

Chapter 9

Long distance agreement and information structure

Johannes Mursell

Goethe-Universität Frankfurt am Main

In this paper, I discuss a specific subtype of long distance agreement (LDA), namely agreement across a finite (CP) clause boundary. I show that most, if not all, instances of LDA depend on information-structural properties of the agreement target in the embedded clause, be it for cross clausal object agreement or case assignment. After discussing LDA in unrelated language families, I argue that LDA is possible due to a complex probe in the left periphery containing information-structural features bundled with ϕ -features. This probe serves as an intermediate agreement step between a verb in the matrix clause and the information-structurally marked element in the embedded clause, so that LDA does not violate the phase impenetrability condition.

1 Introduction

Long distance agreement (LDA) in general refers to a syntactic dependency by which certain features, usually ϕ -features, of a probing head depend on features of a non-local constituent, i.e. a constituent not in the specifier of the probing head (Bhatt & Keine 2016). While long distance agreement provides a strong argument for the operation of AGREE as formulated in Chomsky (2000; 2001), and further refined in Pesetsky & Torrego (2007), which does not rely on local specifier-head agreement, several sub-cases of long distance agreement need to be distinguished.

- (1) AGREE (Pesetsky & Torrego 2007: 268)
 - a. An unvalued feature F (a probe) on a head H at syntactic location α (F_α) scans its c-command domain for another instance of F (a goal) at location β (F_β) with which to agree.



- b. Replace F_α with F_β , so that the same feature is present in both locations.

The three cases of LDA that need to be distinguished differ from each other with respect to the distance between probe and goal. Those two elements can either be part of the same clause, separated by a non-finite clause boundary or by a finite one. The first case of LDA can be exemplified by quirky subjects in Icelandic or object agreement in Zulu. The example in (2) from Zaenen et al. (1985) shows that in sentences with non-nominative subjects, a nominative object controls verbal agreement. Importantly, this nominative object is most likely not in the specifier of the projection that hosts *voru* and thus, qualifies as a case of LDA.

- (2) Icelandic (Zaenen et al. 1985: 460)
Konunginum voru gefnar ambáttir.
the.king.DAT were given.F.PL maidservants.NOM.F.PL
'The king was given female slaves.'

The second type, in which probing head and agreement target are separated by a non-finite clause boundary, can be found in English. In raising constructions, an expletive can be inserted in matrix subject position instead of raising the embedded subject. However, the embedded subject still controls agreement on the matrix verb.¹

- (3) a. Two men seem to be in the garden.
b. There seem to be two men in the garden.
c. * There seems to be two men in the garden.

This type of cross-clausal long distance agreement, i.e. agreement into a non-finite clause, is present in several languages and has frequently been discussed in the literature. Thus, this type of agreement can be found in, among others, Hindi-Urdu (Bhatt 2005, 4a), Godoberi (Haspelmath 1999, 4b), or Basque (Preminger 2009).

- (4) a. Hindi-Urdu (Bhatt 2005: 760)
Vivek-ne [**kitaab** pərh-nii] chaah-ii.
Vivek-ERG **book.F** read-INF.F want-PFV.F.SG
'Vivek wanted to read the book.'
b. Godoberi (Haspelmath 1999: 136)
[wašu-di **quči-be** r-al-u] r-uL-i.
boy-ERG **book-PL.ABS** PL.N-read-CVB.PST **PL.N-finish-AOR**
'The boy finished reading the books.'

¹Two reviewers point out that examples like (3c) are possible in certain registers of English.

If long distance agreement into a non-finite clause is analysed as being based on a restructuring configuration (Wurmbrand 2001), i.e. a configuration consisting of a full-fledged matrix clause and a truncated, smaller embedded clause, the two types of long distance agreement discussed so far do not challenge the locality of the agreement process in (1), since neither violates the strong version of the Phase-Impenetrability Condition in (5).

(5) *Phase-Impenetrability Condition (PIC)* (Chomsky 2001)

In a 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 phases are assumed to be at least CP and vP and if restructuring involves embedding clause smaller than CP, then agreement into a clause smaller than a CP does not violate the PIC.

The last and most surprising possibility of long distance agreement concerns instances in which probe and goal of the agreement process are separated by a finite CP boundary, and thus constitutes a clear violation of the PIC.² Examples for this type of LDA are rather rare but can be found in at least three different language families, namely in some of the Nakh-Dagestanian languages spoken in the north-eastern Caucasus region, in certain Algonquian languages spoken in North America, and in at least one Altaic language.

In those languages it is possible that, in certain circumstances, an argument of an embedded finite CP determines agreement on the matrix verb. In this paper, I will be concerned with the conditions in which this kind of long distance agreement can take place. I will argue that the crucial factor for LDA is that the embedded agreement goal is information-structurally marked, either as topic or as focus. This will also enable an analysis of the phenomenon compatible with PIC by agreement through the edge of the CP based on information-structural features, analyzing LDA as successive cyclic agreement comparable to successive cyclic wh-movement. In the discussion of LDA, an interesting generalization will emerge, namely that languages that allow LDA, allow it either for topics alone or for topics and foci, not however for foci alone.

The paper is structured as follows: I will first introduce the relevant phenomenon of LDA in more detail in Section 2 before I turn to previous analysis and their respective problems in Section 3. In Section 4, I develop my analysis and Section 5 concludes the paper.

²This presupposes that AGREE is subject to the same restrictions as movement. This is not undebated, see Bošković (2007a,b).

2 LDA crosslinguistically

Long distance agreement of the kind I am interested in in this paper is typologically much rarer than the other two types, but can be found in at least three different and unrelated language families. In this section, I am going to present the relevant data, pointing out the specific properties of the construction in the various languages based on the available literature, the properties any theory of LDA should be able to account for. First, the most well-known example, Tsez, will be presented, together with data from two other Nakh-Dagestian languages, Hinuq and Khwarshi. Second, I will present data from the Algonquian languages Blackfoot, Innu-aimûn, and Passamaquoddy. In the last subsection, I will present data from an Altaic language, Uyghur, which displays a slightly different type of LDA, not based on ϕ -features but mostly on case.

As will become clear in this chapter, all languages share one important property: long distance agreement always takes place between an element of a higher phase and an information-structurally marked element of the lower phase. This information-structural marking will be at the heart of the analysis developed in Section 4.

2.1 Nakh-Dagestian languages

In their seminal paper, Polinsky & Potsdam (2001) discuss LDA in the Nakh-Dagestian language Tsez. The basic paradigm is given in (6).

(6) Tsez (Polinsky & Potsdam 2001: 584)

- a. Eni-r [uʒā **magalu** b-āc'ruʒi] **b-iy-xo**.
mother-DAT [boy **bread.III.ABS** ate] III-know-PRS
'The mother knows that, as for the bread, the boy ate it.'
- b. Eni-r [uʒā **magalu** b-āc'ruʒi] **r-iy-xo**.
mother-DAT [boy **bread.III.ABS** ate] IV-know-PRS
'The mother knows that the boy ate the bread.'

In this ergative, verb-final language, only absolutive arguments determine agreement. As can be seen in (6a), it is possible that the absolutive argument of the embedded clause determines noun class agreement on the matrix verb. As the translations of (6) suggest, this is only possible if the embedded absolutive argument is interpreted as a topic; if it is not, the matrix verb shows default agreement (noun class IV). One possible solution to this problem would be to assume that 'bread' in (6) was actually scrambled into the matrix clause. However, as Polinsky

& Potsdam (2001: 590) point out, Tsez does not show independent evidence for cross-clausal scrambling. Thus, the agreement pattern in (6) constitutes an apparent violation of the PIC. The topic status of the agreement target can further be confirmed with overt topic marking, which is generally optional in Tsez. If the embedded absolutive is overtly topic marked, LDA becomes obligatory (7a). Furthermore, LDA is impossible with a focussed absolutive in the embedded clause (7b).

(7) Tsez (Polinsky & Potsdam 2001: 610–611)

- a. Eni-r [uʒā magalu-n/gon b-āc'ruʃi] b/*r-iy-xo.
 mother-DAT [boy bread.III.ABS-TOP ate] III/IV-know-PRS
 'The mother knows that, as for the bread, the boy ate it.'
- b. Eni-r [t'ek-kin y-igu yāʃruʃi] r/*y-iy-xo.
 mother-DAT [book.II.ABS-FOC II-good be] II/IV-know-PRS
 'The mother knows that the BOOK is good.'

The authors show that this agreement indeed crosses a clause boundary and that there is neither movement into the matrix clause nor a covert *pro* co-referent with the embedded absolutive in the matrix clause. More important for the present purpose, there are further restrictions on long distance agreement in Tsez. Non-absolutive topics, either fronted or marked by a topic particle, block LDA.

(8) Tsez (Polinsky & Potsdam 2001: 636)

- Enir [huʃ uʒā magalu bāc'ruʃi] *b/r-iy-xo.
 mother [yesterday boy bread.III.ABS ate] III/IV-know-PRS
 'The mother knows that, yesterday, the boy ate it.'

Non-absolutive *wh*-words, in situ or ex situ, also block LDA. Unfortunately, the absolutive *wh*-word *šebi* 'who, what' shows class IV agreement which cannot be differentiated from agreement with the whole embedded clause/default agreement.

(9) Tsez (Polinsky & Potsdam 2001: 634)

- Enir [ʃu micxir b-ok'ākruʃi] *b/r-iy-xo.
 mother [who.ERG money.III.ABS III-stole] III/IV-know-PRS
 'The mother knows who stole the money.'

D-linked *wh*-words in Tsez belong to different classes and can consequently be used to test the availability of LDA with *wh*-elements. And indeed, d-linked *wh*-elements can trigger long distance agreement.

- (10) Tsez (Polinsky & Potsdam 2001: fn. 20)
 Enir [šebi y-āk'iru-*ti*] y-iy-x-ānu.
 mother [**wh.II.ABS II-went-C**] **II-know-PRS-NEG**
 'The mother does not know who [of women] left.'

Lastly, LDA is also blocked by the presence of the overt complementizer *-λin* (11). In contrast, the complementizer *-ti* does not block LDA, as can be seen from the previous examples.

- (11) Tsez (Polinsky & Potsdam 2001: 635)
 * Enir [użā **magalu** b-āc'-si-**λin**] b-iy-xo.
 mother [boy **bread.III.ABS III-eat-PST.EVID-C**] **III-know-PRS**
 int.: 'The mother knows that, as for the bread, the boy ate it.'

The fact that the two complementizers behave differently is puzzling at first glance. Polinsky & Potsdam (2001: fn 19) suggest that *ti* should not be treated as a complementizer at all but rather as a derivational suffix. A different possibility would be to assume that the two complementizers occupy different positions in a complex left periphery, as has been proposed for Italian (Ledgeway 2005). For the analysis to be presented in Section 4, it is only important to note that at least one type of complementizer blocks LDA.

Long distance agreement in the Nakh-Dagestanian languages is not just restricted to Tsez but also present in at least two other related languages, namely Khwarshi and Hinuq.³

In Khwarshi (Khalilova 2008; 2009) long distance agreement is possible into complement clauses of verbs of cognition,⁴ and, in contrast to Tsez, embedded topics and embedded foci can be targeted. Again, only absolutive arguments show agreement and in addition to LDA with the embedded absolutive, the matrix verb can also show class IV agreement, which can either be treated as agreement with the whole complement clause or as default agreement, comparable to Tsez. In (12), an example for LDA with an embedded topic is shown. In its in situ position, the topic can cause optional LDA with the matrix verb. If the topic is fronted to the matrix clause, however, then the matrix verb obligatory agrees with the fronted topic.

³See Börjesson & Müller (2020 [this volume]) for a discussion of LDA in Nakh-Dagestanian languages.

⁴The other mentions that LDA possible into complement clauses of verbs of cognition but only gives examples for *to know* and also does not provide a reason for this behavior. From the author's discussion of the examples, it might be concluded that the reason for this is related to case: the subject of *to know* surfaces in lative case and the other argument with absolutive which enables the other argument to determine agreement.

(12) Khwarshi (Khalilova 2008: 118)

- a. Uža-l b/I-iq'-še [**zihe-n** b-iti-xx-u].
 boy.OBL-LAT III/IV-know-PRS **cow(III)**-& III-divide-CAUS-PFV.CVB
 'The boy knows that the cow was stolen.'
- b. **Zihe-n** uža-l b-iq'-še [b-iti-xx-u].
cow(III)-& boy.OBL-LAT III-know-PRS III-divide-CAUS-PFV.CVB
 'The boy knows that the cow was stolen.'

In addition to embedded topics, embedded foci, more specifically answers to d-linked wh-questions, can also show long distance agreement. In d-linked wh-questions, the pattern is similar to (12): the wh-element in the embedded clause can determine agreement on the matrix verb, while agreement with the whole complement clause remains a possible but dispreferred option.

(13) Khwarshi (Khalilova 2008: 390)

- [dogu **zihe** b-ot'uq'q-u] b/I-iq'-še uža-l?
 which **cow(III)** III-come-PST.PTCP III/IV-know-PRS boy.OBL-LAT
 'Which cow does the boy know came?'

Similarly in the answer, the LDA pattern is preferred over the local agreement. If the constituent corresponding to the wh-element in the question is fronted, then LDA becomes obligatory. Since constituents that correspond to wh-elements in the respective questions usually carry focus, I assume that not only topic but also focus on the agreement goal can license LDA in Khwarshi.

(14) Khwarshi (Khalilova 2008: 390)

- a. uža-l b/I-iq'-še [k^ʃaba **zihe** b-ot'uq'q-u].
 boy.OBL-LAT III/IV-know-PRS black **cow(III)** III-come-PST.PTCP
 'The boy knows that the black cow has come.'
- b. [k^ʃaba **zihe** b-ot'uq'q-u] b-iq'-še uža-l.
 black **cow(III)** III-come-PST.PTCP III-know-PRS boy.OBL-LAT
 'The boy knows that the black cow has come.'

Those complement clauses that allow LDA in Khwarshi are formed based on a nominalized form of the embedded verb, compatible with a clause union/restructuring analysis (Haspelmath 1999 for Godoberi). However, Khalilova (2009: 386–388) explicitly argues for a bi-clausal analysis based on the behaviour of reflexives and adverbs as well as the scope of negation. Thus, Khwarshi constitutes another language in which cross-clausal long distance agreement is possible, differing from Tsez in the fact that not only topics but also foci can participate in LDA.

The last Nakh-Dagestanian language to be discussed is Hinuq (Forker 2012). Again, only absolutive arguments agree and in cases in which LDA is possible, the verb can either show noun class agreement with the embedded absolutive or agreement with the whole complement clause/default agreement, which is class v. While the class of matrix verbs that allow LDA is bigger than in Khwarshi, the status of the agreement target is similar, it can either be a topic (15a) or a focus (15b).

(15) Hinuq (Forker 2012: 628)

- a. Haylo-z b-ike-s [meši čeq-i-do b-ił'i-š].
 he.OBL-DAT III-see-PST calf(III) forest-IN-DIR III-go-PST
 'He saw that the calf went into the forest.'
- b. Pat'imat-ez y-eq'i-yo [Madina-y t'ek y-ux-iš-t].
 Patimat-DAT IV-know-PRS Madina-ERG book(IV) IV-buy-RES-ABST
 'Patimat know that Madina bought the BOOK.'

In contrast to Tsez, non-absolutive wh-elements do not block LDA, and absolutive wh-elements can themselves be agreement targets in LDA constructions.

(16) Hinuq (Forker 2012: 637)

- a. Šamil-ez r/b-eq'i-yo [ni Madina-y mexer
 Šamil-DAT v/III-know-PRS where Madina-ERG money(III)
 b-uqi-š-ti].
 III-hide-RES-ABST
 'Šamil knows where Madina hid the money.'
- b. Obu-z r/Ø-eq'i-yo [ked-ez lu Ø-ike-s-ti].
 father-DAT v/I-know-PRS girl-DAT who(I) I-see-RES-ABST
 'Father knows who the girl saw.'

Another interesting difference concerning LDA in Hinuq is that it is possible across several clauses. In a sentence with three clauses, LDA is easily possible when the higher verbs all show non-local agreement (17a). It is also possible that only the intermediate verb(s) show non-local agreement (17b), but non-local agreement of the highest verb and local agreement of the intermediate one is dispreferred (17c).

(17) Hinuq (Forker 2012: 633)

- a. ʔali-ž b-eti-yo [[obu-y ec'endiyu mašina
 Ali-DAT III-want-PRS father-ERG new car(III)
 b-ux-ł'os-ti] Madina-z b-eq'-ayaz].
 III-buy-HAB-ABST Madina-DAT III-know-PURP
 'Ali wants Madina to know that father will buy a new car.'

- b. Murad-ez r-eq'i-yo [ħakim-ez y-eti-n [de kayat
Murad-DAT V-know-PRS ruler-DAT IV-want-UW PST I.ERG letter(IV)
cax-a]].

write-INF

'Murad knows that the boss wants me to write a letter.'

- c. ?di-ž b-eti-n [devez r-eq'-a [ʔu-y gulu
I-DAT III-want-UW PST you.SG.DAT V-know-INF who-ERG horse(III)
b-ik'ek-iš-ti]].

III-steal-RES-ABST

'I want you to know who stole the horse.'

Thus, even though all three languages presented in this section allow LDA, the exact implementation varies. The most important point of variation concerns the status of the agreement target, the DP in the embedded clause. While in Tsez it is only possible when the embedded DP is interpreted as a topic, Hinuq and Khwarshi allow LDA with embedded foci as well. In the next section it will be shown that the same type of variation can be found in the Algonquian language family.

2.2 Algonquian languages

A second language family in which some languages show LDA are the Algonquian languages spoken in North America over a large area, stretching from the east coast through to the Rocky Mountains. The patterns of LDA found in this language family are remarkably similar to LDA in Nakh-Dagestanian languages, in that a precondition for LDA is that the agreed-with DP in the embedded clause receives a special information-structural interpretation.⁵

Starting with Innu-aimûn, as discussed in Branigan & MacKenzie (2002), a very similar pattern to Tsez emerges. Certain matrix verbs take complement clauses and can show agreement either with the ϕ -features of the embedded subject (18b, 19a), or the embedded object (18c, 19b). However, LDA is always optional, and agreement with the whole complement clause/default agreement is possible (18a), which in Innu-aimûn is similar to transitive inanimate (TI) agreement, the agreement with inanimate objects. The complement clause is fully specified for tense and can be declarative (18) or interrogative (19), thus strongly suggesting a CP-sized complement clause.⁶

⁵See Fry & Hamilton (2014) for a comprehensive discussion of LDA in Algonquian languages.

⁶Agreement in Innu-aimûn and Passamaquoddy is very complex and a full discussion beyond the scope of the paper. I have marked the relevant agreement marker and the argument it references in bold. For further details about the agreement systems and details on the glossing, the reader is referred to the referenced literature.

- (18) Innu-aimûn (Branigan & MacKenzie 2002: 388)
- a. Ni-tshissenitamû-ânân [mûpishtuât Shûshepa Tshân mâk Maânî].
1PL-know-TI-1PL visit Joseph John and Marie
'We know that John and Marie visited Joseph.'
 - b. Ni-tshissenim-ânân-at [mûpishtuât Shûshepa Tshân mâk Mânî].
1PL-know-1PL-3PL visit Joseph John and Marie
'We know that John and Marie visited Joseph.'
 - c. Ni-tshissenim-ânân [mûpishtuât Shûshepa Tshân mâk Mânî].
1PL-know-1PL-3SG visit Joseph John and Marie
'We know that John and Marie visited Joseph.'
- (19) Innu-aimûn (Branigan & MacKenzie 2002: 399)
- a. Ma tshi-tshissenim-in [tân ishpih na nit-aimâ Mânî]?
Q 2SG-know-1SG when 1SG-called Marie
'Do you know when I called Marie?'
 - b. Ma tshi-tshissenim-âu [tân ishpih na nit-aimâ Mânî]?
Q 2SG-know-3SG when 1SG-called Marie
'Do you know when I called Marie?'

Since Innu-aimûn is a pro-drop language, the agreement target in LDA frequently is a dropped pronoun. Interestingly, if LDA takes place, the agreed-with DP can be moved to the front of the embedded clause. This movement is impossible if LDA is absent. Similar to Passamaquoddy discussed below, the question arises whether the moved embedded DP ends up in a position in the matrix clause or a left-peripheral position of the embedded clause. The authors, in contrast to Bruening (2001) for Passamaquoddy, assume that the DP ends up in the matrix clause. However, (20) is also compatible with the dislocated DP still being in the embedded clause, simply higher than the *wh*-element or complementizer.

- (20) Innu-aimûn (Branigan & MacKenzie 2002: 389)
- a. Tshi-tshissenim-âu-â [Mânî tshekuân kuet aimiât Pûna
2SG-know-TA-3SG-Q Marie why called Paul
utshimâminua]?
boss
'Do you know why Marie called Paul's boss?'

- b. N-uî-tshissenim-âu [kassinu **kâuâpikueshit** tshetshî
 1SG-want-know-3SG every **priest** if
 mûpishtâshkuenit].
 visited-2SG/INV

‘I want to know if every priest visited you.’

The authors argue at length against a proxy-agreement account, to be discussed shortly, and instead propose that the agreement target in the embedded clause carries an unvalued A'-feature that they term *O*-feature which allows it to move covertly into the specifier of the CP so it is available for agreement with the matrix verb. Additionally, the authors claim that the agreement goal in the embedded clause is usually interpreted as the topic of that clause, so that the most likely candidate for the *O*-feature is a topic feature. This is in line with the observation mentioned earlier that the LDA goal is frequently a dropped pronoun, a highly topical element. The only qualification to this assumption is presented by *wh*-elements, since they also can appear as targets for LDA. Thus, the left periphery seems to play the crucial role in licensing LDA: either the LDA agreement goal is information-structurally marked or moved high enough in the left periphery of the embedded clause due to independent reasons like *wh*-movement.⁷

Thus, Innu-aimûn shows many properties of LDA already discussed for Nakh-Dagestanian languages, since it is possible either with embedded topics or with *wh*-elements in the left periphery of the embedded clause. Similarly, if the embedded agreed-with DP is overtly topic marked by dislocation, LDA is obligatory, just as in Tsez when the topic is marked overtly by a particle.

Another Algonquian language that shows long distance agreement is Passamaquoddy (Bruening 2001). Here, the picture is more complex, just as the data in Hinuq appear to be more complex than in Tsez. Passamaquoddy has a raising to object construction, raising the embedded object across embedded C, that causes object agreement on the matrix verb (21). However, the actual raising part of this operation is optional, such that a LDA configuration is created (22), in which the matrix verb does not need to agree with the highest argument in the embedded clause. If LDA takes place, the agreed-with argument in the lower clause is usually interpreted as topical or focussed.

⁷Unfortunately, the authors do not discuss whether the *wh*-element can be targeted for LDA when a topicalised element precedes it.

(21) Passamaquoddy (Bruening 2001: 258)

- a. 'Kosiciy-a-l [yaq uhsimis-ol eli keka
3SG-know.TA-DIR-OBV QUOT 3.younger.sib-OBV c almost
peciya-li-t].
come-OBV.S-3SG.CONJ

'She knew that her brother had almost arrived.'

- b. Susehp 'kosiciy-à [akòm eli Muwin
Susehp 3SG-know.TA-DIR.OBV.PL snowshoe.OBV.PL c M.
kisi-mil-at Wiphun].
PERF-give-3SG.CONJ W.

'Susehp knows that Muwin gave Wiphun snowshoes.'

(22) Passamaquoddy (Bruening 2001: 259)

- a. N-wewitaham-a-k [ma=te nomiy-a-w-ik
1SG-remember-DIR-3PL NEG=EMPH see-DIR-NEG-PART.3PL
mawsuwinuw-ok Kehlis-k].
person-3PL Calais-LOC

'I remember that I didn't see people in Calais.'

- b. N-kosicihtun-ol [eli Piyel nokkaht-aq sukolis-ol
LSG-know.TI-INAN.PL c Piyel eat.up-3SG.CONJ candy-INAN.PL
wikahtm-an-pon-il].
like.eat-L.CONJ-PRET-PART.INAN.PL

'I know that Piyel ate up the candies that I liked.'

LDA is also possible in embedded questions, either with the *wh*-element (23a) or a different argument. If LDA in embedded questions takes place with an argument different from the *wh*-element and is also accompanied by overt movement, this agreed-with argument may end up in a position to the left of the embedded *wh*-element (23b). Since Passamaquoddy is a *wh*-movement language, the fronted argument consequently either occupies a position in the left periphery even higher than the *wh*-element in the embedded clause or a low position already in the matrix clause, similar to what Branigan & MacKenzie (2002) argue for Innu-aimûn. The author argues extensively for the first option, which thus makes LDA in Passamaquoddy similar to LDA in the Nakh-Dagestanian languages.

(23) Passamaquoddy (Bruening 2001: 259)

a. Tihtiyas ma=te wewitaham-a-wiy-il [wen-il amsqahs
 Tihtiyas NEG=EMPH remember-*DIR*-NEG-OBV **who-OBV** first
 kis-aqosom-uw-iht kiwhosu].
 PERF-cook-APPL-3SG.CONJ.INV muskrat.OBV.PL
 ‘Tihtiyas doesn’t remember who first cooked muskrat for her.’

b. N-kosiciy-a-k [**nuhuw-ok** muwinuw-ok keq
 1SG-know.TA-*DIR*-3PL **three-3PL** **bear-3PL** what
 kis-temu-htit].
 PERF-eat-3PL.CONJ
 ‘I know what the three bears ate.’

Similarly to Branigan & MacKenzie (2002) for Innu-aimûn, Bruening (2001) conclusively argues against possible alternatives to the movement analysis. However, the exact nature of the landing site of the movement in the left periphery of the embedded clause is not clear. The agreement target in the embedded clause is compatible with either a topic or a focus interpretation. This is shown by either marking the embedded argument with the contrastive topic marker *olu*, (24a), or modifying it with the focus sensitive particle *tehpû* ‘only’, (24b).

(24) Passamaquoddy (Bruening 2001: 282)

a. Ma=te n-kosiciy-a-wi [**wot** **olu n-tatat**, tan-iyut
 NEG 1SG-know.TA-*DIR*-NEG **this.AN TOP** 1SG-father WH-this.INAN
 keti-nomkuwal-s-it atomupil].
 IC.FUT-lend-INTRANS-3SG.CONJ car
 ‘I don’t know which car, my father, he’s going to buy.’

b. N-kosiciy-a [**tehpû** Susehp oc menuwa-c-ihî
 1SG-know.TA-*DIR* **only** Susehp FUT IC.buy-3SG.CONJ-PART.OBV.PL
 nuhu akom].
 three.OBV.PL snowshoe.OBV.PL
 ‘I know that only Susehp would buy three snowshoes.’

If two possible agreement targets compete, i.e. when the embedded clause contains a focus and a topic, the topic can be the target of LDA while skipping the focussed argument, as already shown in (23b), where a topic serves as goal for LDA despite the presence of a focussed element, the *wh*-word.

This leads Bruening (2001) to the conclusion that the landing position of the raising to object movement cannot be a single dedicated topic or focus projection. Assuming a split-CP analysis (Rizzi 1997), the data are compatible with an analysis of the landing position in terms of discourse projections: both the topic projection and the focus projection can serve as a landing site for raising to object movement and agree with the respective constituent.⁸ These data show striking similarities with respect to LDA not only between Passamaquoddy and Innu-aimûn, but also between those two languages from the Algonquian family and the languages from the Nakh-Dagestanian family, Tsez, Khwarshi, and Hinuq.

As already pointed out by Polinsky (2003), it is important when discussing LDA in Algonquian languages to carefully distinguish genuine LDA from cases of proxy-agreement or prothetic agreement which is not a case of LDA even though it appears to be one on the surface. In those languages that are viable to that kind of analysis, a silent pronoun in the matrix clause, which is co-referential with the element in the embedded clause serves as the actual agreement target of the matrix predicate.⁹ Due to the co-referentiality of the proxy argument in the matrix clause and the embedded argument, such a structure appears to show LDA even though only clause bound agreement is involved. Such a construction would be similar to English (25), with the matrix pronoun being a silent *pro*.

⁸Some data appear to be incompatible with this analysis. Singular, non-referential quantifiers can appear in embedded questions preceding wh-elements (i).

- (i) Passamaquoddy (Bruening 2001: 282)
 Sapet 'kosiciy-a-l [psi=te wen-il tan-iyuhtol
 Sapet 3SG-know.TA-DIR-OBV all=EMPH someone-OBV WH-that.OBV
 nucitqonkelic-il kisi-tqon-at].
 policeman-OBV PERF-arrest-3SG.CONJ
 'Sapet knows which policeman arrested everybody.'

However, first, it is not obvious that the fronted quantifiers cause long distance agreement on the matrix verb. If the language provides other means to move an element in the embedded clause above the wh-element, then LDA could for example also be based on overt QR. Second, Bruening (2001) argues that sometimes the element that appears to have undergone raising to object is actually initially merged in the matrix clause, which could also be the case for (i). Thus the data do not constitute counterevidence to the information-structural dependence of LDA.

⁹At least one argument used by Polinsky (2003) in favor of a proxy-agreement account is directly addressed and dismissed by Bruening (2001) as well as Branigan & MacKenzie (2002). In Passamaquoddy as well as Innu-aimûn, agreement on the matrix verb can be with a subset of the goal in LDA configurations. Polinsky argues this is due to a *pro* in the matrix clause, but Bruening (2001: 269) shows that it can also occur in contexts which definitely involve movement like relative clauses and thus dismisses this argument.

(25) * Peter knows of her that Mary went to the movies.

Note the ungrammaticality of (25), which is due to a violation of Principle C. Even though Principle C is not necessarily active in all Algonquian languages, Branigan & MacKenzie (2002) argue that it does play a role in Innu-aimûn and thus an analysis along the lines of proxy agreement is not feasible. This is supported by further arguments, for example the impossibility of having an English (26) with a proxy argument in the matrix clause, i.e. having a prothetic pronoun co-referring with a *wh*-element.

(26) * Do you know of him who is laughing?

For Passamaquoddy, Bruening (2001) also discusses and dismisses a proxy agreement account, relying on the same arguments as Branigan & MacKenzie (2002). He shows that having an overt pronoun in the matrix clause doubling the embedded DP is ungrammatical, probably due to a Principle C violation, (27). The observation that LDA in Passamaquoddy can also target the *wh*-element of the embedded clause and that embedded *wh*-elements cannot be coreferential with a matrix pronoun (cf. 26), provides another argument against a proxy agreement analysis.¹⁰

(27) Passamaquoddy (Bruening 2001: 270)

- a. N-kosiciy-a [eli Piyel koti-nathula-t
1SG-know.TA-DIR c Piyel FUT-pick.up.in.boat-3SG.CONJ
Susehp-ol].
Susehp-OBV
'I know that Piyel will pick up Susehp in a boat.'
- b. * N-kosiciy-a **nekom** [eli Piyel koti-nathula-t
1SG-know.TA-DIR 3SG c Piyel FUT-pick.up.in.boat-3SG.CONJ
Susehp-ol].
Susehp-OBV
'I know about him that Piyel will pick up Susehp in a boat.'

Before closing this section, it should be pointed out that even though the proxy-agreement account of LDA cannot be applied to LDA in Passamaquoddy or Innu-aimûn, other Algonquian languages do seem to present cases of agreement by

¹⁰The author also argues against a different kind of the proxy agreement analysis in which either a proxy or the full DP is base generated in the left periphery of the embedded clause, comparable to a left dislocation construction. I cannot discuss this here for reasons of space, but Bruening (2001: 263ff) provides ample evidence that movement inside the embedded clause does indeed take place.

proxy. Blackfoot for example does show apparent LDA (28). Blackfoot also presents an apparent counterexample to the generalization stated in the beginning that either topic and focus can serve as LDA targets in a given language or topic alone, since Bliss (2009) argues that LDA in Blackfoot marks contrastive foci.

- (28) Blackfoot (Bliss 2009: 1)
nit-iksstaat-a an-wa Leo nin-aahk-sspommo-a-hsi.
1SG-want.TA-1SG:3SG DEM-PROX Leo 1-MOD-help.TA-1SG:3SG-CONJ
'I want to help Leo.'

However, LDA in Blackfoot has been analysed as proxy agreement and thus does not constitute a proper case of LDA and therefore also not a counterexample to the generalization stated above. Polinsky (2003), going back to Frantz (1978), discusses several phenomena that can be linked to LDA by proxy. One of them concerns binding. In Blackfoot, since the proxy in the matrix clause constitutes a full pronoun, it can actually bind a reflexive on the verb.

- (29) Blackfoot (Frantz 1978: 99, via Polinsky 2003: 286)
noxkówa ki niistówa nits-iksstat-tsiiyi-xpinnaani
my.son.3 and 1SG 1SG-want-RECP-1PL
n-áxk-a'po'tak-ss-innaan.
1-might-work-CONJ-1PL
lit.: 'My son and I want of each other that we work.'

I have argued in this section that, in addition to languages in the Nakh-Dagestian family, some Algonquian languages also show LDA, also conditioned by information-structural properties of the embedded agreement target. Similarly to Nakh-Dagestian languages, two types of LDA targets can be distinguished in Algonquian languages: either agreement with both topics and foci is possible, as in Passamaquoddy, or LDA is restricted to topics as in Innu-aimûn. Additionally, I have also distinguished LDA proper from LDA by proxy, as exemplified by Blackfoot. Concerning LDA by proxy in Nakh-Dagestian languages, Polinsky (2003) has argued that Tsez show proper LDA. This has not been discussed explicitly for Hinuq or Khwarshi, but data strongly suggest an analysis in the line of Tsez. The next subsection will discuss LDA in yet another unrelated language family, however in a slightly different context.

2.3 LDA in Uyghur

Another, less frequently discussed, case of an apparent PIC violating dependency is exceptional subject case marking in Altaic languages. In many languages from

this family, subjects in certain types of finite embedded clauses can occur with a case other than nominative, mostly genitive but also accusative. Even though the occurrence of non-canonical subject case is different from the LDA agreement pattern discussed above, in which a matrix verb agrees with an argument in a complement clause, those cases could still constitute instances of LDA, namely when the exceptional case is licensed from outside the clause and the clause itself is of CP size. Thus, these two aspects need to be considered carefully, and Table 1 shows that, as expected, significant variation along these two dimensions can be found in the Altaic languages.

Table 1: Exceptional Subject Case in Altaic.

Language	Emb. clause size	Licenser for subject
Turkish (Kornfilt 2008)	CP	clause internal C
Dagur (Hale 2002)	AspP (<CP)	clause external D
Japanese (Miyagawa 2011)	TP (<CP)	clause external D
Uyghur (Asarina & Hartman 2011a)	CP	clause external D

As can be seen in the table, Uyghur seems to present the right configuration that exceptional subject case could be analyzed as LDA, since the case of the subject in a CP is licensed by an element from outside that CP. If case licensing is taken to be based on agreement, and this analysis is correct, then Uyghur shows a case of LDA. Examples of this configuration are given in (30).

(30) Uyghur (Asarina & Hartman 2011b: 4)

- a. [**Men-ij** ji-gen] tamaq-**im** jaxʃi.
 1SG-GEN eat-RAN food-1SG.POSS good
 ‘The food that I ate is good.’
- b. [**Ötkür-nij** oqu-ƙan] kitav-i uzum.
 Ötkür-GEN read-RAN book-3SG.POSS long
 ‘The book that Ötkür read is long.’
- c. [Ötkür oqu-ƙan] kitap uzum.
 Ötkür read-RAN book long
 ‘The book that Ötkür read is long.’

In order to make this point more clear, I will summarize the discussion in Asarina & Hartman (2011a), first showing that it is indeed clause external D that licenses genitive subject case in relative clauses and NP complement clauses in Uyghur

and second, also presenting arguments in favour of a CP sized embedded clause. The first argument for licensing GEN on the embedded subject by an external head comes from agreement with the embedded subject that shows up on the external D head only when the subject carries genitive case, (30a–30b). If the embedded subject is unmarked (which is a free variant), then there is no agreement on the external D (30c). The second piece of evidence comes from the observation that clause external D can only assign genitive once. Thus the sentence in (31a) is ambiguous, and double genitives as in (31b) are impossible.

- (31) Uyghur (Asarina & Hartman 2011a: 3)
- a. Ajgül-nuñ resim-i
Aygül-GEN picture-3SG.POSS
'picture belonging to Aygül' or 'picture depicting Aygül'
 - b. *Ötkür-niñ Ajgül-nuñ resim-i
Ötkür-GEN Aygül-GEN picture-3SG.POSS
int.: 'picture that depicts Aygül and belongs to Ötkür'

If in relative clauses and nominal complements the clause external D head is responsible for genitive assignment to the subject of the embedded clause, genitive subjects should be in complementary distribution with genitive possessors while unmarked subjects should be able to occur with genitive possessors since the genitive of the clause external D is still available when it is not assigned to the subject. The data in (32) show that this prediction is borne out.

- (32) Uyghur (Asarina & Hartman 2011a: 3)
- a. * [Ötkür-niñ oqu-ğan] Ajgül-nuñ kitav-i uzum.
Ötkür-GEN read-RAN Aygül-GEN book-3SG.POSS long
int.: 'Aygül's book that Ötkür read is long.'
 - b. [Ötkür oqu-ğan] Ajgül-nuñ kitav-i uzum.
Ötkür read-RAN Aygül-GEN book-3SG.POSS long
'Aygül's book that Ötkür read is long.'

Following the authors, I conclude from this that the genitive subject is indeed licensed from outside the clause. However, to analyze this phenomenon as a case of LDA proper, it is also necessary to show that the embedded clauses are CPs and not smaller constituents as in other Altaic languages. The authors support the assumption of CP sized embedded clauses with two arguments, CP adverbs and embedded questions. First, if the embedded clause can host CP adverbs, then it must be treated as a CP. The clause initial adverb *xeqiqi* 'truly' in (33) shows the possibility of CP adverbs.

(33) Uyghur (Asarina & Hartman 2011a: 8)

- a. [χ eqiqi **Aygül-niŋ** jaz-**ƚan**] kitiv-i-ni korset!
 truly **Aygül-GEN** write-RAN book-3SG.POSS-ACC show
 ‘Show me the book that Aygül truly wrote.’
- b. [χ eqiqi **men-ij** jaxfi kör-i-gen] tamaq-**im-ni** ber!
 truly **I-GEN** well see-IMP-F-RAN food-1SG.POSS-ACC give
 ‘Give (me) the food that I truly like.’

Second, the embedded clause can also be a question and it is argued that interpretation as a question necessarily requires a CP layer in the embedded clause to host the clause type information.

(34) Uyghur (Asarina & Hartman 2011a: 8)

- Men [**Ajgöl-nuŋ** katfan ket-ken-(lik)] heqiqet-i-ni sordum.
 1SG **Aygül-GEN** when leave-RAN-LIQ fact-3SG.POSS-ACC asked
 ‘I asked when Aygül left.’

Summarizing the discussion, genitive subject licensing in Uyghur constitutes a genuine case of long distance agreement. However, Asarina & Hartman (2011a) reject an analysis in the spirit of Polinsky & Potsdam (2001) to account for the apparently PIC violating pattern, since not all genitive subjects are topics. As I have discussed above, not only topics can participate in LDA though, but also foci. The following data show exactly that, namely that in addition to the LDA cases above, which Asarina & Hartman (2011a) claim involve topical subjects, foci are also possible in LDA in Uyghur. The examples (35a) and (35b) contain a focused subject marked by the focus sensitive clitic *-la* ‘only’, while (36) contains a contrastive focus.¹¹

(35) Uyghur (Asarina & Hartman 2011a: 10)

- a. [**Ötkür-niŋ-la** kel-gen-lik] xever-i muhim.
Ötkür-GEN-only come-RAN-LIQ news-3SG.POSS important
 ‘The news that only Ötkür came is important.’
- b. [**men-ij-la** jaxfi kör-gen] kitav-**im** uzun.
 1SG-GEN-only well see-RAN book-1SG.POSS long
 ‘The book that only I like is long.’

¹¹This stands in contrast to subjects in finite ECM clauses in Turkish. As Şener (2008) shows, the subjects, which receive exceptional accusative case, need to be topics. Thus, Turkish ECMs might yet present another case of LDA in an Altaic language, with the variation with respect to the information-structural features of the agreement target being in line with the variation found in other language families.

(36) Uyghur (Asarina & Hartman 2011a: 10)

Q: Ötkür [Aǵǵül-nuñ kel-gen-lik-i-ni] didi-mu?

Ötkür Aǵǵül-GEN come-RAN-LIQ-3.POSS-ACC said-Q

‘Did Ötkür say that Aǵǵül came?’

A: Yaq, Ötkür [Mehemmet-niñ kel-gen-lik-i-ni] didi.

no Otkür Mehemmet-GEN come-RAN-LIQ-3.POSS-ACC said

‘No, Ötkür said that MEHEMMET came.’

Thus, LDA in Uyghur seems to behave parallel to Hinuq, Khwarshi, and Passamaquoddy, in that not only embedded topics can serve as agreement target but embedded foci as well. This sensitivity to information structure seems to be the uniting property of LDA in all languages discussed in this section, with either topic, or topics and foci being possible agreement targets for an element in a higher clause, a verb in the case of Algonquian and Nakh-Dagestanian languages, and a D-head in the case of Uyghur. The findings from this section are summarized in Table 2.

Table 2: LDA and IS property of goal.

IS property	Dependency	
	ϕ -features	case
TOP	Tsez (Polinsky & Potsdam 2001) Innu-aimûn (Branigan & MacKenzie 2002)	Turkish (Şener 2008)
TOP+FOC	Hinuq (Forker 2012) Khwarshi (Khalilova 2009)	Uyghur (Asarina & Hartman 2011a)

The next section presents previous analyses of this phenomenon, their advantages and their problems, before I turn to my own analysis in Section 4.

3 Previous analyses

Long distance agreement and the challenge it presents to standard locality theories, more specifically the Phase Impenetrability Condition (PIC), have made it a frequently discussed phenomenon in the syntactic literature. Two possible ways to solve the apparent incompatibility between LDA and the PIC are immediately obvious. One could either assume that agreement is simply not subject to locality

constraints as strong as they are for movement, an approach taken for example by Bošković (2007b). In contrast, the other approach relies on the strong parallel between agreement and movement by claiming that the way moved constituents can, on the surface, violate the PIC, is also the solution to LDA, namely successive cyclicity. Thus, similarly to elements moving out of phases by successive cyclically moving first into a phase edge position, agreement crossing a phase boundary should be possible via an element in the phase edge serving as intermediate agreement step. Arguments in favor of such a cyclic agreement process are frequently found in the literature, for example in Legate (2005) for a theory of agreement of T with a DP in situ but also mentioning LDA data, and, in the framework of Tree Adjoining Grammar, in Frank (2006).

Worked-out proposals for long distance agreement are much rarer in the literature than general approaches suggesting agreement through phase edges. The two most discussed proposals are Polinsky & Potsdam (2001), who actually suggest LF movement of the agreement goal in the embedded clause into the phase edge, and, more recently Bjorkman & Zeijlstra (2019) who assume an agreement process through the periphery of the embedded clause. Since Polinsky & Potsdam (2001) as well as Bjorkman & Zeijlstra (2019) (including its discussion in Preminger 2013 and Preminger & Polinsky 2015), in part deal with data that are also the focus of this paper, I will discuss their contributions in more detail in this section, starting with Polinsky & Potsdam (2001). The main facts that all proposals of LDA need to account for are the dependence on information-structural properties of the agreement goal as well as the locality restrictions to which LDA is subject in the various languages.

3.1 Polinsky & Potsdam (2001)

In their paper on LDA in Tsez, Polinsky & Potsdam (2001) propose to account for the pattern by assuming that the embedded topic moves on LF into a left peripheral topic position (Rizzi 1997) in the embedded clause. This position then is close enough to the matrix verb (cf. also Bobaljik 2008), part of the phase edge in our terms, to serve as the agreement goal of the probing matrix v.¹² Alternatives, for example that the agreement goal has continued to move covertly to the matrix clause or that there is a *pro* co-indexed with the embedded topic in the matrix clause, are conclusively argued against by the authors. A structural representation of their approach is given in (38).¹³

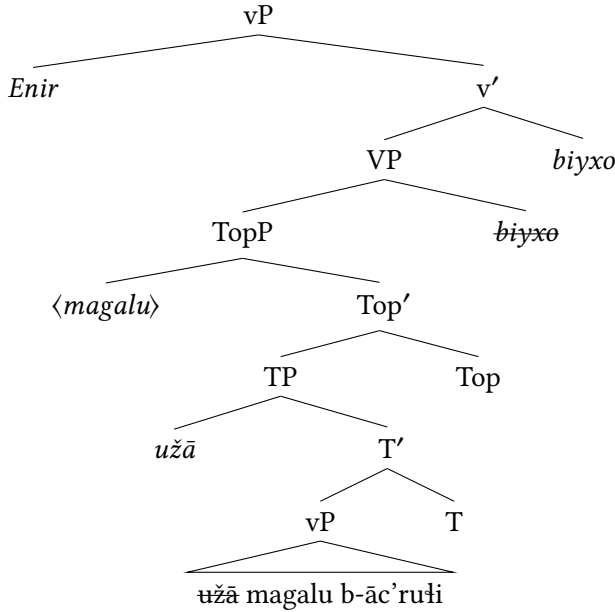
¹²I assume the ϕ -probe in the matrix clause is on v whereas the authors seem to relegate it to V. Nothing depends on that.

¹³Note that ~~strike through~~ represents traces of overt movement while $\langle XP \rangle$ is used for covertly moved elements.

(37) Tsez (Polinsky & Potsdam 2001: 584)

Eni-r [užā magalu bāc'ruḥi] b-iy-xo.
 mother-DAT boy bread.III.ABS ate III-know-PRS
 'The mother knows that, as for the bread, the boy ate it.'

(38)



From this assumption nearly all properties of LDA in Tsez can be derived. First, the topic status of the agreement goal in the embedded clause is crucial. If the element does not serve as the topic of the embedded clause, then it will not move to spec-TopP on LF, meaning it will be too far away from matrix *v* to serve as agreement goal. This is further supported by the observation that if the embedded absolutive argument is overtly topic marked (which is usually optional), the LDA even becomes obligatory. Second, the fact that non-absolutive topics, overtly marked or not, block LDA, can also be easily accounted for. Only absolutive arguments show agreement in Tsez in general. If the only position that can serve as goal for LDA is occupied by a non-absolutive element, LDA becomes impossible. The same holds for non-absolutive *wh*-elements, either in situ in the embedded clause or moved to the left periphery of the embedded clause. If it is assumed that the landing site of *wh*-elements in the embedded clause is above TopP they provide a closer goal hosting ϕ -features for the probing matrix *v* than the topic and they thus block LDA since ϕ -features of non-absolutive elements cannot participate in agreement and therefore lead to default agreement on matrix *v*.

Note on the other hand that this implies that *wh*-elements should not block LDA when they are absolutive and instead provide suitable agreement targets. This is indeed the case, as can be seen in (10), repeated for convenience in (39).

- (39) Tsez (Polinsky & Potsdam 2001: fn. 20)
 Enir [*šebi* y-āk'iruʃi] y-iy-x-ānu.
 mother **wh.II.ABS** II-went II-know-PRS-NEG
 'The mother does not know who [of women] left.'

The blocking effect of complementizers can be accounted for in the same fashion. Assume that complementizers also carry a set of valued ϕ -features in Tsez. Then it is expected that if the complementizer is situated structurally above the TopP, it will block LDA due to providing a closer agreement goal. If the complementizer is situated below TopP, it should not have this effect. Both types of complementizers can be found in the language. The high complementizer *-λin* blocks LDA, as can be seen in (11), repeated in (40). In contrast, the complementizer *-ti* does not block LDA, as can be seen from the previous example. Polinsky & Potsdam (2001) only assume one possible high complementizer position and therefore argue that *-ti* should not be analysed as complementizer but as nominalizer. However, assuming two possible positions for complementizers in the left periphery, a high one and a low one, as has been proposed for Romance languages (Ledgeway 2005), allows an analysis of *-ti* as complementizer.

- (40) Tsez (Polinsky & Potsdam 2001: 635)
 * Enir [uʒā **magalu** b-āc'-si-**λin**] b-iyxo.
 mother [boy **bread.III.ABS** III-eat-PST.EVID-C] III.know
 int.: 'The mother knows that, as for the bread, the boy ate it.'

Turning to languages other than Tsez, it appears easy to extend the proposal by Polinsky & Potsdam (2001) to cases in which not only embedded topics but also embedded foci can participate in LDA. The only additional assumption that needs to be made is that it seems to be a language specific property whether only topics or both topics and foci covertly move to peripheral positions. However, this assumption is more problematic than it appears at first glance, since it is generally assumed that all information-structural information, is always encoded in the CP-periphery. Thus, if covert movement is assumed, then the distinction is usually not made between whether elements move or do not move (overtly or covertly) to a peripheral position but only whether or not this movement is overt or covert. Consequently, it is actually expected that if LDA is possible for (covertly moved) topics it should always also be possible for (covertly moved) foci, contrary to

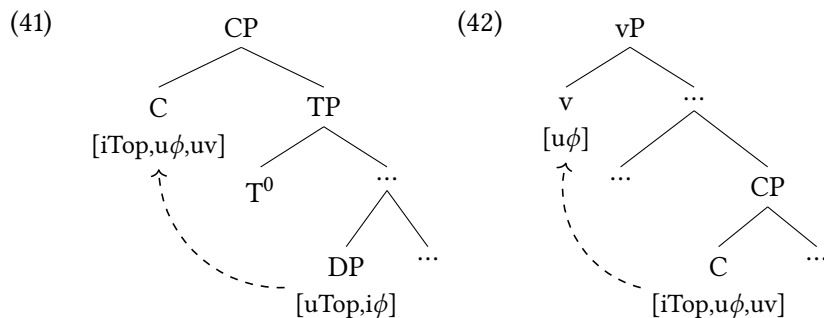
what can be observed. A second, more theoretical problem concerns the reliance of Polinsky & Potsdam (2001) on LF movement of the topic. As already pointed out in earlier versions of Bjorkman & Zeijlstra (2019), if LF movement is taken to be a post-syntactic process, then its output should not be able to feed narrow syntactic processes like agreement. The argument of Preminger & Polinsky (2015) that if the embedded clause is taken to be its own domain, LF movement in that domain or phase can very well serve as input to the narrow syntax of the next higher domain/phase, also does not hold since it would wrongly predict long distance QR. A possible solution to that would be a single output syntax as proposed by Bobaljik (2002) in which all movement happens in the syntax and spell out can target different copies. I will not follow this assumption and instead propose an analysis that is compatible with the standard model of grammar that does not rely on LF movement but on agreement in the next section. Before that, however, I will first discuss another proposal from the literature for LDA that also relies on agreement, albeit a non-standard version of it, namely Bjorkman & Zeijlstra (2019).

3.2 Bjorkman & Zeijlstra (2019)

Bjorkman & Zeijlstra (2019) discuss long distance agreement as an argument supporting their general theory of upwards instead of downwards agreement (see also Baker 2008 and Wurmbrand 2012; 2016a,b). Even though the direction of agreement is different from what I assume in my own proposal, their general argument can also be made in a standard, downward agreement framework. The two authors assume a configuration very much in the spirit of the successive cyclic agreement idea of Legate (2005) without any recourse to LF movement. It is assumed that the agreement goal, the DP in the embedded clause, carries an uninterpretable topic feature, [uTop], and also interpretable valued ϕ -features. The [uTop] feature probes upwards and establishes an agreement relation with the [iTop] feature on the C head, allowing the C head to value its own $u\phi$ -features (41). Even though not discussed explicitly, the authors appear to assume a feature bundle on a C-head which comprises information-structural features and ϕ -features, with the valuation of one being dependent on an agreement relation established based on the other. This will also be the core of my own proposal and elaborated on in more detail in the next section.

To establish the second link in the agreement chain, the authors assume that the C head also carries an unvalued feature, namely an unvalued case feature [uv], which allows the C-head to establish an agreement relation with the matrix *v*, so that *v* can value its ϕ -features (42), which then show up as LDA on the matrix

verb, following the frequently assumed connection between case assignment and agreement.¹⁴



Several points of criticism can be raised against this proposal, many of which have already been discussed in Preminger & Polinsky (2015). From an empirical point of view, there is no obvious explanation for the various blocking effects observed in Tsez. Neither complementizers, nor non-absolute topics or *wh*-elements should block the relationship between the C head and the higher *v* since those elements cannot assign case, therefore do not carry a valued *v* feature and thus cannot intervene in the relation between embedded C and matrix *v* based on this feature. Second, it seems not straight-forward in this system to derive cross-clausal long distance agreement as can be found in Hinuq (17). While Polinsky & Potsdam (2001) could stipulate successive cyclic LF-movement of the topic DP, Bjorkman & Zeijlstra (2019) would have to assume that each *v* head that participates in cross-clausal LDA carries its own [uTop] feature to establish an agreement relationship with the next higher C head.

From a theoretical point of view, Bjorkman & Zeijlstra (2019) have to assume that in an extended left periphery in the sense of Rizzi (1997), it is a language specific property whether the topic head or the focus head hosts the [uv] case feature to account for the fact that in some languages foci can also trigger LDA. However, if it is assumed that this is an idiosyncratic property of a specific head in the left periphery, it becomes impossible to account for the generalization that if foci can trigger LDA in a specific language, this language will have LDA based on topics as well. Additionally, as has been discussed for Uyghur, LDA is sometimes also possible in relative clauses and DP complements. For the approach of Bjorkman & Zeijlstra (2019), it follows that in these languages, the C head in

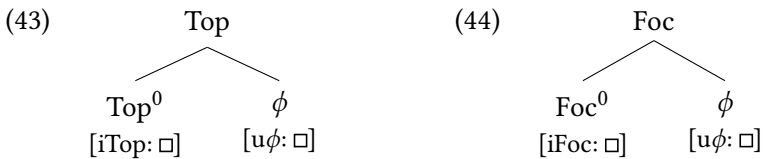
¹⁴This assumption has frequently been questioned, for example in Bobaljik (2008) and more recently in Baker & Bobaljik (2015) and Bány (2015).

the periphery of the embedded clause does not carry a [uv]-feature but a [uD]-feature, since, in this case, D is the case assigner, assigning genitive case to the embedded subject. Lastly, turning again to Tsez and other ergative-absolutive languages, linking the assignment of absolutive case to *v* appears problematic, as absolutive is frequently linked to a higher case assigner since it is usually the unmarked case and shows up, for example, in unaccusatives and passives.

Summing up the discussion, both approaches presented in this section assume some kind of cyclical process that connects matrix verb with embedded agreement goal via an intermediate step in the left periphery. However, both approaches suffer from empirical and theoretical problems that are either linked to the assumption of LF movement as input to later syntactic process or due to the features that are assumed to be involved in the agreement process. In the next section, I will argue for another proposal, that keeps the idea of a cyclic agreement process similar to Bjorkman & Zeijlstra (2019) but does not involve case features.

4 LDA conditioned by information structure

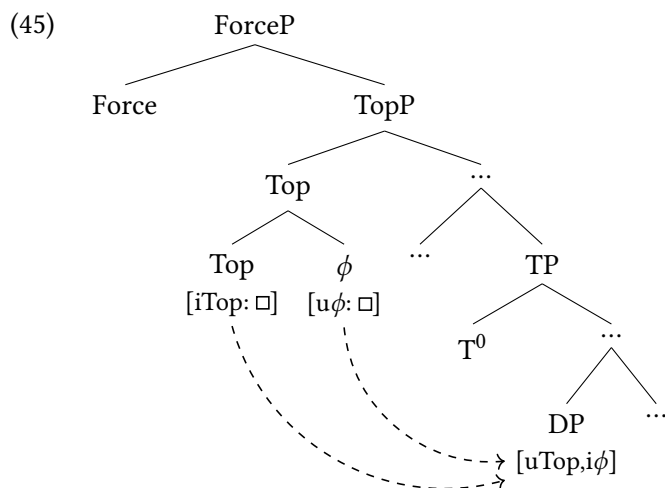
I propose to keep the idea of Bjorkman & Zeijlstra (2019) of a cyclic agreement process through the left periphery by capitalizing on the idea that ϕ -features can be bundled with information-structural features in the left periphery and that the valuation of the former depends on an agreement process established based on the latter. Concretely, I propose the following structures for the information-structural heads in the left periphery, (43) for languages like Tsez in which only topics participate in LDA and (44) for languages in which this is also possible for foci.



Following Chomsky (2008) and the idea of feature inheritance, I assume that the agreement features that are manifested in T in English and other languages are actually introduced in C, since C is the relevant phase head, and then only inherited by T. The particular type of feature inheritance is, however, a language specific property and Miyagawa (2010; 2017) assumes that this is the factor that conditions whether a language is agreement driven or discourse configurational.

In the case of the former, for example English, ϕ -features are inherited by T and information-structural δ -features remain in C. In case of the latter, for example Hungarian, it is the reverse, with ϕ -features staying in C and δ features being inherited by T. In line with Jiménez-Fernández (2010) and Miyagawa (2017), I assume that there is a third type, namely languages in which both types of features can remain bundled on the same head giving rise to the structures in (43) and (44).¹⁵

Assuming that the left periphery of the embedded clause in LDA contexts contains a head in which ϕ - and δ -features are bundled together provides the intermediate agreement step necessary to connect the higher verb with the lower DP. The derivation for a language like Tsez, in which topicality of the agreement goal is the decisive factor, proceeds as follows. In a first step, represented in (45), the topic head in the left periphery agrees with the embedded topical DP, based on the [uTop]/[iTop] feature pair. At the same time, this allows the ϕ -features of the topic head to be valued by the agreed-with DP. In other words, the valuation of the ϕ -features depends on an agreement relation established by information-structural features.

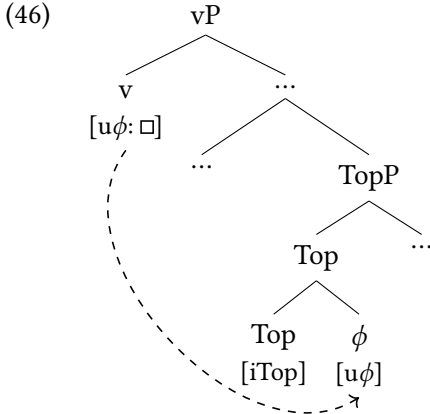


In a second step, matrix *v* with its own set of unvalued ϕ -features probes in its c-command domain and agrees with the ϕ -features that are now hosted by the topic head in the periphery of the embedded clause.¹⁶ Under the assumption that

¹⁵The question then arises whether T really inherits the features from C in all cases or whether it is also possible in certain languages that the features are not inherited but copied so that another instance of the feature(-bundle) remains in C.

¹⁶In contrast to Bjorkman & Zeijlstra (2019), I only link ϕ -feature agreement to *v* and not to absolutive case.

the verb in the matrix clause moves at least as high as v, these ϕ -features then end up being spelled out on the matrix verb.



In this proposal, most of the properties of LDA that have been discussed above can easily be accounted for. First, the blocking effect of non-absolute topics is simply due to the fact that only absolute arguments can participate in ϕ -feature agreement. Thus, even though an agreement relation between topic head and embedded non-absolute topic can be established, the ϕ -features on the topic head remain unvalued, leading to the spell-out of the default agreement on matrix v. Similarly, under the assumption that ϕ -features are also present on complementizers, those features provide a closer agreement goal for v than the ϕ -features of the embedded topic. However, since the complementizer does not carry absolute case, v cannot agree with it, leading again to the spell out of the default feature value. Appealing to locality can also account for the observation that absolute wh-elements can lead to LDA. LDA is possible when the agreement goal manages to get high enough in the periphery of the embedded clause. Since wh-elements are usually connected to the left periphery by either movement or agreement, their participation in LDA is expected, as long as they carry absolute case.¹⁷

Turning to properties of LDA in languages other than Tsez, the generalization that if foci can participate in LDA in a particular language, topics can participate as well follows directly from the dependence on information structure and the fine structure of the left periphery as assumed by Rizzi (1997), as exemplified in (47).

¹⁷This property of LDA has been exploited by Heck & Cuartero (2012) in their analysis of relative clauses as involving a particular form of LDA.

(47) [_{ForceP} Force [_{TopP*} Top [_{FocP} Foc [_{TopP*} Top [_{FinP} Fin [...]]]]]]]

It is generally assumed that, in a left periphery hosting information-structural projections, the focus phrase is sandwiched between two topic phrases. Thus, the presence of a focus phrase always entails the presence of a higher topic phrase. Consequently, if a focus phrase counts a sufficiently local to the higher matrix *v*, the topic phrase dominating the focus phrase necessarily is local enough as well so that if foci can participate in LDA, so will topics.

LDA across more than one CP boundary is also expected in this approach. Nothing prevents intermediate topic heads from probing for a topicalized element in their *c*-command domain. For these heads, the initial topicalized DP in the lowest clause is not accessible since it is not local enough, i.e. separated by at least one CP boundary. However, they can agree with the topic head in the left periphery of the next lower clause. Since this topic head also hosts ϕ -features which have been valued by the topicalized DP in the lowest clause, the intermediate topic head can also value its ϕ -features, in effect transmitting the ϕ -features of the topicalized DP up until *v* in the highest clause.¹⁸

Note that this approach predicts that it is impossible for LDA to skip an intermediate clause. Thus, in a structure with three clauses, it is impossible for the verb in the highest clause to show LDA with the topic of the lowest clause if the verb in the intermediate clause does not show LDA as well. Since cross clausal LDA is successive cyclic as well, all intermediate topic heads need to participate in it, otherwise the ϕ -feature of the topic in the lowest clause cannot be accessed by higher verbs. As has been discussed above, this prediction is borne out in Hinuq (17) and cross-clausal LDA cannot skip an intermediate clause.

Lastly, the approach presented in this section is independent of the element in the matrix clause that establishes the actual agreement relation into the embedded clause. In the Nakh-Dagestanian and Algonquian languages, this element appears to be *v*, whereas in Uyghur it is a D head in the higher clause.

In sum, the present approach appears to be advantageous over the approaches of Polinsky & Potsdam (2001) as well as Bjorkman & Zeijlstra (2019) discussed above in that it easily derives the cross-linguistic properties of long distance agreement without any stipulations that cannot independently be motivated.

¹⁸I ignore the *v*P here. If *v*P is indeed a phase, then its periphery might host information structural projections as well which then participate in successive-cyclic agreement.

5 Conclusion

In this paper, I have presented a successive cyclic approach to long distance agreement that is based on the bundling of ϕ - and δ -features on heads in the left periphery of the embedded clause. Starting from an extensive typological discussion, I have shown that LDA can be found in various unrelated language families, namely in Nakh-Dagestanian, Algonquian and Altaic languages, albeit with slightly different properties. A generalization emerged from this discussion, namely that if a language permits LDA based on embedded foci it will also allow LDA based on embedded topics.

In the second part, after discussing previous analyses of LDA, I presented an approach that derives long distance agreement via successive cyclic agreement through the periphery of the embedded clause, thus analyzing LDA in accordance with the PIC, and similar to successive cyclic wh-movement. The nature of the intermediate head in the left periphery was then the main ingredient of the analysis in which I assumed a particular notion of feature inheritance. In addition to either transmitting ϕ - or δ -features from C to T, both can remain in C bundled together on one head. The valuation of one part of the feature bundle then depended on an agreement relation established based on the other part. More concretely, the head in the left periphery established an agreement relation with the relevant DP based on δ -features which allowed the ϕ -features of the left-peripheral head to be valued as well. For the second step in this successive cyclic agreement, the ϕ -features on this head then served as agreement goal for a higher probing head in the matrix clause, v or D. These assumptions, combined with a standard analysis of the left periphery, were then able to derive the cross-linguistic properties of LDA.

If this analysis is on the right track, more instances of information-structurally influenced ϕ -feature agreement are to be expected in various languages, see for example D'Alessandro (2020 [this volume]) for topic agreement in Ripano, Dalrymple & Nikolaeva (2011) for the connection between objects and information structure more generally, and (van der Wal 2015; 2020 [this volume]) for an overview of object marking in Bantu.

Abbreviations

Roman numerals refer to noun classes, Arabic numerals followed by *SG/PL* refer to person.

ABS	absolutive	F	feminine	PERF	perfect
ABST	abstract suffix	FOC	focus	PFV	perfective
ACC	accusative	FUT	future	POSS	possessive
AI	animate intrans.	GEN	genitive	PRET	preterite
AN	animate	HAB	habitual	PROX	proximal
AOR	aorist	IMPF	imperfective	PRS	present
APPL	applicative	INAN	inanimate	PST	past
AUG	augmentative	INV	inverse	PTCP	participle
C	complementizer	LAT	lative	PURP	purposive
CAUS	causative	LOC	locative	Q	question particle
CONJ	conjunct	N	neuter	QUOT	quotative
CVB	converb	NEG	negation	RECP	reciprocal
DAT	dative	NOM	nominative	RES	resultative
DEM	demonstrative	OBL	oblique	SM	subject marker
DIR	directional	OBV	obviative	TA	trans. animate
ERG	ergative	OM	object marker	TI	trans. inanimate
EMPH	emphatic particle	PL	plural	TOP	topic
EVID	evidential	PART	participle agr.	UWPST	unwitnessed pst

Acknowledgements

The topic was (in part) presented at the IGG 2017, at Goethe University Frankfurt and the EGG 2016, and I thank the audiences for their comments and suggestions. The research has also benefited greatly from discussions with Katharina Hartmann, Peter Smith, Jennifer Tan, Hedde Zeijlstra and Jonathan Bobaljik, as well as the comments and suggestions of two reviewers. All shortcomings remain my own.

References

- Asarina, Alya & Jeremy Hartman. 2011a. Genitive subject licensing in Uyghur subordinate clauses. In Andrew Simpson (ed.), *Proceedings of the 7th Workshop on Altaic Formal Linguistic (WAFL)*, 17–31. Cambridge, MA: MIT Working Papers in Linguistics.
- Asarina, Alya & Jeremy Hartman. 2011b. Uyghur genitive subjects and the phase impenetrability condition. Talk at CUNY Syntax supper.
- Baker, Mark C. 2008. *The syntax of agreement and concord*. Cambridge: Cambridge University Press.
- Baker, Mark C. & Jonathan D. Bobaljik. 2015. On inherent and dependent theories of ergative case. In Jessica Coon, Diane Massam & Lisa Travis (eds.), *Handbook of ergativity*, 111–135. Oxford: Oxford University Press.
- Bárány, András. 2015. *Differential object marking in Hungarian and the morphosyntax of case and agreement*. University of Cambridge. (Doctoral dissertation).
- 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. 2002. A-Chains at the PF-interface: Copies and covert movement. *Natural Language & Linguistic Theory* 20(2). 197–267.
- Bobaljik, Jonathan D. 2008. Where's phi? Agreement as a post-syntactic operation. In Daniel Harbour, David Adger & Susana Béjar (eds.), *Phi-Theory: Phi features across interfaces and modules*, 295–328. Oxford: Oxford University Press.
- Börjesson, Kristin & 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
- Bošković, Željko. 2007a. Agree, phases, and intervention effects. *Linguistic Analysis* 33(1–2). 54–96.
- Bošković, Željko. 2007b. On the locality and motivation of Move and Agree: An even more minimal theory. *Linguistic Inquiry* 38(4). 589–644.

- 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).
- 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.
- D’Alessandro, Roberta. 2020. Agreement across the board: Topic agreement in Ripano. In Peter W. Smith, Johannes Mursell & Katharina Hartmann (eds.), *Agree to Agree: Agreement in the Minimalist Programme*, 235–270. Berlin: Language Science Press. DOI:10.5281/zenodo.3541757
- Dalrymple, Mary & Irina Nikolaeva. 2011. *Objects and information structure*. Cambridge: Cambridge University Press.
- Forker, Diana. 2012. *A grammar of Hinuq*. Berlin: de Gruyter.
- Frank, Robert. 2006. Phase theory and tree adjoining grammar. *Lingua* 116(2). 145–202.
- Frantz, Donald G. 1978. Copying from complements in Blackfoot. In Eung-Do Cook & Johnathan Kaye (eds.), *Linguistic studies of Native Canada*, 89–110. Vancouver: University of British Columbia Press.
- Fry, Brandon J. & Michael David Hamilton. 2014. Long-distance agreement in Algonquian: Accounting for syntactic variation. In Ross Burkholder, Carlos Cisneros & Emily R. Coppess (eds.), *Proceedings of the 50th annual meeting of the Chicago Linguistics Society (CLS 50)*, 159–173. Chicago: Chicago Linguistic Society.
- Hale, Kenneth. 2002. On the Dagur object relative: Some comparative notes. *Journal of East Asian Linguistics* 11(2). 109–122.
- Haspelmath, Martin. 1999. Long distance agreement in Godoberi (Daghestanian) complement clauses. *Folia Linguistica* 33(1–2). 131–151.
- Heck, Fabian & Juan Cuartero. 2012. Long distance agreement in relative clauses. In Artemis Alexiadou, Tibor Kiss & Gereon Müller (eds.), *Local modelling of non-local dependencies in syntax*, 49–85. Berlin: De Gruyter.
- Jiménez-Fernández, Ángel L. 2010. Discourse-agreement features, phasal C and the edge: A minimalist approach. *Diacritica* 24(1). 25–49.

- Khalilova, Zaira. 2008. 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.
- Kornfilt, Jaklin. 2008. Subject case and Agr in to types of Turkish RCs. In Cedric Boeckx & Suleyman Ulutas (eds.), *Proceedings of the 4th Workshop on Altaic Formal Linguistic (W AFL)*, 145–168. Cambridge, MA: MIT Working Papers in Linguistics.
- Ledgeway, Adam. 2005. Moving through the left periphery: The dual complementiser system in the dialects of Southern Italy. *Transactions of the Philological Society* 103(3). 339–396.
- 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.
- Miyagawa, Shigeru. 2010. *Why agree? Why move?: Unifying agreement-based and discourse configurational languages*. Cambridge, MA: MIT Press.
- Miyagawa, Shigeru. 2011. Genitive subjects in Altaic and specification of phase. *Lingua* 121. 1265–1282.
- Miyagawa, Shigeru. 2017. *Agreement beyond phi*. Cambridge, MA: MIT Press.
- Pesetsky, David & Esther Torrego. 2007. The syntax of valuation and the interpretability of features. In Simin Karimi, Vida Samiiian & Wendy K. Wilkins (eds.), *Phrasal and clausal architecture: Syntactic derivation and interpretation*, 262–294. Amsterdam: John Benjamins.
- Polinsky, Maria. 2003. Non-canonical agreement is canonical. *Transactions of the philological society* 101(2). 279–312.
- 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 & Maria Polinsky. 2015. Agreement and semantic concord: A spurious unification. <http://ling.auf.net/lingbuzz/002363>. Ms, University of Maryland.
- Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of grammar: A handbook in generative syntax*, 281–337. Dordrecht: Kluwer.
- Şener, Serkan. 2008. Non-canonical case licensing is canonical: Accusative subjects of CPs in Turkish. Ms, University of Connecticut.

- van der Wal, Jenneke. 2015. Bantu object clitics as defective goals. *Revue Roumaine de Linguistique* 60(2-3). 277–296.
- van der Wal, Jenneke. 2020. The AWSOM correlation in comparative Bantu object marking. In Peter W. Smith, Johannes Mursell & Katharina Hartmann (eds.), *Agree to Agree: Agreement in the Minimalist Programme*, 199–234. Berlin: Language Science Press. DOI:10.5281/zenodo.3541755
- Wurmbrand, Susi. 2001. *Infinitives: Restructuring and clause structure*. Berlin: Mouton De Gruyter.
- Wurmbrand, Susi. 2012. Agree(ment): Looking up or looking down? Ms, University of Connecticut.
- Wurmbrand, Susi. 2016a. Feature dependencies: The ups and downs of agree. Lecture notes, University of Connecticut.
- Wurmbrand, Susi. 2016b. The ups and downs of agree. Lecture notes, University of Connecticut.
- Zaenen, Annie, Joan Maling & Höskuldur Thráinsson. 1985. Case and grammatical functions: The Icelandic passive. *Natural Language & Linguistic Theory* 3(4). 441–483.

