic locality requirement imposed by the Phase Impenetrability Condition (PIC, Chomsky 2000; 2001; 2008; 2013) in (6) is respected by the operation.

(6) *Phase Impenetrability Condition* (PIC):
In phase α with head H, the domain of H is not accessible to operations outside α; only H and its edge are accessible to such operations.

If α in (1–5) qualifies as a phase in the sense of (6), the very existence of long-distance agreement seems to be at variance with the PIC.[2]
In what follows, we will argue that the PIC-induced locality problem with long-distance agreement is real, and has not been convincingly solved yet in any of the approaches to long-distance agreement that have been developed so far. In view of this state of affairs, we will propose a radically new approach in terms of complex predicate formation plus reprojection where long-distance agreement can be analyzed as a strictly local operation: The new analysis involves (i) a verb V_1 which enters the syntactic derivation as part of a complex predicate V_1-V_2 that is merged with the embedded internal argument, agreeing with it locally early in the derivation; and (ii) subsequent reprojection movement of V_1 out of V_2's clause, which eventually produces a biclausal structure (and thereby leads to a counter-bleeding effect with agreement).

The article is structured as follows. In Section 2, we will sketch the different types of existing approaches to the phenomenon of long-distance agreement. Given the PIC-based premise of strictly local application of all syntactic operations, we will point out individual problems the different approaches face as well as introduce some new data that turn out to be problematic for almost all of them. We will then go on to introduce the new approach in terms of reprojection in Section 3. In Section 4, which concludes the article, we will summarize the main features of the new approach and provide an outlook into how it might also be put to use in other contexts involving extraction from DPs, where there is evidence for an extremely local relation of two heads that show up in two separate domains in syntactic output structures.

2 Existing analyses

At least for our present purposes, four different kinds of analysis of long-distance agreement can be distinguished. We will refer to these as (i) non-local analyses (where long-distance agreeement can apply in a non-local fashion), (ii) small structure analyses (where there is no phase boundary), (iii) Cyclic Agree analyses (where the information relevant for agreement is locally passed on through

[2] As a matter of fact, simple cases of T agreeing with a nominative object DP in VP in a language like Icelandic are already problematic from the point of view of the PIC in (6). For this reason, Chomsky (2001) also envisages a second, somewhat more liberal version of the PIC, where a phase domain becomes opaque only when the next phase is reached. However, even this less restrictive version of the PIC would not suffice to straightforwardly derive the possibility of long-distance agreement. Furthermore, one might argue that the head actually responsible for this agreement into VP is not T but v. We will abstract away from this issue in what follows, and presuppose the PIC in (6) for the remainder of the paper.

the tree, originating with the controller and ultimately reaching the matrix verb), and (iv) feeding analyses (where movement of the controller makes local agreement possible). We will discuss these four analysis types in turn.[3]

2.1 Type (i): Non-local analyses

Non-local analyses either assume that the locality constraint on Agree is weaker than the original PIC (see Chomsky 2001 and footnote 2), or that Agree is not, in fact, subject to such a strict locality constraint in the first place (see Sells 2006, Bošković 2007, Keine 2016).

For example, Bošković's (2007) analysis relies on a revised Agree operation, one which is not subject to either the Activity Condition (7d in the original Agree definition in 7) or the PIC.

(7) *Agree*
α may Agree with β iff:
   a. α carries at least one unvalued and uninterpretable feature and β carries a matching interpretable and valued feature
   b. α c-commands β
   c. β is the closest goal to α
   d. β bears an unvalued uninterpretable feature.

Thus, the matrix verb can look all the way down to the embedded absolutive DP to check its ϕ-features. Bošković (2007) assumes that finite complement clauses can in principle be CPs or TPs. Thus, in complement clauses in which there is no evidence for a CP layer, one might as well asume that those clauses actually lack it. The idea, then, is that CPs block long-distance agreement in Tsez while TPs allow it. This is because, Bošković (2007) claims, CPs, in contrast to TPs, may carry ϕ-features that need to get checked. In the case of Tsez, CPs (and those of the languages that lack long-distance agreement altogheter), he assumes that they do carry such (default-valued) ϕ-features, which makes it possible for the matrix verb to locally agree with the CP as such, leading to local agreement. Long-distance agreement in cases that involve CPs is impossible due to the condition on Agree to involve the potential goal closest to the probe (see 7c).

[3]We will not consider a fifth type of analysis, where the matrix verb locally agrees with a covert pronoun, which in turn is coindexed (and therefore shares ϕ-features) with an embedded DP. As shown in Polinsky & Potsdam (2001) and Bhatt & Keine (2016), such "proxy agreement" is not a viable alternative in general. (Also, it is worth noting that such an analysis solves one locality problem (seemingly non-local agreement) by shifting it to another, well-known locality problem (seemingly non-local binding chains)).

Whatever the merits of this proposal, it is clear that it is incompatible with the strict PIC in (6): Since the potential long-distance agreement controller in transitive sentences is the internal argument bearing absolutive case, which is base-generated in VP, the PIC will predict that long-distance agreement should not be possible, independently of whether the embedded vP phase has both a TP and a CP projection on top of it, or just a TP projection.

2.2 Type (ii): Small structure analyses

In small structure analyses, it is argued that the configuration of matrix verb and long-distance agreement-trigger is local after all. The structure of the complement clause is assumed to be smaller than might be thought from the surface data (e.g., a VP in Boeckx 2004; an InflP in Bhatt 2005).

Like Bošković's (2007) approach, Bhatt's (2005) analysis is based on a revised Agree operation – in this case, it is one which also is not subject to the Activity Condition, but which does respect the PIC. To account for apparent long-distance agreement in Hindi, Bhatt (2005) assumes that complement clauses to matrix verbs that allow long-distance agreement are in fact only InflPs/VPs, which lack an external argument (i.e., they have no PRO) and are thus not phases. In contrast, long-distance agreement out of finite clauses is not possible in Hindi because finite complement clauses are CPs; the same prediction arises for infinitival structures that contain a PRO subject. In line with this, as far as the optionality of long-distance agreement in Hindi is concerned, Bhatt (2005) explains it by assuming that matrix verbs that allow long-distance agreement have an option of selecting either a restructuring infinitive or a non-restructuring infinitive, where the latter involves a syntatically projected PRO subject. This PRO intervenes between the matrix verb and the embedded object and, thus, blocks long-distance agreement (in the same way, a problem with the PIC would arise in this context).

Bhatt's (2005) analysis may work well for a language like Hindi, but it does not carry over to long-distance agreement in languages of the Nakh-Daghestanian type, where the external argument is clearly present (bearing ergative) in the embedded clause; see, e.g., (2b) from Tsez and (3b) from Hinuq. Given the strict version of the PIC in (6), it is clear that the presence of an external argument DP uncontroversially implies the presence of a vP phase, and this should make agreement of a matrix V with an internal argument DP included in the complement domain of v impossible, independently of whether vP qualifies as α (in the above sense) or not (i.e., independently of whether there is additional structure on top of vP in the complement of the matrix V).[4]

[4] As a matter of fact, to account for evidence of the Nakh-Daghestanian type, Bhatt (2005: 791) ultimately concludes that a feeding account along the lines of Polinsky & Potsdam (2001) is independently called for; see below.

2.3 Type (iii): Cyclic Agree analyses

In Cyclic Agree analyses, it is assumed that what looks like long-distance agreement actually is to be decomposed into a series of shorter agreement steps, all of of which obey strict locality. On this view, first the embedded verb agrees with the embedded agreement controller DP; second, the matrix verb agrees with the embedded verb; third, by transitivity, this implies that the matrix verb will eventually agree with the embedded DP, albeit indirectly. This kind of analysis has been pursued by Butt (1995), Legate (2005), Keine (2008), Preminger (2009), and Lahne (2012), among others. As an illustration, consider the specific approach developed in Legate (2005).

The basic premise of this approach is that at no stage of the derivation is there an Agree relation between the matrix verb and the embedded DP. Rather, the agreement controller DP's ϕ-features first valuate an $[u\phi]$ probe feature of a phase head, which by definition (cf. the PIC in 6) is also part of the higher phase. The matrix verb then probes the embedded phase head's ϕ-features. Thus, the embedded phase head acts as a hinge between the matrix and embedded domains. This accounts for the observation that long-distance agreement presupposes the existence of local agreement in the embedded clause.

The Cyclic Agree approach solves the locality problem with long-distance agreement in a very simple manner that directly corresponds to the analogous (and by now well-established) treatment of long-distance movement in terms of successions of smaller movement steps.[5] However, there are both conceptual and empirical problems raised by Cyclic Agree approaches to long-distance agreement. On the one hand, Cyclic Agree is conceptually problematic from a minimalist perspective, given standard assumptions about probe features, goal features, and the Agree operation: It looks as though one and the same set of ϕ-features (on the phase head in the middle) must act as a probe in one case, and as a goal in another (see Bhatt 2005). On the other hand, there is an empirical problem (see Polinsky & Potsdam 2001, Bhatt & Keine 2016) that is due to the fact that Cyclic Agree approaches rely on transitivity. The problem is that if two verbs V_1 and V_2 can in principle participate in local agreement in some long-distance agreement constructions, and V_2 and DP_{abs} can participate in local agreement, then

[5] As a matter of fact, an alternative local analysis of long-distance agreement that mimicks SLASH feature percolation as it has been proposed for movement dependencies (see Gazdar 1981) might in principle also be an option. This would then express a similarity of the two operations in long-distance contexts (viz., movement and agreement) even more straightforwardly. However, to the best of our knowledge, such an analysis has not yet been proposed. That notwithstanding, it would be subject to the same empirical problem mentioned in the main text below.

long-distance agreement involving V_1 and DP_{abs} must also be possible. However, this is not always the case: For instance, in Tsez, long-distance agreement, unlike local agremeent, requires DP_{abs} to be a topic (see Polinsky & Potsdam 2001). This is shown by the examples in (8), where (i) local agreement of the embedded predicate and the absolutive DP as the agreement controller is possible throughout, (ii) the matrix and embedded predicates can participate in long-distance agremeent in principle, but (iii) long-distance agreement in this configuration is blocked nevertheless because the absolutive DP is not interpreted as a topic: It is interpreted as a focus in (8a), and it shows up as an – inherently non-topicalizable – reflexive pronoun in (8c). In both cases, the matrix verb can only carry out agreement with the embedded clause (α) itself; cf. (8b–8d).

(8) a. * Eni-r [$_{\alpha_{IV}}$ t'ek-kin y-igu yāł-ru-łi]
mother-DAT bookII.ABS-FOC II-good be-PSTPRT-NMLZ
y-iy-xo
II-know-PRES
'The mother knows that the BOOK is good.'
(Polinsky & Potsdam 2001: 611)

b. Eni-r [$_{\alpha_{IV}}$ t'ek-kin y-igu yāł-ru-łi]
mother-DAT bookII.ABS-FOC II-good be-PSTPRT-NMLZ
r-iy-xo
IV-know-PRES
'The mother knows that the BOOK is good.'
(Polinsky & Potsdam 2001: 611)

c. * Eni-r [$_{\alpha_{IV}}$ už-ā nesā že žāk'-ru-łi]
mother-DAT boy-ERG REFL.I.ABS beat-PSTPRT-NMLZ
∅-iy-xo
I-know-PRES
'The mother knows that the boy beat himself up.'
(Polinsky & Potsdam 2001: 612)

d. Eni-r [$_{\alpha_{IV}}$ už-ā nesā že žāk'-ru-łi]
mother-DAT boy-ERG REFL.I.ABS beat-PSTPRT-NMLZ
r-iy-xo
IV-know-PRES
'The mother knows that the boy beat himself up.'
(Polinsky & Potsdam 2001: 612)

2.4 Type (iv): Feeding analyses

2.4.1 Movement feeds agreement

Fourth and finally, Polinsky & Potsdam (2001) have argued that a local approach to long-distance agreement in Tsez is both technically feasible and empirically supported. In their view long-distance agreement involves feeding of local agreement by movement. The basic assumption is that the agreement controller (an absolutive DP in Tsez) moves to a position in which it can locally agree with the matrix verb. However, there are two complications to this simple picture. First, Polinsky & Potsdam present strong arguments against the assumption that displacement of the agreement controller DP ends up in the matrix clause itself. For one thing, all established movement operations in Tsez are strictly clause-bound, so the operation that feeds long-distance agreement would be the only type of movement that could leave a clause. For another, Polinsky & Potsdam (2001) observe that long-distance agreement in Tsez never co-occurs with scope reversal; in other words, a DP that participates in long-distance agreement with a matrix verb can never take scope over quantifed items in the matrix clause. This latter property is illustrated in (9): Independently of whether long-distance agreement takes place (see 9b) or not (see 9a), the embedded absolutive DP (with a universal quantifier in 9) cannot take scope over a matrix subject (with an existential quantifier in the case at hand).

(9) a. Sis učiteler [$_{\alpha}$ šibaw uži ∅-ik'ixosi-ɬi] r-iy-xo
one teacher every boyI.ABS I-go-NMLZ IV-know-PRS
'Some teacher is such that he knows that every boy is going.'
*'Every boy is such that some teacher knows that he is going.' (Polinsky & Potsdam 2001: 618)

b. Sis učiteler [$_{\alpha}$ šibaw uži ∅-ik'ixosi-ɬi] ∅-iy-xo
one teacher every boyI.ABS I-go-NMLZ I-know-PRS
'Some teacher is such that he knows that every boy is going.'
*'Every boy is such that some teacher knows that he is going.' (Polinsky & Potsdam 2001: 619)

From these considerations it follows that the postulated movement operation cannot actually end up in the matrix clause in cases of long-distance agreement. The second complication involves the overt/covert distinction of movement operations. Since the absolutive DP that participates in long-distance agreement does not have to be overtly displaced and typically shows up in its in situ position (or, more generally, given that Tsez exhibits variable word order: it shows

up in its unmarked position), it is clear that the movement operation that feeds long-distance agreement must be a covert one.

Against this background, Polinsky & Potsdam's (2001) proposal is the following: There is covert, information structure-driven movement of the long-distance agreement-controller into a higher domain (phase) of the same clause, and the position thus reached provides a local enough configuration with the matrix verb to make Agree with it possible. More specifically, Polinsky & Potsdam's (2001) analysis works as follows.

A crucial basic assumption is that the size of embedded clauses in Tsez is variable. (10) gives the maximal syntactic structure for a clause in Tsez. This structure is only fully built up when needed; i.e., clauses are CPs if they exhibit material that belongs in this layer (e.g., a C head) but not otherwise; a TopP is projected if the clause contains a topic; and so forth.

(10) *Clause structure for Tsez*:
$[_{CP}\ [_{C'}\ C\ [_{TopP}\ [_{Top'}\ Top\ [_{TP}\ DP_{erg}\ [_{T'}\ T\ [_{VP}\ DP_{abs}\ V]]]]]]]$

On this basis, Polinsky & Potsdam (2001) assume that, generally, long-distance agreement-allowing matrix verbs can select a TP as complement. However, to derive long-distance agreement in Tsez, it is postulated that a topic-marked long-distance agreement-controlling element is covertly moved to the specifier of TopP in the left periphery of the complement clause. This movement brings the triggering element into a sufficiently local relation to the matrix verb to allow the latter to check its uninterpretable ϕ-features against those of the covertly topicalised element, resulting in long-distance agreement. Long-distance agreement is, thus, taken to be a reflex of the topic-status of the triggering element.

A CP is predicted to block long-distance agreement, as is a bare TP, given that a DP_{abs} agreement controller does not occupy SpecT (at this point the analysis is not fully PIC-compatible). The relation that Polinsky & Potsdam (2001) assume to underlie long-distance agreement is Head Government (not Agree, as in 7), which, following Rizzi (1990), is understood as in (11).

(11) *Head Government*:
X head-governs Y iff:
   a. X ∈ {A, N, P, V, H[+ tense]}.
   b. X m-commands Y.
   c. No barrier intervenes.
   d. Relativized Minimality is respected.

It is then postulated that it can be derived from (11) that "a head governs its specifier, its complement, an element adjoined to its complement, and the specifier of its complement" (Polinsky & Potsdam 2001: 627), but not, say, the specifier of the complement of the head's complement. However, as a matter of fact, it is not quite clear why this should be the case. Consider (12).

(12) $[_{XP}\ X\ [_{ZP}\ Z\ [_{WP}\ YP\ [_{W'}\ \ldots\ ]]]]$

Does X head-govern YP in (12)? First, suppose that X is one of the possible items mentioned in (11a). Second, X clearly m-commands YP. Third, Relativized Minimality is respected in (12) because there is no intervening phrase that could induce a Relativized Minimality effect for X and YP (there is no phrase that c-commands YP and is c-commanded by X – ZP and WP both dominate YP). That leaves the presence of a barrier as the only possible source of a failure of head government of YP by X. Whether ZP or WP is a barrier in (12) depends on the exact definition of this concept (which Polinsky and Potsdam do not provide). The first thing to note is that both ZP and WP are complements, i.e., sisters of X^0 categories. This will suffice to exempt them from barrier status in most of the available conceptions of barriers (see, e.g., Cinque 1990). In contrast, according to the more complex, two-stage definition of barrier in Chomsky (1986), ZP might in fact emerge as a barrier in (12) if WP can be classified as a blocking category that passes on its status as a "virtual barrier" to the phrase immediately above it. So it seems that only under this complex approach, based on blocking categories vs. real barriers, can it be derived that X does not head-govern YP in (12).

Based on the assumption that all agreement relations are subject to (11) (or at least the general consequences that (11) is supposed to have), a number of restrictions that Polinsky & Potsdam observe for long-distance agreement in Tsez follow. First, a CP can never be projected in long-distance agreement contexts because "it would block government of SpecTop by the verb" (Polinsky & Potsdam 2001: 638). Note that this presupposes that ZP (= CP) would indeed qualify as a barrier in (12) that makes head government of YP (= DP_{abs} in SpecTop) by matrix V (= X) impossible. (As we have just seen, this consequence is far from straightforward.) However, with this qualification, it can be derived that long-distance agreement is impossible (i) in the presence of a wh-phrase in a clause (which inherently activates the CP layer, whether or not wh-movement takes place overtly), and (ii) in the presence of the element *λin*, which is assumed to be a designated C element.[6] A third prediction is that long-distance agreement with

[6] As noted by Polinsky & Potsdam (2001: fn. 20), long-distance agreement in the presence of a wh-phrase would ceteris paribus be expected to be possible if the wh-phrase is itself the absolutive argument and occupies SpecC.

D_{abs} is impossible if some other XP functions as the topic in a clause.[7] These qualifications notwithstanding, Polinsky & Potsdam's (2001) analysis would seem to derive licit and illicit cases of long-distance agreement in a very simple and elegant way that furthermore respects locality considerations. Still, as shown in the following subsection, there are conceptual and empirical problems with this approach.

2.4.2 Problems with the feeding approach

2.4.2.1 The nature of covert topic movement

First, as noted by Bošković (2007), the crucial postulation of a covert topicalization operation for Tsez is far from innocuous. There is virtually no independent evidence that such an operation exists. Also, there is a real danger of an ordering paradox: It is not really clear how *covert* movement at LF can trigger *overt* agreement – if the movement takes place at LF, it comes too late.[8]

[7] Again, though, details of the account are somewhat unclear. One might think that this effect is due to intervention, i.e., the Relativized Minimality part of the definition of Head Government. However, if there can only be one topic per clause in Tsez, and that is not DP_{abs}, then DP_{abs} can never reach the position (viz., SpecTop) it needs to reach to enable long-distance agreement, and a resort to intervention is not necessary. If, on the other hand, there can be more than one topic per clause, then it is not obvious why DP_{abs} should not qualify as the structurally highest one of them, thereby circumventing an intervention effect.

[8] This problem can in principle be solved by assuming that covert movement is actually movement taking place in the narrow syntax, with the only difference to overt movement being that the lowest copy of a complex chain is subject to phonological realization (rather than the highest member, as with overt movement). However, even if one were to adopt such an approach based on the copy theory of movement, the intended effect does not seem to arise anywhere else. As far as we can tell, other instances of covert movement that have been suggested in the literature (e.g., for certain cases of wh-in situ) never feed Agree operations. Thus, compare (i.a), where overt wh-movement gives rise to new options for reflexivization (assumed here to be an instance of Agree, see Reuland 2011 for discussion), with (i.b), where covert wh-movement fails to produce the same effect (see Barss 1986).

(i) a. $John_1$ wonders [$_{CP}$ [$_{DP}$ which book about himself$_1$] Bill bought]
b. * $John_1$ wonders [$_{CP}$ why Bill bought [$_{DP}$ a book about himself$_1$]]

In the same way, covert wh-movement (unlike overt wh-movement) never seems to feed case assignment or agreement in the world's languages. – All that said, Polinsky & Potsdam (2001: 626) would seem to exclude a reinterpretation of covert movement as overt movement plus pronunciation of the lower copy when they explicitly state that the "syntactic agreement configuration between the probe and the absolutive trigger is created at LF".

2.4.2.2 **Complementizers**

There are two complementizer-like items in Tsez viz. *łi*, which permits long-distance agreement (and shows up in all Tsez examples exhibiting long-distance agreement disussed above), and *λin*, which blocks long-distance agreement. As noted, Polinsky & Potsdam (2001) assume that *λin* is indeed a regular C item, and assuming that the presence of a CP makes head government impossible, it is correctly predicted that there is no long-distance agreement across *λin*. However, for the same reason, it must be assumed that *łi* is not a C element. The problem here is that this is exactly what it looks like, given that, like *λin*, it is the outermost head in the word containing V. What is more, it does not yet suffice to assume that *łi* is not a C element – *łi* must be assumed not to be structurally represented *at all*. The reason is that if *łi* were the head of a phrase (outside of TopP), it would block long-distance agreement in the same way as *λin*. It remains unclear whether there is any independent evidence for such a radically different treatment (projecting complementizer vs. structure-less morphological marker) of the two items. (In the analysis to be developed in Section 3 below, we will presuppose that both *λin* and *łi* are regular C items.)

We take these first two problems to be potentially worrisome but certainly not decisive. Arguably, things are different with the next two issues raised by Polinsky & Potsdam's analysis, concerning a semantic problem based on the assumed covert DP movement, and an incompatibility of the analysis with what look like clear cases of long-distance agreement across a CP boundary.

2.4.2.3 **Topic interpretation within the embedded clause**

Polinsky & Potsdam (2001) assume that the landing site of the abstract movement is in the left periphery of the embedded clause. Accordingly, the long-distance agreement-controlling DP is interpreted as the topic of the embedded clause. A problem with this analysis is that information structure phenomena – and root phenomena in general – are usually confined to clauses that have some illocutionary force; see Hooper & Thompson (1973), Ebert et al. (2008), Krifka (2014), and Matić et al. (2014). Most long-distance agreement-allowing matrix verbs, however, are factives, which semantically take sentence radicals (see Stenius 1967), i.e. propositions, as complements (see Krifka 2004) and also syntactically involve smaller structures (see de Cuba & Urogdi 2010). Thus, complements of factives do not involve any illocutionary operator in their syntax/semantics – only the matrix clause does. Under a structured proposition approach (see Krifka 1992) to information structural phenomena, this leads to a semantic representation in

which the topic can only be understood as the topic of the whole sentence. That is, the predicted structure for (2b) (repeated in 13) is as in (14).

(13) Eni-r [$_{\alpha}$ užā magalu b-āc'-ru-łi] b-iy-xo
mother-DAT boy-ERG bread.III.ABS III-eat-PSTPRT-NMLZ III-know-PRS
'The mother knows that the boy ate the bread.' (Polinsky & Potsdam 2001: 584)

(14) ASSERT(⟨$_{T}$ λP.**mother**(λz.**boy**(λx.P(λy.**eat**(x,y))) (λp.**know**(z,p))), **bread**⟩)

The felicity conditions of the ASSERT-operator for topic-comment structures make reference to the first part of the structured proposition as a whole. Thus, it is unclear whether the embedded topic interpretation advocated by Polinsky & Potsdam (2001) is actually available; i.e., whether (13) *is* actually understood as paraphrased in (15a), or not rather as in (15b).

(15) *Readings for topics in long-distance agreement*
a. The mother knows that, as for the bread, the boy ate it.
b. As for the bread, the mother knows that the boy ate it.

More specifically, complex sentences with factive matrix verbs presuppose the truth of the proposition denoted by the respective complement clause. In terms of information structure, factive presuppositions belong to the "background" of an utterance and they are "taken for granted". Thus, they are not "at-issue" or "under discussion". This characterization is inconsistent with Polinsky & Potsdam's (2001) assumption that a DP embedded under a factive verb can act as topic of the embedded clause.

2.4.2.4 Long-distance agreement across a CP boundary

A severe empirical problem for the analysis developed in Polinsky & Potsdam (2001) (but also for analyses of the small structure type) is that there is evidence from other Nakh-Daghestanian languages that strongly suggests that long-distance agreement is in principle possible across a CP boundary, and without movement to SpecC (recall footnote 6).

Thus, Khwarshi (see Khalilova 2009) and Hinuq (see Forker 2012), two Nakh-Daghestanian languages closely related to Tsez, also exhibit long-distance agreement. Similarly to Tsez, this also goes along with a prominent information structural status of the triggering DP. In contrast to Tsez, however, in these languages,

the triggering NP can have either topic or focus status. For instance, in the Khwarshi examples in (16), the absolutive DP is interpreted as the topic of the embedded clause, and long-distance agreement is possible (see Khalilova 2009: 387). However, long-distance agreement is also possible in answers to information questions, as in (17), suggesting that the long-distance agreement-controlling DP_{abs} may also function as the focus of the embedded clause.

(16) a. Išet'u-l l-iq'-še goli uža bataxu y-acc-u
mother.OBL-LAT IV-knows-PRS COP boy.ERG bread(V) V-eat-PST.PTCP
'Mother knows that the boy ate bread.' (Khalilova 2007: 116)

b. Išet'u-l y-iq'-še goli uža bataxu y-acc-u
mother.OBL-LAT V-knows-PRS COP boy.ERG bread(V) V-eat-PST.PTCP
'As for the bread, mother knows that the boy ate it.'
(Khalilova 2007: 117)

(17) a. (Which cow does the boy know came?)

b. Uža-l l/b-iq'-še kˤaba zihe b-ot'uq'-u
boy.OBL-LAT IV/III-know-PRS black cow(III) III-come-PST.PTCP
'The boy knows that the black cow has come.'
(Khalilova 2007: 118)

Similar facts obtain in Hinuq.

Importantly, long-distance agreement in Khwarshi and Hinuq is also less restricted in another respect: There are cases in which a *wh*-element occurs in an interrogative complement clause, and long-distance agreement is nevertheless available. This is shown for Khwarshi in (18) (see Khalilova 2007).

(18) Uža-l l/b-iq'-še [$_{CP(IV)}$ łu$_{[foc]}$ zihe
boy.OBL-LAT IV/III-know-PRS who.ERG cow(III)
b-iti-xx-u]
III-divide-CAUS-PST.PTCP
'The boy knows who has stolen the cow.'

Here a wh-phrase bearing ergative case shows up in the embedded interrogative clause. Given standard assumptions about the semantics of questions (see, e.g., von Stechow 1996), the interrogative interpretation of a clause is inherently, and invariably, tied to the presence of a C element. Therefore, (18) proves that long-distance agreement across a CP boundary is possible in Khwarshi independently of whether the ergative wh-phrase can be assumed to be located in SpecC in the syntax, and of whether or not there is an overt C item present.

Next, (19) and (20) illustrate the possibility of long-distance agreement across a CP boundary in Hinuq (see Forker 2012). In (19), the embedded interrogative clause contains a subject wh-phrase that does not block long-distance agreement with the (non-wh) absolutive DP.

(19) [$_{CP}$ łu rek'^{w}e goɬiš] diž ∅-eq'i-yo gom
who man.I be.CVB I.DAT I-know-PRS be.NEG
'I do not know who he was.' (Forker 2012: 637)

In contrast, (20) shows that long-distance agreement is also possible with an absolutive DP that is a wh-element itself; but there is no reason to assume that DP_{abs} is not in its base position here.

(20) Debez r/∅-eq'i-ye [$_{CP(V)}$ k'ačaɣ-za-y łu ∅-uher-iš-łi] ?
you.SG.DAT V/I-know-Q bandit-OBL.PL-ERG who I-kill-RES-ABST
'Do you know whom the bandits killed?' (Forker 2012: 637)

Note that (19) and (20) contain clause-final elements that would seem to correspond to *łi* rather than *λin* in Tsez. However, notwithstanding the problems raised by Polinsky & Potsdam's (2001) analysis of *łi* mentioned above, and notwithstanding the arguments that Forker presents for a uniform CP analysis of these contexts in Hinuq after all, it is clear that the status of *łi* has no bearing on the question of whether there is a CP present in (19) and (20) (and 18, for that matter): There must be a CP boundary here because of the combination of interrogative semantics and a full clausal structure, including assignment of ergative (which by itself is not yet decisive, given that there are good arguments for assuming that the ergative is assigned within vP in Nakh-Daghestanian languages; see Gagliardi et al. 2014, Polinsky 2016).

Thus, there is strong evidence from Khwarshi and Hinuq for the general availability of long-distance agreement across what must qualify as a CP.

What is more, unlike Tsez, the non-Tsezic Nakh-Dahgestanian language Tsakhur also permits what can be called "super-long-distance agreement", i.e., long-distance agreement across two clause boundaries; see Kibrik (1999).

(21) Ičī-s w=uk̄īkɨn-na [$_{CP}$ jičo-j-s [$_{CP}$ gaba hāʔ-as]
girl-DAT 3=want.PF-AA sister-OBL-DAT carpet.3 3.do-POT
XaIr-qi=w=x-es]
learn-3=become-POT
'The girl wants her sister to learn to make a carpet.'
(Kibrik 1999)

Similar phenomena have also been reported by Forker (2012) for Hinuq; see (22).

(22) Iyo-z b-eqʼi-yo [$_{CP}$ Patʼimat-ez [$_{CP}$ tort b-acʼ-a]
mother-DAT III-know-PRS Patimat-DAT cake(III) III-eat-INF
b-eti-š-łi]]
III-want-RES-ABST
'The mother knows that Patimat wanted to eat the cake."
(Forker 2012: 633)

If there are two CP boundaries present, there is no way for the feeding approach to account for the option of long-distance agreement since the covert movement postulated in this approach always has to be clause-bound.

2.5 Interim conclusion

We take it that the conclusion that can be drawn on this basis is that all four existing approaches face significant problems with long-distance agreement from the point of view of a grammar that incorporates a strict locality principle like the PIC. Thus, there is every reason to pursue a new approach; and given the general availability of long-distance agreement across a CP, this approach must be such that it preserves strict locality even if there can be no denying the fact that the matrix verb and the embedded agreement controller DP can be far away from one another in structural terms in syntactic surface representations.

As a basic premise, we will assume that the only way to locally model non-local dependencies is via *movement* (see Hornstein 2001; 2009 for this general point, and Müller 2014 for some specific proposals in a priori recalcitrant domains). Given the PIC, a matrix V and an embedded (agreement-controlling) DP have to enter a local relation at some point of the derivation. As we have seen, there is evidence against the assumption that DP moves to the matrix V domain (or to a position of the embedded domain that is accessible from it); this excludes feeding analyses. The only remaining possibility then is that it is actually V that moves to the matrix domain: If the mountain won't come to the prophet, the prophet will go to the mountain.

For concreteness, we would like to propose that locality in long-distance agreement is not established *late* in the derivation (as in Polinsky & Potsdam 2001's approach, where movement feeds long-distance agreement), but, in fact, *early*. This approach thus involves counter-bleeding (rather than feeding): Agreement with the embedded internal argument DP takes place at a stage in the derivation when DP and the two verbs involved are all clause-mates. It is only due to subsequent reprojection movement of what will eventually become the matrix verb

that on the surface it looks as if agreement takes place long-distance; reprojection movement of V thus comes too late to bleed (i.e., it counter-bleeds) Agree.

3 A new analysis

3.1 Head movement as reprojection

Let us begin by sketching the outlines of a general approach to head movement in terms of reprojection, a concept that has been widely pursued for various empirical domains over the last decades (see Pesetsky 1985; von Stechow & Sternefeld 1988; Sternefeld 1989; Holmberg 1991; Ackema et al. 1993; Kiss 1995; Koeneman 2000; Haider 2000; Bhatt 2002; Hornstein & Uriagereka 2002; Fanselow 2003; 2009; Bury 2003; Surányi 2005; Donati 2006; Bayer & Brandner 2008; Georgi & Müller 2010; Müller 2011; S. Müller 2016, among others).[9] The basic idea behind head movement as reprojection is that an X^0 head is moved out of a projection that dominates it and takes this projection as its own complement by merging with it, projecting anew in the derived position. This solves the notorious c-command and Extension Condition (cf. Chomsky 1995) problems with head movement as adjunction to an X^0 category: In a head-movement-as-adjunction structure like (23), the moved head Y fails to extend the tree (since XP must, be definition, have been in place before movement of Y), and Y does not c-command its trace (because the next branching node containing Y is the higher X segment, which does not dominate Y's trace).

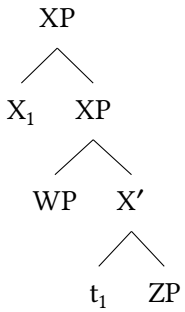
(23) $[_{XP}\ [_{X}\ Y_1\ X\ ]\ [_{WP}\ \ldots\ t_1\ \ldots\ ]]$

These problems disappear under a reprojection approach: Head movement has now extended the tree, and the moved item is able to c-commands its trace. Furthermore, this approach does not necessitate (i) a relocation of head movement to PF (see Chomsky 2000), (ii) a reinterpretation as XP movement (see Koopman & Szabolcsi 2000, Mahajan 2001, and Nilsen 2003, among many others), or (iii) the postulation of a complex operation integrating both regular syntactic movement and syntactically irregular morphological merger (see Matushansky 2006).

There are basically three different reprojection scenarios. A first possibility is that a head moves out of its own projection, merges with the XP of which it was the head prior to the movement, and projects anew. Such local reflexive reprojection is shown in (24).

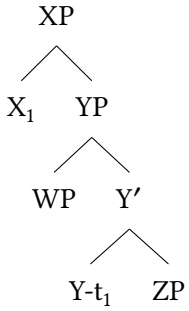
[9] Pesetsky (1985) suggests that reprojection after head movement at LF serves to circumvent bracketing paradoxes. As far as we can tell, this qualifies as the first instance of a reprojection approach to head movement in the literature.

(24) *Local reflexive reprojection*:



A second possibility is that a reprojection movement is still highly local (in the sense that the moved head attaches to the minimal phrase that dominated it before the movement step was carried out), but not reflexive. In this scenario, the moved head excorporates from a complex head structure that was formed by an earlier (possibly pre-syntactic) operation combining two primitive X^0 categories (in accordance with c-command and Extension Condition requirements), or that is stored as such in the lexicon; after the movement, the moved head projects its own XP in the derived position.[10] Local non-reflexive reprojection is illustrated in (25).

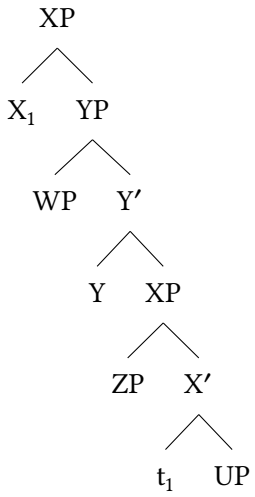
(25) *Local non-reflexive reprojection*:



Finally, reprojection can be non-local (by definition, it is then also non-reflexive). In (26), the moved head skips over two maximal projections and reprojects in the derived position.

[10]Thus, strictly speaking, this is not actually an instance of *re*-projection: X in (25) projects for the first time in the derived position.

(26) *Non-local non-reflexive reprojection*:



Assuming these three scenarios to be available in the world's languages, it can be concluded that head movement can involve excorporation (see Roberts 1991; 1997), and that head movement does not obey the Head Movement Constraint (see Roberts 2009; 2010 vs. Travis 1984 for arguments to this effect). Given that the data that originally motivated stipulation of the excorporation and Head Movement Constraint restrictions can be derived otherwise, this would seem to permit a simpler, more attractive theoretical approach, and to correspond to the null hypothesis. Furthermore, one should expect that head movement as reprojection obeys the same constraints that hold of all movement operations; this includes the PIC (see 6). Thus, for the operation to be legitimate, it can be concluded that YP is not a phase in (26); and that, more generally, head movement as reprojection can cross phases by carrying out intermediate movement steps to phase edges, in accordance with the PIC.

As for the concrete mechanics of reprojection movement, we will make the following assumptions. First, all syntactic operations are feature-driven: On the one hand, there are designated structure-building features (edge features, subcategorization features) that trigger (external or internal) Merge; we will refer to these as [•F•] features. On the other hand, there are probe features that trigger Agree. To simplify exposition and simultaneously avoid commitment to one of the existing options in various domains (e.g., valuation vs. checking, interpretability vs. uninterpretability), we will refer to probes as [⁎F⁎] features throughout. All these features triggering syntactic Merge and Agree operations are ordered on lexical items; and they are discharged (i.e., rendered syntactically inactive) one after the other after having induced the respective operations that they encode.

Finally (although this assumption will not actually be crucial), we postulate that all phrases are phases. As a consequence, movement must take place via all intermediate phrase edges that intervene between a base position and the ultimate landing site of some moved item (except for the minimal specifier domain if the item is already part of the phase edge, as is the case with reprojection movement of heads). Given this assumption, YP in (26) must be a phase, and X_1 must therefore carry out an intermediate step to SpecY on its way to its ultimate position.[11]

Suppose further that Featural Cyclicity holds, as in (27).[12]

(27) *Featural Cyclicity*:
A non-root XP cannot contain a feature δ in the non-edge domain of X that is supposed to trigger an operation ([•F•] or [⁎F⁎]).

In the normal course of events, the head X of some XP has discharged all the Merge-inducing features ([•F•]) and Agree-inducing features ([⁎F⁎]) it contains before XP is merged with some other category. However, suppose that the head X has not been able to discharge a [•F•] or [⁎F⁎] (plus, possibly, other features that are lower on the list of operation-inducing features of the head, and that can only be accessed if the topmost feature has been discharged). In such a situation, one of two Last Resort operations may take place: Either the [•F•] or [⁎F⁎] feature is deleted (see Béjar & Řezáč 2009, Preminger 2014, and Georgi 2014 for proposals along these lines); or the item containing the incriminating feature is moved to the edge domain of the current phrase, so as not to violate Featural Cyclicity in (27). The two Last Resort options for [•F•] and [⁎F⁎] features are stated in (28).

(28) *Last Resort*:
If a feature δ on X that triggers an operation cannot be discharged in XP, there are two basic options:
a. δ is deleted.
b. δ is moved to the edge of XP, pied-piping the minimal category containing it.

Thus, a head X with a non-discharged δ ([•F•] or [⁎F⁎]) feature undergoes intermediate movement to phrase edges for as long as it takes to reach a position in

[11] One might think that allowing heads to move to specifier positions might give rise to various over-generation problems. However, it is worth bearing in mind that specifier positions can only ever be used as *intermediate* escape hatches (required by the PIC) by head movement under present assumptions; thus, the situation is completely analogous to, say, wh-movement via an intermediate Specv position in English – the wh-phrase can use this position as an intermediate escape hatch, but can never ultimately show up in it (for essentially the same reason, viz., that the trigger can only be saturated in the final landing site).

[12] This constraint can plausibly be derived as a theorem under various conceptions of cyclic spell-out of complements of phase heads.

which δ can eventually be discharged. Following Fanselow (2003; 2009), Surányi (2005), Matushansky (2006), and Georgi & Müller (2010), these kinds of features can then be viewed as triggers for reprojection movement.[13] Note that it can in principle be both probe features on some head X that trigger (intermediate or final) reprojection movement (e.g., if there is no matching goal for a probe in the structure, or if the goal is not c-commanded by the probe feature on X), and structure-building features (e.g., if there is no accessible matching category, or if two heads simultaneously need to discharge their [•F•] feature but only one can do this at any given stage of the derivation). However, in the reprojection approach to long-distance agreement to be developed in the next section, it is the need to discharge a structure-building feature that triggers the movement of (what thereby becomes) the matrix verb.

3.2 Long-distance agreement by reprojection

3.2.1 Complex predicates

Long-distance agreement typically encompasses verbs that in many languages are restructuring verbs. In fact, for another Nakh-Daghestanian language, Godoberi, Haspelmath (1999) shows with a series of tests that apparent long-distance agreement in the language actually involves only a monoclausal structure with a complex predicate. However, Forker (2012) and Khalilova (2009) show with similar tests that this is not the case for Hinuq or Khwarshi, both of which involve truly biclausal structures, with an embedded CP.

In view of this state of affairs, we would like to suggest that despite this biclausal character, long-distance agreement in Hinuq and Khwarshi (and Tsez, and perhaps more generally) does indeed involve some form of restructuring, albeit in the form of a special type of complex predicate formation. In standard lexical approaches to complex predicate formation (see, e.g., Haider 1993; 2010, Kiss 1995, Stiebels 1996, and S. Müller 2002 on German, or Butt 1995 on Hindi/Urdu), all lexical subcategorization information of the verbs that participate in the operation is unified by functional composition. This results in one featural array for the complex predicate and monoclausality throughout. Against the background of the present approach, this would imply a unique list of structure-building and probe features associated with the complex predicate, with the features discharged one after the other. In contrast, we adopt a version of pre-syntactic complex predicate formation where two predicates (two verbs, in the case at hand) are combined into a complex category in a way that, crucially, leaves the verbs'

[13] Fanselow (2003; 2009) and Georgi & Müller (2010) refer to these kinds of features as *Münchhausen* features, based on the literary character Baron Münchhausen who escapes from a swamp (where he is trapped on the back of his horse) by pulling himself up by his hair.

individual lexical information intact – i.e., there are still two separate lists of features triggering syntactic operations.[14]

3.2.2 Derivations

Let us now look at how long-distance agreement in Nakh-Dahestanian languages (and possibly elsewhere) can be derived on the basis of an approach in terms of reprojection and pre-syntactic complex predicate formation. Throughout, we will assume a CP status of the embedded clause, with both *λin*-type and *łi*-type markers qualifying as C heads. The definition of Agree that we will adopt is similar but not identical to the one in (7) from page 312; it is given in (29).[15]

(29) *Agree*:
α can Agree with β iff:
   a. α carries a probe feature [∗F∗], and β carries a maching goal feature [F].
   b. α c-commands β, or β c-commands α.
   c. There is no δ that is closer to β than α and also carries [∗F∗], and there is no γ that is closer to α than β and carries an active [F].
   d. β bears an active feature.

The syntactic derivation of a sentence such as (20) in Hinuq, where the matrix verb undergoes long-distance agreement with the embedded absolutive wh-phrase, starts with the complex predicate in (30); (20) is repeated here as (31) (with the default agreement option ignored).[16]

[14] For present purposes, it is immaterial whether this pre-syntactic component is conceived of as the lexicon, or as a pre-syntactic morphology domain; for concreteness, we will generally assume the former here.

[15] Since the PIC holds for all syntactic operations, the fact that Agree is also subject to this constraint does not have to be mentioned explicitly. As noted above, we will not address the question here of how exactly other cases of Agree that would at first sight seem to violate the PIC can be accounted for; but note that this issue is even more prominent (though not categorially different) in an approach where all phrases are phases.

The requirement in (29b) permits both upward and downward Agree; see Zeijlstra (2012) and Bjorkman & Zeijlstra (2019) vs. Preminger (2013). (The local Agree operation initiated by (what will become) the matrix verb in long-distance agreement will involve upward Agree.) (29c) ensures minimality, with closeness definable in terms of minimal path length. There is no defective intervention here: Discharged features on intervening heads and checked features on intervening phrases can be ignored. Finally, (29d) encodes the Activity Condition: An active feature is one that has not participated in Agree.

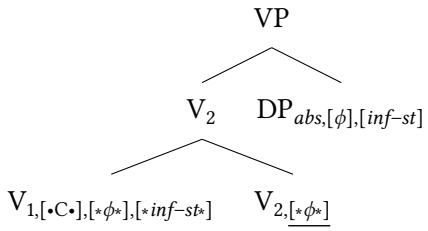
[16] One may ask why it is that V_2 (which will eventually become the embedded verb) projects in this structure, rather than V_1 (which will end up as the matrix verb). As a matter of fact,

(30) $[_{V_2}\ [_{V_1}$ know] $[_{V_2}$ kill]]

(31) Debez ∅-eq'i-ye $[_{CP(V)}$ k'ačaɣ-za-y ɬu ∅-uher-iš-ɬi] ?
you.SG.DAT I-know-Q bandit-OBL.PL-ERG who I-kill-RES-ABST
'Do you know whom the bandits killed?' (Forker 2012: 637)

In the first step, *$[_{V_2}[_{V_1}$know]-$[_{V_2}$kill]]* is merged with the internal argument DP, triggered by [•D•] on V_2. The resulting representation is shown in (32).

(32) *Long-distance agreement by reprojection, first stage*:



Each of the verbs involved has its own ϕ-probe (see Béjar & Řezáč 2009), which is checked through Agree with the ϕ-feature on the internal argument DP. Since V_2 is the head of the complex predicate, its ϕ-probe intervenes between V_1's ϕ-probe and the internal DP. Thus, V_2's ϕ-probe has to be discharged first; afterwards, V_1 can discharge its ϕ-probe via Agree with DP. This derives the generalization that long-distance agreement (i.e., under present assumptions, extremely local agreement of V_1 and the absolutive DP) is possible only if embedded agreement (i.e., agreement of V_2 and the absolutive DP) has taken place.[17]

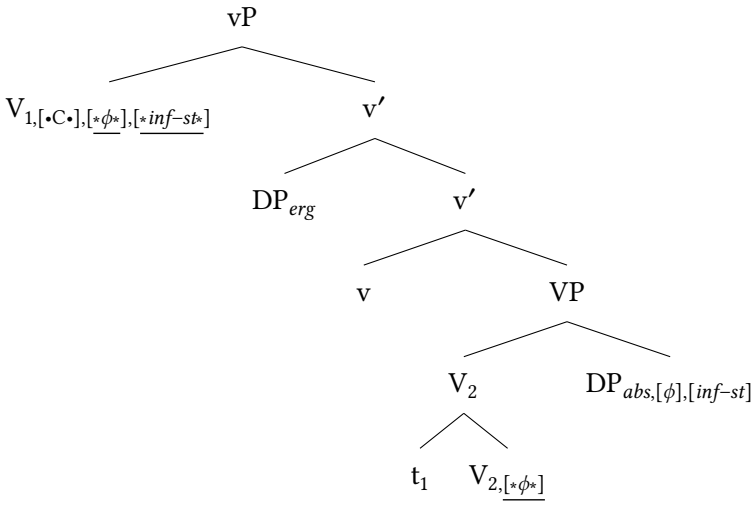
there does not seem to be a good reason why the alternative representation where V_1 projects – *$[_{V_1}[_{V_1}$know]-$[_{V_2}$kill]]* – should be excluded as such: With two bare X^0 heads forming a complex structure, labelling can be expected to be free. However, as will become clear when we look at the derivation for (31), choosing an initial representation *$[_{V_1}[_{V_1}$know]-$[_{V_2}$kill]]* of the complex predicate (rather than *$[_{V_2}[_{V_1}$know]-$[_{V_2}$kill]]*) can never lead to a well-formed derivation: V_2 ultimately needs to combine with a DP, but after (extended) projection of V_1 has been completed, only a CP is available for V_2, and this makes it impossible to discharge the [•D•] feature of V_2. In the same way, V_1 needs to combine with a CP (due to [•C•] on its feature list), but such a CP is not available at the beginning of the derivation.

[17] Given that only absolutive DPs can act as agreement controllers in the languages currently under consideration, the question arises whether this information is already locally available in the structure in (32). There are two possibilities, both of which strike us as viable. First, the absolutive (vs. lexically case-marked) nature of an internal DP argument might indeed already be visible at this stage (e.g., because V_2 does not have a lexical case feature). Second, if absolutive is not identifiable yet at this stage, agreement could simply take place in the hope that it will later emerge (e.g., be assigned by a functional head like T) – if it does not, the derivation will eventually crash.

In addition to the dependence of long-distance agreement on local agreement, a second generalization about long-distance can be derived at this point: There must be an obligatory information-structural reflex on the DP participating in long-distance agreement (with an interpretation as topic in Tsez, as topic or focus in Hinuq and Khwarshi, etc.); this is simply signalled by [inf-st] in (32) (i.e., [inf-st] stands for [topic], [topic, focus], or other information-structural features). Here is why: Given Chomsky's (2001) *Activity Condition*, after ϕ-Agree with V_2, DP_{abs} in (32) can only undergo ϕ-Agree with V_1 if it still has *a different, active* feature that V_1 is looking for; and [inf-st] fulfills this role. This explains the presence of [∗inf-st∗] on V_1 that needs to undergo Agree with [inf-st] on DP_{abs}. As a consequence of this second Agree operation involving V_1 and DP_{abs}, V_1 is equipped with the information that DP_{abs} is a topic.

In the further course of the derivation, V_2 first discharges all its structure-building features (if it has any such features left). Subsequently, v merges with VP; after that it merges with an external argument DP; and then it assigns ergative case to it. Up to this point, V_1 has not yet had a chance to discharge its [•C•] feature.[18] Therefore, before the vP is completed, V_1 needs to move to v's specifier position, so as to comply with Featural Cyclicity (cf. 27). The resulting representation is shown in (33).

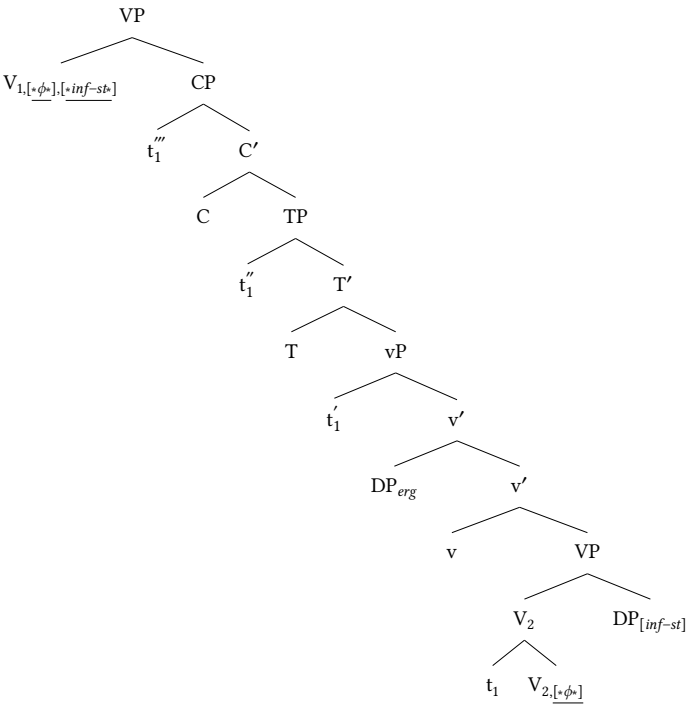
(33) *Long-distance agreement by reprojection, second stage*:



[18] This feature is either lower on the list of operation-triggering features of V_1 than the probe features for agreement with DP_{int}, or there are actually two separate stacks involved here (as indicated in 32): one for structure-building features, and one for probe features. This second option might be preferable on conceptual and empirical grounds; see Müller (2004; 2009) for discussion.

In further steps, the TP and CP structures of (what will become) the embedded clause are generated by Merge and Agree operations, while V_1 moves up the developing syntactic structure, via intervening phase edges. Finally, when the CP is completed, and V_1 has moved to C's edge domain because it still has not been able to discharge its structure-building feature [•C•], V_1 is in a position from which it can undergo reprojection movement, take the CP generated so far as its complement (thereby discharging [•C•]), and create a matrix VP.[19] This is shown in (34).[20]

(34) *Long-distance agreement by reprojection, third stage*:



The resulting representation is opaque in Kiparsky's (1973) sense as it involves a counter-bleeding interaction of operations (also cf. Chomsky 1951, Chomsky 1975: 25–26): Reprojection movement of V_1 would bleed Agree with DP_{abs} (which requires strict locality, due to the PIC) but fails to do so because it applies too

[19]In addition to subcategorization, V_1 carries out an Agree operation with C that reflects the embedding of an interrogative ([+wh]) clause.

[20]The mechanics here are similar to Martinović's (2015) analysis of the left periphery of Wolof in terms of head splitting and reprojection.

late: When V_1 has left the local domain in which agreement with DP_{abs} can legitimately be carried out, this agreement has already taken place.

From this point onwards, everything happens exactly as one would expect it to (with matrix vP, TP, and CP generated by Merge and Agree operations), and there is basically no difference anymore to derivations in which there is no complex predicate formation to begin with. Of course, given that pre-syntactic (lexical) complex predicate formation is an optional process, this second kind of derivation can be assumed to underlie minimally different sentences in which long-distance agreement does not occur. Thus, the two strategies differ substantially as far as earlier stages are concerned, but they end up with exactly the same structures once the matrix domain has been reached. There is one qualification, though. As a consequence of reprojection movement of V_1, [inf-st] of DP_{abs} is transported into the matrix clause.[21] The information that the embedded $DP_{[inf\text{-}st]}$ is interpreted as a topic is therefore shifted to the matrix sentence, and consequently, a DP that is affected by long-distance agreement is interpreted as the topic of the entire complex sentence. The analysis is thus consistent with usual assumptions concerning the impossibility of information-structural elements in clauses without illocutionary force (like non-assertive, presuppositional declarative clauses). Whereas information-structural features of an embedded DP can thus be interpreted in the matrix clause, there is no way how an embedded DP could take relative scope in the matrix clause as well (cf. the sentences in 9): Relative scope is determined by the position of an item, not by features, and there is no stage of the derivation where the embedded DP_{abs} would show up in the matrix clause.

3.2.3 Further consequences

The example of long-distance agreement that we have considered here on the basis of the sample derivation in (32), (33) and (34), involves a DP_{abs} controller that is also a wh-phrase. However, it should be clear that the approach generalizes to all the other cases of long-distance agreement mentioned above. For instance, an account in terms of reprojection of the part of a complex predicate straightforwardly derives long-distance agreement as in (2b) in Tsez and in (3b) in Hinuq

[21] Note that this implies that *discharged* features, while syntactically inert, are not actually deleted. This assumption must independently be made for discharged probe features more generally that give rise to morphological realization; see Adger (2003).

(where it can now be assumed that α stands for a full CP). Similarly, examples like (16b) in Khwarshi (with a DP_{abs} controller acting as a topic, i.e., [inf-st] representing [topic]) and (17b) in Hinuq (with a DP_{abs} controller acting as a focus, i.e., [inf-st] representing [focus]) are directly accounted for under the present analysis. Examples (18) (from Khwarshi) and (19) (from Hinuq) have subject wh-phrases (one marked by ergative, one not) that do not block long-distance agreement with the absolutive DP. Again, this is expected under present assumptions: Independently of whether the wh-phrase here occupies SpecC in overt syntax or not, reprojection movement of the verb to the matrix domain is possible (given the general option of multiple specifiers, particularly for intermediate movement steps).

Next, instances of of super-long-distance agreement where the agreeing verb and the agreement controller DP are separated by two intervening CP boundaries, like (21) in Tsakhur or (22) in Hinuq, can also be addressed under the reprojection approach: Here a complex predicate is formed pre-syntactically where V_1 (which will become the highest verb) and V_2 (which will become the intermediate verb) are first combined, with V_2 projecting (in a successful derivation; cf. footnote 16), and then the complex V_2 category is combined with V_3 (which will become the most deeply embedded verb), with V_3 projecting, as shown in (35).

(35) $[_{V_3}\ [_{V_2}\ V_1\ V_2]\ V_3]$

Here V_3, V_2, and V_1 first carry out Agree operations with V_3's internal argument (DP_{abs}), and then a CP is generated on top of VP_3, with the complex V_2 moving successive-cyclically to intermediate phase edge positions, until it finally merges with the CP. Then, the second, intermediate, CP is generated, with V_1 excorporating from the complex $[_{V_2}\ V_1\ V_2]$ category and moving via the intermediate CP's phases edges until, finally, the intermediate CP has been completed and V_1 can take this CP as its internal argument, via reprojection. The Tsakhur example in (21) and the Hinuq example in (22) fully correspond to this scenario, with V_1, V_2, and V_3 all participating in agreement with DP_{abs}. However, examples involving super-long-distance agreement like the one in (36) (from Hinuq) can also be found.

(36) Dіž y-eqʼi-yo [$_{CP}$ ʔumar-i [$_{CP}$ Madina y-aqʼ-es=ƛen]
I.DAT II-know-PRS Umar-ERG Madina(II) II-come-PST=QUOT
ese-s-ɬi]
tell-RES-ABST
'I know that Omar said that Madina came.' (Forker 2012: 633)

In (36), the intermediate verb V_2 does in fact not exhibit overt agreement marking even though both the matrix verb V_1 and the most deeply embedded verb V_3 do. Still, (36) does not call into question the present approach: It can plausibly be assumed that ϕ-feature agreement is indeed present on V_2, but fails to be registered overtly (there are many verbs that fail to exhibit visible agreement marking despite showing up in the proper syntactic context in Nakh-Daghestanian languages, and the reason for this is presumably simply a morphological one). Thus, all in all, super-long-distance agreement can be derived.[22]

A further property of long-distance agreement that needs to be accounted for concerns Polinsky & Potsdam's (2001) observation that the C element *λin* blocks the operation in Tsez (cf. Section 2.4 above). Given that there is good evidence that long-distance agreement across CP is possible in principle in Nakh-Daghestanian languages, and given that we have analyzed the transparent morpheme *łi* as a C item, too, a recourse to a general blocking nature of C is not available in the present approach. Also, it is not possible to claim that a reprojecting V_1 cannot merge with a CP headed by *λin*: First, the [•C•] feature responsible for reprojection movement is not sensitive to a difference between C heads, and an additional selection relation (mediated by Agree) would have to be stipulated; second (and more importantly), [•C•] on a reprojecting V is exactly the same feature as [•C•] on a V that fails to undergo complex predicate formation, and successfully takes CP complements headed by *λin* in environments without long-distance agreement. In view of this, we would like to suggest that the blocking effect of a C head *λin* is due to the fact that it does not permit a specifier. Thus, the problem with long-distance agreement in these contexts can be traced back to the unavailability of intermediate movement of a V_1 that is initially part of

[22]It should be mentioned that there are two further complications, though. First, recall that the present account of the obligatory information-structural reflex of long-distance agreement in terms of the Activity Condition would, strictly speaking, require two different additional features (next to the ϕ-probes) on V_1 and V_2, and not just one, as in the cases discussed so far. It is not a priori clear what this extra feature might be. However, it has been argued that information-structural features like topic and focus do not qualify as primitives, but are rather composed of more primitive binary features (so as to capture natural classes of information-structural categories), like [±new], [±prom] (with, say, topic emerging as [–new,+prom]); see Choi (1999), based on Vallduví (1992). If so, V_2 and V_1 can be equipped with separate pieces of [inf-st] information.

Second, Forker (2012) also maintains that it is not completely impossible in Hinuq to have super-long-distance agreement involving V_1, V_3, and D_{abs} in the most deeply embedded clause, not merely in the absence of agreement on V_2 (as in 36), but in the presence of a *different* agreement on V_2. If such sentences (which Forker assigns an intermediate status, signalled by "?") can be substantiated as grammatically well formed, additional assumptions that complement the present analysis will be called for.

a complex predicate, to SpecC: As a consequence, the final reprojection step of V_1 will have to fatally violate the PIC (V_1 can only reach SpecT, which is not accessible anymore once CP has been completed).

Finally, we would like to point out that the present approach in terms of reprojection makes a very simple prediction: Reprojection movement of a verb by definition creates a head-complement structure; there is no way how a specifier or adjunct could be involved (since this would require a non-X^0 category to move). Therefore, long-distance agreement is expected never to occur into subject clauses or adjunct clauses. This prediction is borne out: Long-distance agreement always involves complement clauses.

4 Conclusion

We have argued that from the point of view of a model of syntax where all operations apply in strictly local domains (as defined by the Phase Impenetrability Condition, PIC), and in the face of empirical evidence showing that long-distance agreement can involve a matrix verb and an agreement-controlling DP separated by a CP, none of the existing approaches to long-distance agreement (non-local analyses, small structure analyses, Cyclic Agree analyses, and analyses where movement to the edge feeds agreement) work satisfactorily. In view of this, we have developed a new approach in terms of pre-syntactic complex predicate formation and reprojection: The derivation starts out with a complex verb V_1-V_2 headed by V_2, so that agreement of V_1 with DP can apply early in the derivation (not late, as in other approaches), in an extremely local domain, and subsequent reprojection movement of V_1 turns the latter into a matrix verb, thereby masking the locality of agreement and creating opacity (viz., counter-bleeding) in syntax.

This approach may at first sight look quite radical. However, it is worth bearing in mind that it suggests itself without further ado once two widely employed operations are adopted and combined, viz., (i) pre-syntactic complex predicate formation, and (ii) head movement as reprojection. The properties that these two operations must have for the analysis to work all qualify as independently motivated, and they often correspond to standard assumptions in the field (in analyses that adopt the operations). As a matter of fact, the only innovative assumption that we have come up with is that pre-syntactic complex predicate formation does not (or does not have to) result in a single list of structure-building and agreement-inducing features (via a process of functional composition), but can maintain the integrity and independence of the two individual lists of structure-building and agreement-inducing features.

Nevertheless, ideally there should be independent evidence for the type of interaction of complex predicate formation and reprojection movement that is at the heart of the present analysis of long-distance agreement. To end this paper, we would like to briefly sketch an approach to an entirely different phenomenon that works in the same way, viz., extraction from DPs in German.
As for the empirical evidence, extraction is impossible from subject DPs and indirect object (dative-marked) DPs. This is shown (with wh-movement as the extraction operation) in (37a) and (37b), respectively.

(37) a. * [$_{PP}$ Über wen]$_1$ hat [$_{DP}$ ein Buch t_1] den Karl
about whom has a book.NOM the Karl.ACC
beeindruckt?
impressed
'For which person is it the case that a book about that person impressed Karl?'

b. * [$_{PP}$ Über wen] hat sie [$_{DP}$ einem Buch t_1] keine
about whom has she.NOM a book.DAT no
Chance gegeben?
chance.ACC given
'For which person is it the case that she gave a book about that person no chance?'

With extraction from direct object (accusative-marked) DPs, things are somewhat more variable: With some combinations of V and N, extraction is possible (see 38a), with other combinations, it is not (see 38b).

(38) a. [$_{PP}$ Über wen]$_1$ hat Karl [$_{DP}$ ein Buch t_1] gelesen ?
about whom has Karl.NOM a book.ACC read
'Who did Karl read a book about?'

b. * [$_{PP}$ Über wen]$_1$ hat Karl [$_{DP}$ ein Buch t_1] geklaut ?
about whom has Karl.NOM a book.ACC stolen
'For which person is it the case that Karl stole a book about that person?'

Thus, both structural and lexical factors play a role: On the one hand, extraction from DP can be well formed in German if DP is a complement (as with direct objects in (38a) and (38b)), but not if it is a specifier (as with subjects and indirect objects in (37a) and (37b), which can be assumed to occupy Specv and SpecAppl positions, respectively). On the other hand, extraction from DP also

requires V and N to form a tight unit, or a "natural predicate". This latter status is arguably determined both by semantic considerations and by extralinguistic factors (frequency, entrenchment), and it may to some extent vary from speaker to speaker. Still, it must be modelled in the grammar in some way. In Müller (1991) and Müller & Sternefeld (1995), it is proposed that the relevant concept is that of abstract incorporation (in Baker 1988's sense, conceived of as incorporation at LF that is signalled already by co-indexation of heads in overt syntax): V (*read*) and N (*book*) in (38a) undergo abstract incorporation and thus form a natural predicate, whereas V (*steal*) and N (*book*) in (38b) do not (for most speakers). Given that the theory of locality constraints on movement is sensitive to this difference (as well as to the structural difference between complements and specifiers), the data in (37) and (38) can then all be accounted for. Similar approaches in terms of abstract incorporation have subsequently been developed by Davies & Dubinsky (2003) and Schmellentin (2006). However, there are problems with this kind of approach. In particular, on this view abstract incorporation of N into V must either be able to apply non-locally, across an intervening DP projection (plus, possibly, other functional projections in the DP that may intervene between D and N); or the analysis must abandon the DP-over-NP hypothesis. To be sure, there are ways out for the abstract incorporation approach.[23] Still, it can be noted that an approach based on complex predicate formation plus reprojection can account for the data in a very simple way.

Thus, suppose that some combinations of V and N can undergo pre-syntactic (lexical) complex predicate formation whereas others cannot do so. This means that complex heads like the one in (39) can be primitive inputs of Merge operations in the syntax.[24]

(39) $[_{N_2}\ V_1\ N_2]$

In the ensuing derivation, N_2 first discharges its structure-building and probe features (thereby undergoing Merge with a PP); see (40a). Then DP is added on top of NP (and possibly also other functional projections before that), with V_1 undergoing intermediate, Last Resort-driven movement to SpecD; see (40b). After that, V_1 undergoes the final movement step to take DP as its complement and thereby discharge its [•D•] feature; see (40c). From this point onwards, everything proceeds exactly as in a derivation where there is no complex predicate formation as in (40); such a derivation produces the VP in (41).

[23]For instance, in Müller (2011), abstract incorporation is viewed as a regular syntactic Agree operation, with no actual movement involved.

[24]As before, the head that will ultimately come to occupy a higher (c-commanding) position must be the one that fails to project initially in a well-formed derivation; cf. footnote 16.

(40) a. $[_{NP}\ [_{N_2}\ V_1\ N_2]\ PP]$
b. $[_{DP}\ V_1\ [_{D'}\ D\ [_{NP}\ [_{N_2}\ t_1\ N_2]\ PP]]]$
c. $[_{VP}\ V_1\ [_{DP}\ t'_1\ [_{D'}\ D\ [_{NP}\ [_{N_2}\ t_1\ N_2]\ PP\ ]]]]$

(41) $[_{VP}\ V_1\ [_{DP}\ D\ [_{NP}\ N_2\ PP]]]$

Importantly, both the NP and the PP in (40c) (based on complex predicate formation of V and N) have been in an extremely local (object-like) relation with V whereas the NP and PP in (41) (based on regular, separate projection of V and N) have never been in a local relation with V. Without going into the details of how exactly this will best be implemented in a given theory of locality restrictions on movement, it seems plausible to assume that it is the extremely local relation of V and NP/PP at an earlier derivational step that makes extraction of PP from DP possible in (38a), and it is the absence of such a relation that blocks the movement in (38b). As before, reprojection movement of V by definition cannot take place from specifiers, which then accounts for the illformedness of (37a) and (37b).

Abbreviations

ABS	absolutive	NOM	nominative
ABST	abstract suffix	NMLZ	nominalizer
ACC	accusative	OBJ	object
CAUS	causative	OBL	oblique
CL	clitic	PERF	perfective
COP	copula	PL	plural
CVB	narrative converb	PR(E)S	present
DAT	dative	PST	past
ERG	ergative	PSTPRT	past participle
FEM	feminine	PTCP	participle
FOC	focus	Q	question particle
INF	infinitive	QUOT	quotative enclitic
INF-ST	information structure	RES	resultative participle
INTRANS	intransitive	SG	singular
LAT	lative	SUBJ	subject
MASC	masculine	TI	transitive inanimate

Acknowledgments

For comments and discussion, we are grateful to Josef Bayer, Elena Pyatigorskaya, Tonjes Veenstra, Martina Martinović, Johannes Mursell (and the other editors of the present volume), and audiences at the *GGS* meeting at Universität Konstanz (2014), the workshop on *Agreement* at the University of York (2014), and the workshop *Heads, Phrases, Ts and Nodes* at Humboldt-Universität zu Berlin (2017).

References

Ackema, Peter, Ad Neeleman & Fred Weerman. 1993. Deriving functional projections. In Amy J. Schafer (ed.), *Proceedings of NELS 23*, 17–31. Amherst: GSLA.

Adger, David. 2003. *Core syntax*. Oxford: Oxford University Press.

Alexiadou, Artemis, Tibor Kiss & Gereon Müller. 2012. Local modelling of non-local dependencies in syntax: An introduction. In Artemis Alexiadou, Tibor Kiss & Gereon Müller (eds.), *Local modelling of non-local dependencies in syntax*, 1–49. Berlin: De Gruyter.

Baker, Mark C. 1988. *Incorporation: A theory of grammatical function changing*. Chicago: University of Chicago Press.

Barss, Andrew. 1986. *Chains and anaphoric dependence*. MIT. (Doctoral dissertation).

Bayer, Josef & Ellen Brandner. 2008. On wh-head-movement and the Doubly-Filled-Comp Filter. In Charles B. Chang & Hannah J. Haynie (eds.), *Proceedings of the 26th West Coast Conference on Formal Linguistics*, 87–95. Somerville, MA: Cascadilla Proceedings Project.

Béjar, Susana & Milan Řezáč. 2009. Cyclic Agree. *Linguistic Inquiry* 40(1). 35–73.

Bhatt, Rajesh. 2002. The raising analysis of relative clauses. *Natural Language Semantics* 10(1). 43–90.

Bhatt, Rajesh. 2005. Long distance agreement in Hindi-Urdu. *Natural Language & Linguistic Theory* 23(4). 757–807.

Bhatt, Rajesh & Stefan Keine. 2016. Long distance agreement. In Martin Everaert & Henk van Riemsdijk (eds.), *The Blackwell companion to syntax*, 2nd edn., 2291–2321. Hoboken, NJ: Blackwell-Wiley.

Bjorkman, Bronwyn & Hedde Zeijlstra. 2019. Checking up on (ϕ-)Agree. *Linguistic Inquiry* 50(3). 527–569.

Bliss, Heather. 2009. Structuring information in Blackfoot: Against an A-bar-agreement analysis of cross-clausal agreement. In Susie Jones (ed.), *Proceedings of the 2008 Canadian Linguistics Association annual conference*.

Bobaljik, Jonathan D. & Susi Wurmbrand. 2005. The domain of agreement. *Natural Language & Linguistic Theory* 23(4). 809–865.
Boeckx, Cedric. 2004. Long-distance agreement in Hindi: Some theoretical implications. *Studia Linguistica* 58(1). 23–36.
Bošković, Željko. 2007. Agree, phases, and intervention effects. *Linguistic Analysis* 33(1–2). 54–96.
Branigan, Phil & Marguerite MacKenzie. 2002. Altruism, A-bar movement, and object agreement in Innu-aimûn. *Linguistic Inquiry* 33(3). 385–407.
Bruening, Benjamin. 2001. *Syntax at the edge: Cross-clausal phenomena and the syntax of Passamaquoddy*. MIT. (Doctoral dissertation).
Bury, Dirk. 2003. *Phrase structure and derived heads*. University College London. (Doctoral dissertation).
Butt, Miriam. 1995. *The structure of complex predicates in Urdu*. Standford, CA: CSLI Publications.
Butt, Miriam. 2008. Revisiting long-distance agreement in Urdu. Ms., Universität Konstanz.
Chandra, Pritha. 2005. Hindi-Urdu long distance agreement: Agree, Agree or Spec-Head? Ms, University of Maryland.
Choi, Hye-Won. 1999. *Optimizing structure in context: Scrambling and information structure*. Stanford: CSLI Publications.
Chomsky, Noam. 1951. *Morphophonemics of Modern Hebrew*. University of Pennsylvania. (MA thesis).
Chomsky, Noam. 1975. *The logical structure of linguistic theory*. New York: Plenum Press.
Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press.
Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by step: Essays on Minimalist syntax in honor of Howard Lasnik*, 89–155. Cambridge, MA: MIT Press.
Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 1–52. Cambridge, MA: MIT Press.
Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos Otero & Maria Luisa Zubizarreta (eds.), *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, 133–166. Cambridge, MA: MIT Press.
Chomsky, Noam. 2013. Problems of projection. *Lingua* 130. 33–49.
Cinque, Guglielmo. 1990. *Types of A-bar dependencies*. Cambridge, MA: MIT Press.
Davies, William & Stanley Dubinsky. 2003. On extraction from NPs. *Natural Language & Linguistic Theory* 21(1). 1–37.

de Cuba, Carlos & Barbara Urogdi. 2010. Clearing up the "facts" on complementation. *University of Pennsylvania Working Papers in Linguistics* 16(1). 41–50.
Donati, Caterina. 2006. On Wh-head movement. In Lisa Cheng & Norbert Corver (eds.), *Wh-movement: Moving on*, 21–46. Cambridge, MA: MIT Press.
Ebert, Christian, Cornelia Endriss & Stefan Hinterwimmer. 2008. Intermediate scope readings as embedded speech acts. In Atle Grønn (ed.), *Proceedings of SuB 12*, 122–136. Oslo: ILOS.
Fanselow, Gisbert. 2003. Münchhausen-style head movement and the analysis of verb-second. In Anoop Mahajan (ed.), *Head movement and syntactic theory*, 40–76. Los Angeles & Potsdam: UCLA & Universität Potsdam Working Papers in Linguistics.
Fanselow, Gisbert. 2009. Bootstrapping verb movement and the clausal architecture of German. In Artemis Alexiadou, Jorge Hankamer, Thomas McFadden, Justin Nuger & Florian Schäfer (eds.), *Advances in comparative Germanic syntax*, 85–118. Amsterdam: Benjamins.
Forker, Diana. 2012. *A grammar of Hinuq*. Berlin: de Gruyter.
Gagliardi, Annie, Michael Goncalves, Maria Polinsky & Nina Radkevich. 2014. The biabsolutive construction in Lak and Tsez. *Lingua* 150. 137–170.
Gazdar, Gerald. 1981. Unbounded dependencies and coordinate structure. *Linguistic Inquiry* 12(2). 155–184.
Georgi, Doreen. 2014. *Opaque interactions of Merge and Agree*. Universität Leipzig. (Doctoral dissertation).
Georgi, Doreen & Gereon Müller. 2010. Noun phrase structure by reprojection. *Syntax* 13(1). 1–36.
Haider, Hubert. 1993. *Deutsche Syntax – generativ*. Tübingen: Narr.
Haider, Hubert. 2000. Branching and discharge. In Peter Coopmans, Martin Everaert & Jane Grimshaw (eds.), *Lexical specification and insertion*, 135–164. Amsterdam: John Benjamins.
Haider, Hubert. 2010. *The syntax of German*. Cambridge: Cambridge University Press.
Haspelmath, Martin. 1999. Long distance agreement in Godoberi (Daghestanian) complement clauses. *Folia Linguistica* 33(1–2). 131–151.
Holmberg, Anders. 1991. Head scrambling. Handout of talk at GLOW 1991, Leiden.
Hooper, Joan & Sandra Thompson. 1973. On the applicability of root transformations. *Linguistic Inquiry* 4(4). 465–497.
Hornstein, Norbert. 2001. *Move: A Minimalist theory of construal*. Oxford: Blackwell.

Hornstein, Norbert. 2009. *A theory of syntax: Minimal operations and universal grammar*. Cambridge: Cambridge University Press.
Hornstein, Norbert & Juan Uriagereka. 2002. Reprojections. In Samuel D. Epstein & Daniel Seely (eds.), *Derivation and explanation in the Minimalist Program*, 106–132. London: Blackwell.
Keine, Stefan. 2008. Long-distance Agreement und zyklisches Agree. Ms, Universität Leipzig.
Keine, Stefan. 2016. *Probes and their horizons*. University of Massachusetts at Amherst. (Doctoral dissertation).
Khalilova, Zaira. 2007. Long-distance agreement in Khwarshi. In Miltiadis Kokkonidis (ed.), *Proceedings of LingO 2007*, 116–124. Oxford: Faculty of Linguistics, Philology, & Phonetics, University of Oxford.
Khalilova, Zaira. 2009. *A grammar of Khwarshi*. Utrecht: LOT.
Kibrik, Aleksandr. 1999. *Elementy grammatiki caxurskogo jazyka v tipologičeskom osveščenii*. Moskva: Nasledie.
Kiparsky, Paul. 1973. Abstractness, opacity and global rules. In Osamu Fujimura (ed.), *Three dimensions in linguistic theory*, 57–86. Tokyo: TEC.
Kiss, Tibor. 1995. *Infinite Komplementation*. Tübingen: Niemeyer.
Koeneman, Olaf. 2000. *The flexible nature of verb movement*. Utrecht University. (Doctoral dissertation).
Koopman, Hilda & Anna Szabolcsi. 2000. *Verbal complexes*. Cambridge, MA: MIT Press.
Krifka, Manfred. 1992. A compositional semantics for multiple focus constructions. In Joachim Jacobs (ed.), *Informationsstruktur und Grammatik*, 17–53. Opladen: Westdt. Verlag.
Krifka, Manfred. 2004. Semantics above and below speech acts. Handout of talk at Stanford University.
Krifka, Manfred. 2014. Embedding speech acts. Ms., Humboldt-Universität zu Berlin.
Lahne, Antje. 2012. Specificity-driven syntactic derivation. In Myriam Uribe-Etxebarria & Vidal Valmala (eds.), *Ways of structure building*, 271–296. Oxford: Oxford University Press.
Legate, Julie A. 2005. Phases and cyclic agreement. In Martha McGinnis & Norvin Richards (eds.), *Perspectives on phases*, 147–156. Cambridge, MA: MIT Working Papers in Linguistics.
Mahajan, Anoop. 1990. *The A/A-bar distinction and movement theory*. MIT. (Doctoral dissertation).
Mahajan, Anoop. 2001. Word order and remnant VP movement. Ms, UCLA.

Martinović, Martina. 2015. *Feature geometry and head-splitting: Evidence from the Wolof clausal periphery*. University of Chicago. (Doctoral dissertation).

Matić, Dejan, Rik van Gijn & Robert D. Van Valin Jr. 2014. Information structure and reference tracking in complex sentences: An overview. In Rik van Gijn, Jeremy Hammond, Dejan Matić, Saksia van Putten & Ana Vilacy Galucio (eds.), *Information structure and reference tracking in complex sentences*. Amsterdam: John Benjamins.

Matushansky, Ora. 2006. Head-Movement in linguistic theory. *Linguistic Inquiry* 37(1). 69–109.

Müller, Gereon. 1991. Abstrakte Inkorporation. In Susan Olsen & Gisbert Fanselow (eds.), *DET, COMP und INFL*, 155–202. Tübingen: Niemeyer.

Müller, Gereon. 2004. Argument encoding and the order of elementary operations. Ms., Universität Leipzig.

Müller, Gereon. 2009. Ergativity, accusativity, and the order of Merge and Agree. In Kleanthes K. Grohmann (ed.), *Explorations of phase theory: Features and arguments*, 269–308. Berlin: Mouton de Gruyter.

Müller, Gereon. 2011. *Constraints on displacement: A phase-based approach*. Amsterdam: John Benjamins.

Müller, Gereon. 2014. *Syntactic buffers*. Universität Leipzig: Linguistische Arbeitsberichte.

Müller, Gereon & Wolfgang Sternefeld. 1995. Extraction, lexical variation, and the theory of barriers. In Urs Egli, Peter E. Pause, Christoph Schwarze, Arnim von Stechow & Götz Wienold (eds.), *Lexical knowledge in the organization of language*, 35–80. Amsterdam: John Benjamins.

Müller, Stefan. 2002. *Complex predicates*. Stanford, CA: CSLI Publications.

Müller, Stefan. 2016. *German clause structure: An analysis with special consideration of so-called multiple frontings*. Berlin: Language Science Press.

Nilsen, Øystein. 2003. *Eliminating positions*. Universiteit Utrecht. (Doctoral dissertation).

Pesetsky, David. 1985. Morphology and Logical Form. *Linguistic Inquiry* 16(2). 193–246.

Polinsky, Maria. 2016. Agreement in Archi from a Minimalist perspective. In Oliver Bond, Greville Corbett, Marina Chumakina & Dunstan Brown (eds.), *Archi: Complexities of agreement in cross-theoretical perspective*, 184–232. Oxford: Oxford University Press.

Polinsky, Maria & Eric Potsdam. 2001. Long-distance agreement and topic in Tsez. *Natural Language & Linguistic Theory* 19(3). 583–646.

Preminger, Omer. 2009. Breaking agreements: Distinguishing agreement and clitic doubling by their failures. *Linguistic Inquiry* 40(4). 619–666.

Preminger, Omer. 2013. That's not how you agree: A reply to Zeijlstra. *The Linguistic Review* 30(3). 491–500. DOI:10.1515/tlr-2013-0015
Preminger, Omer. 2014. *Agreement and its failures*. Cambridge, MA: MIT Press.
Reuland, Eric. 2011. *Anaphora and language design*. Cambridge, MA: MIT Press.
Rizzi, Luigi. 1990. *Relativized minimality*. Cambridge, MA: MIT Press.
Roberts, Ian. 1991. Excorporation and minimality. *Linguistic Inquiry* 22(1). 209–218.
Roberts, Ian. 1997. Restructuring, head movement, and locality. *Linguistic Inquiry* 28(3). 423–460.
Roberts, Ian. 2009. On the probable existence of head movement and the definite non-existence of the Head Movement Constraint. Ms, University of Cambridge.
Roberts, Ian. 2010. *Agreement and head movement: Clitics, incorporation, and defective goals*. Cambridge, MA: MIT Press.
Schmellentin, Claudia. 2006. *PP-Extraktionen: Eine Untersuchung zum Verhältnis von Grammatik und Pragmatik*. Tübingen: Niemeyer.
Sells, Peter. 2006. Using subsumption rather than equality in functional control. In Miriam Butt & Tracy King (eds.), *Proceedings of LFG-06*. Universität Konstanz: CSLI Publications.
Stenius, Erik G. 1967. Mood and language game. *Synthese* 17(1). 254–274.
Sternefeld, Wolfgang. 1989. V-movement, extraction from V/2 clauses, and the ECP. *Working Papers in Scandinavian Syntax* 44. 119–140.
Stiebels, Barbara. 1996. *Lexikalische Argumente und Adjunkte: Zum semantischen Beitrag von verbalen Präfixen und Partikeln*. Berlin: Akademie Verlag.
Surányi, Balázs. 2005. Head movement and reprojection. *Annales Universitatis Scientiarum Budapestinensis de Rolando Eötvös Nominatae. Sectio Linguistica. ELTE* 26. 313–342.
Travis, Lisa deMena. 1984. *Parameters and effects of word order variation*. MIT. (Doctoral dissertation).
Vallduví, Enric. 1992. *The informational component*. New York: Garland.
von Stechow, Arnim. 1996. Against LF pied-piping. *Natural Language Semantics* 4(1). 57–110.
von Stechow, Arnim & Wolfgang Sternefeld. 1988. *Bausteine syntaktischen Wissens*. Opladen: Westdeutscher Verlag.
Zeijlstra, Hedde. 2012. There is only one way to agree. *The Linguistic Review* 29(3). 491–539.