Ke, Alan Hezao. 2023. Search downward: Minimal Search-based Agree. *Glossa: a journal of general linguistics* 8(1). pp. 1–38. DOI: https://doi.org/10.16995/glossa.6526

GLOSSA
a journal of general linguistics

Open Library of Humanities

Search downward: Minimal Search-based Agree

Alan Hezao Ke, Michigan State University, Department of Linguistics, Languages, and Cultures, US, kehezao@msu.edu

This paper explores empirical merits of a version of Agree that is defined based on Minimal Search (MS-Agree). Compared to the standard Agree, MS-Agree, essentially a search algorithm, uniquely allows the independent assignment of its search target and search domain. This unique feature enables MS-Agree to accommodate both upward and downward agreement phenomena, and offers a unified downward search analysis for negative concord, inflection doubling, multiple case-assignment, cyclic agreement, and complementizer agreement observed across languages. This paper thus argues that these core empirical data that have served as the main motivation for Upward Agree can be successfully reanalyzed with MS-Agree. It is also argued that the proposed MS-Agree analysis makes better predictions than Upward Agree regarding intervention effects in apparent upward agreement phenomena.

1 Introduction

Being considered a core syntactic operation in syntax, Agree gains significant attention from the field, along with debates on its essential aspects. This paper engages with the debate on the direction of search involved in Agree, based on the influential assumption that Agree is a syntactic operation that is implemented by a search algorithm (i.e., Minimal Search, see Chomsky 2013; Ke 2019; and Ke to appear(a) and citations therein). A unified analysis of both downward and upward agreement will be proposed, with Minimal Search into either (i) the sister of the probing head for downward agreement or (ii) the next built phase above the probing head for (apparent) upward agreement (cf. Béjar & Rezac 2009; Carstens 2016).[1] This paper then argues that the MS-Agree analysis makes better predictions than Upward Agree regarding intervention effects in upward agreement phenomena. The interaction between this analysis and the theory of phases (phasal Transfer in particular) and Labeling will also be discussed.

Chomsky (2000) provides a formal definition of Agree which brings forth the concept of probing[2]:

(1) *Definition of Agree*

Agree is a syntactic operation taking place between a probe P and a goal G in the domain of P, D(P), between which a Matching relation holds.

a. Matching is identity of feature attributes;
b. D(P) is the sister of P;
c. Locality reduces to "closest c-command";

(2) *Definition of closest c-command*

A matching feature G is closest to P if there is no G′ in D(P) matching P s.t. G is in D(G′).

The concept of probing is intriguing as it seems to be related to the application of search in Agree later in Chomsky (2013). (1b) indicates that the domain of probing is determined by sisterhood relations. The definition of "closest c-command" in (2) is then based on this conception of domain. Although this definition of Agree does not explicitly mention another important structural aspect besides sisterhood, namely, (immediate) containment, the phrases "*in* D(P)" and "*in* D(G′)" imply such containment relations. An equivalent (derivational) way to state this

---

[1] In this paper, "downward," "upward," and "cyclic agreement" are used as descriptive terms referring to agreement patterns, not analyses of these phenomena. Terms in capital, e.g., "Agree," "Upward Agree," and "Cyclic Agree" instead represent certain syntactic analyses or operations.

[2] Chomsky (2000) considers matching as feature identity. This conception of matching may be incorrectly interpreted as identity of both feature attributes and feature values. (See Adger 2010 for a theory of features where a feature is considered an ordered pair of feature attributes and feature values.) I revise this part of definition as "matching is identity of feature attributes," which does not involve feature values, due to Chomsky's (2004) re-conception of matching.

definition is that Agree is an operation that probes into a syntactic domain based on sisterhood and containment relations: the probe looks into the domain, its sister/co-member, to find matching feature attributes, then looks into the members that its sister immediately contains, and then the members of the members of its sister; so on and so forth.[3] Probing must be terminated whenever a goal is returned to ensure that "closest c-command" is conformed to.

Crucially, this definition implies that the probe (uninterpretable features) always searches for a goal (interpretable features) inside its sister. That is, the probe always probes downward into its c-command domain, which is a direct consequence of the interaction of the sisterhood and containment relations encoded in the definition. This formalization of the direction of probing in Agree is controversial and has recently produced much discussion.

However, there are important empirical challenges to the direction of probing in Agree: Agree is not always probing into the domain of the probe-bearer. The central argument against the probing-based Agree approach comes from studies on negative concord (Zeijlstra 2012) and inflection doubling (or parasitic participles) (Wurmbrand 2012a,b). Additional evidence comes from multiple case licensing (or Multiple Agree) in Japanese. Zeijlstra (2012) and Bjorkman & Zeijlstra (2019) argue that negative concord, inflection doubling, and multiple case licensing are all morphosyntactic phenomena, and they are subject to syntactic locality relations such as clause boundedness, similar to subject–verb agreement. They further assume that feature agreement is the mechanism behind all of these morphosyntactic phenomena. An Upward Agree-based analysis is proposed by Zeijlstra (2012) and Bjorkman & Zeijlstra (2019) to account for all these upward (and downward) agreement phenomena.

This paper reconsiders the core empirical data that have directly motivated the Upward Agree operation and will argue that these data can be accounted for without resorting to Upward Agree. These cases are all compatible with a downward search algorithm that is independently developed for other reasons. The idea then is, even if we are empathetic to Bjorkman & Zeijlstra's (2019) viewpoint that negative concord and other phenomena that serve as the basis of their theory are agreement phenomena in nature, bypassing a potential opposing view that φ-feature agreement and concord are distinctive phenomena such that they should not be unified under the same analysis, we still find a downward search/probing-based analysis more preferable both theoretically and empirically. I will then present additional data where upward agreement appears to occur, in particular cyclic agreement in Georgian (e.g., Béjar & Rezac 2009) and subject–complementizer agreement in Lubukusu (e.g., Diercks 2013; Carstens 2016), to demonstrate that these data can also be neatly covered by such a downward search-based analysis, if Minimal Search for Agree can take the next built phase above the probing head, in addition to its sister, as the search domain. Specifically, Minimal Search always starts from the sister of the probing head;

[3] See Epstein et al. (1998) for such a derivational view of c-command.

only when this search fails, Minimal Search will search down into the next built phase above the probing head. Such a downward search-based analysis makes critical predictions regarding the absence and presence of intervention effects in agreement, providing an essential diagnostic to distinguish downward search from upward search.

In the next section, I begin by introducing Minimal Search for Agree and Labeling developed in Ke (2019) and Ke (to appear[a]), which forms the basis for the re-analysis of the data that has been used to argue for Upwards Agree and Cyclic Agree.

2 Minimal Search and Agree

Ke (2019, to appear[a]) proposes a search algorithm, i.e., Minimal Search, for both Agree and Labeling and then develops a definition of Agree based on Minimal Search. As shown in (3), the definition of Minimal Search includes a search algorithm that applies iteratively to a search domain (SD) to look for a search target (ST).[4]

(3) *Definition of Minimal Search* (Ke 2019: p. 44)
MS = <SA, SD, ST>, where MS = Minimal Search, SA = search algorithm, SD = search domain (the domain that SA operates on), ST = search target (the features that SA looks for).
Search Algorithm (SA):
a. Given SD and ST, match against every head member of SD to find ST.
b. If ST is found, return the heads bearing ST and go to Step (c); Otherwise, get the set members of SD and store them as a list L.
   i. If L is empty, search fails and go to Step (c); otherwise
   ii. assign each of the sets in L as a new SD and go to Step (a) for all these new SDs in parallel.
c. Terminate search.

Readers are referred to Ke (to appear[a]) for theoretical and empirical considerations behind the following properties of this definition of Minimal Search (MS): why the search algorithm underlying both Agree and Labeling should be breadth-first rather than depth-first search, parallel rather than serial, global rather than modular.

As an illustration of the MS defined in (3), let us walk through an example. In **Figure 1**, an MS is initiated to search for an ST = feature [F], in the SD = set α. Three runs of search are conducted before the target is found. In the first run, MS looks into α. X, the head member of α, is returned by the first run of search. However, since it does not bear the ST [F], and its sister is

[4] MS as defined in Ke (2019, to appear[a]) cannot be straightforwardly applied to internal or external Merge, because MS requires an ST, which is not specified in the case of Merge. That is, the syntax system does not necessarily search for particular types of syntactic objects to which Merge applies.

a set, the search is not terminated. The set member of α, that is, the set β, is then assigned as a new SD.[5] This is the second run. The second run of search finds only set members, i.e., γ and κ, and these two set members are stored as a list L. The two sets in L are then assigned as the new SDs for the third run of search, and two independent searches are initiated in parallel. The head $K_{[F]}$ is found inside the set κ as it bears the ST. $K_{[F]}$ is finally returned by the search algorithm. The search over the set γ fails and is terminated without returning a result.

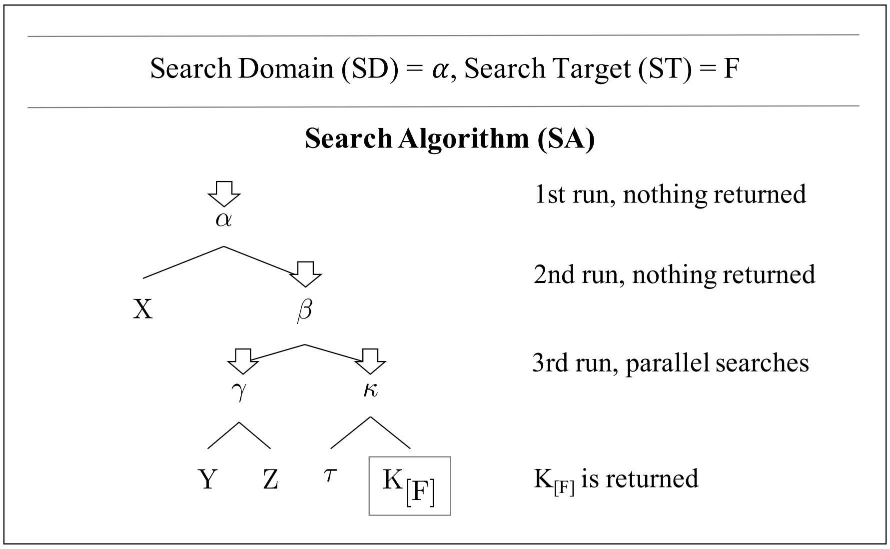


Figure 1: An illustration of Minimal Search.

As shown in (3), MS involves three components, the SD, ST, and SA. The SA is an independent, parallel, breadth-first, and global search algorithm. Note that the values of SD and ST are assigned by MS-based Agree, a unique feature (compared to the standard definition of Agree by Chomsky 2000) that has far-reaching empirical consequences, as I will argue in the rest of this paper. The mechanism for SD and ST assignment will be discussed in detail in Section 2.2 as well as Section 3 and 4. It will be shown that this unique feature of MS-based Agree does not necessarily give rise to a non-predictive theory while ensuring a wide empirical coverage.

2.1 Minimal Search-based Agree

Based on the above definition of MS, which is a general algorithm incorporated in both Agree and Labeling, Ke (2019, to appear[a]) gives an MS-based definition for Agree (MS-Agree):

[5] Note that the MS defined in (3) does not allow a single MS applying to various different sets. In order to search into a different set, a new MS must be initiated taking that set as a new SD.

(4) *Definition: Minimal Search-based Agree (MS-Agree)* (Ke 2019:52–53)

Agree = Minimal Search + valuation

i. **Minimal Search**
   a. Input to Minimal Search:
      - SD = sister of the head bearing unvalued features $H_{[uF]}$ (to be revised)
      - ST = feature attributes of unvalued feature uFs (e.g., [Person:, Num:])
   b. Output of MS: heads bearing the ST

ii. **Valuation**
   - Trigger: the head bearing uFs
   - Valuation: copy the value of the ST in the search output to the corresponding uFs in the trigger

Below, I give an example to illustrate the implementation of MS-Agree. *Be* in (5) can agree with *a man*, or more accurately, the determiner *a*, the head of *a man*. This is a case of long-distance agreement.[6] The steps of MS-Agree are shown in (6).

(5) Example: *(there) be likely to be a man outside*

(6)
a. SD = {likely, {to be a man outside}}
b. ST = ϕ-feature attributes of the unvalued ϕ-features on *be*
c. Search into the SD, the target is not found
d. Set {to, {be a man outside}} as the SD
e. Search into the SD, the target is not found
f. Set {be, {a man outside}} as the SD
g. Search into the SD, the target is not found
h. Set {{a man}, outside} as the SD
i. Search into the SD, the target is not found
j. Set {a, {man}} as the SD
k. Search into the SD, the target is found on *a*, return *a*
l. Valuation: copy feature values to *be* from the matching features on *a*

A few notes should in place before we move on. (i) MS as a general SA defined in (3) always searches downward, and never searches upward. (ii) MS-Agree is split into two components, MS and valuation. This split can also be found in a broad ranges of previous studies, including Arregi & Nevins (2012), Bhatt & Walkow (2013), Deal (2015), Smith (2017), Kučerová (2018), and Bjorkman & Zeijlstra (2019). It is worth mentioning that this split is adopted here not due to a stipulation. Instead, it is a natural consequence of the MS defined in (3), as there is no

[6] When *be* is connected to a plural associate, e.g., in 'there {be} many opportunities,' *be* can be inflected as singular or plural; whereas the plural version seems to be a result of agreement, I do not have an account as to where the singular version comes from.

principled way for a search algorithm to do the job of valuation. (iii) MS-Agree determines which syntactic object will be returned by the search algorithm, but it does not determine which syntactic object will undergo movement to the specifier position of the probing head (Bobaljik & Wurmbrand 2005), consistent with Chomsky's (2015) Labeling-based account of movement. (iv) The SD and ST in MS for Agree are independently assigned by Agree (Ke to appear(a)), and this is a unique feature of the definition of MS. Previous definitions of Agree do not provide this option of assigning the values of the SD and ST; however, this is now a natural consequence of the SA: an SA must have the information of what to search for and which domain to search into, although the ST and SD may not be explicitly defined in a relevant theory. The ST is the unvalued feature (uFs) on a head that triggers MS. MS-Agree then is to look for a head with the ST inside the SD. Crucially, it is this unique feature that the SD and ST are independently assigned by Agree, together with other assumptions, enables us to analyze apparent upward agreement as downward MS-based Agree. Below I would like to explain how the SD and ST are determined for MS-Agree. In brief, I assume that the sister of the trigger will by default be taken as the SD. When the search into this SD fails, then the next built phase above the trigger could be taken as the new SD, a pattern that has been observed in cyclic agreement (Béjar & Rezac 2009) and the Delayed Valuation analysis (Carstens 2016).

2.2 How to determine SD and ST for MS-Agree?

The SD of MS for Agree is determined in accordance with the purpose of Agree. Agree connects two heads, one carrying unvalued features and the other carrying matching valued features. In order for two heads in an agreement relation to be connected to each other syntactically, they must form a syntactic relation at a certain stage of derivation. That is, they must be contained in the same root set or under a single node in a syntactic tree at a certain stage of derivation. Given a head X with an unvalued feature uF, represented as $X_{[uF]}$, in order to value the uF on X by another head Y with a corresponding valued feature vF, represented as $Y_{[vF]}$, there are three logically possible situations for X to be connected to Y at a certain point of the derivation, with regard to the c-command relations between X and Y:

(7) a. $X_{[uF]}$ merges with $Y_{[vF]}$ or a set containing $Y_{[vF]}$: $X_{[uF]}$ c-commands $Y_{[vF]}$;
b. $Y_{[vF]}$ merges with $X_{[uF]}$ or a set containing $X_{[uF]}$: $Y_{[vF]}$ c-commands $X_{[uF]}$;
c. A set α containing $Y_{[vF]}$ merges with a set β containing $X_{[uF]}$. Neither $X_{[uF]}$ nor $Y_{[vF]}$ c-commands the other.

All three options are utilized in natural language. Downward (long-distance) Agree, e.g., where the matrix verb agrees with an embedded nominal for noun classes in Tsez (Polinsky & Potsdam 2001) is an instance of (7a). Upward Agree in negative concord (Zeijlstra 2012), inflection doubling/parasitic participles (Wurmbrand 2012a) and multiple case licensing (Hiraiwa 2001) are instances of (7b). And Ke (2019, ch. 3) argues that reflexive binding is an instance of (7c).[7] Due to space limitations, we will set (7c) aside for the discussion below.

Regular downward Agree in (7a) such as long-distance Agree in Tsez observed by Polinsky & Potsdam (2001) composes a strong argument for downward Agree. MS-Agree, building upon a downward search algorithm, can easily implement downward agreement. The searching of the target which is buried deep in the root structure can be accomplished by MS-Agree in a way similar to what we have seen in the *there*-construction in (5). What is potentially challenging for MS-Agree are cases of upward agreement in (7b) that provide crucial support for an Upward Agree analysis. Below I will provide an analysis of apparent upward agreement as a special case of downward MS-Agree.

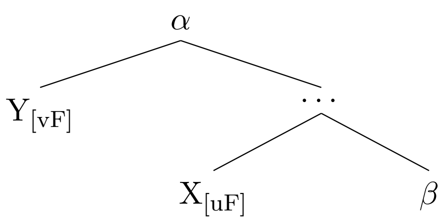
3 Upward agreement as downward MS-Agree

The main goal of this section is to review critical cases where apparent upward agreement are observed, and reanalyze upward agreement as MS-Agree, thus addressing the debate on the direction of probing in Agree. I argue that, based on a unique feature of MS-Agree, namely, the ST and SD are independently assigned by Agree, the SD can be the next built phase above the trigger (cf. Carstens's 2016 Delayed Valuation analysis; see also Bobaljik & Wurmbrand 2005); consequently, cases of apparent upward agreement can be implemented as instances of MS-Agree.

3.1 Empirical challenges to downward search Agree

We have mentioned that there are three logically possible configurations in which a head X with an unvalued feature can be related to and agree with another head Y with a corresponding valued feature, as listed in (7), and the focus of this paper would be the second configuration based on which Upward Agree is proposed. The structure of the second configuration is shown in (8) (The dots in the tree diagram indicate optional additional structure):

(8) $Y_{[vF]}$ merges with $X_{[uF]}$ or a set containing $X_{[uF]}$: $Y_{[vF]}$ c-commands $X_{[uF]}$;



[7] To give an (oversimplified) sketch of the proposal, in an example such as "this$_{[vF]}$ woman loves herself$_{[uF]}$," if projection is not assumed, and if reflexive binding is formalized as Agree (Reuland 2001; Hicks 2009, among many others), this$_{[vF]}$ does not c-command *herself*$_{[uF]}$ but can agree with it via MS-Agree, through a mechanism similar to what we will see in second cycle MS-Agree in Section 4.

The upward agreement configuration in (8) does raise a problem to the classical probing-based Agree operation. We have mentioned that negative concord, inflection doubling/parasitic participle, and Multiple Agree as instances of the upward agreement configuration in (8) provide strong evidence for the existence of Upward Agree. To illustrate, below I quote one example from each of these constructions, which is coupled with a corresponding syntactic analysis.[8] Let us start with negative concord in Czech.

(9) *Negative concord in Czech* (Zeijlstra 2012: p. 501)

a. Dnes nikdo *(ne)volá nikomu
Today n-body NEG.calls n-body
'Today nobody is calling anybody.'

b. *Two possible structural analyses of (9)*
Analysis 1: [$_{\text{CP}}$ [$_{\text{TP}}$ nikdo$_{\text{[uNEG]}j}$ [$_{\text{NegP}}$ *ne*volá$_{\text{[iNEG]}}$ t$_j$ nikomu$_{\text{[uNEG]}}$]]]
Analysis 2: [$_{\text{CP}}$ [$_{\text{TP}}$ Op¬$_{\text{[vNEG]}}$ [$_{\text{TP}}$ nikdo$_{\text{[uNEG]}j}$ *ne*volá$_{\text{[uNEG]}}$ [$_{v\text{P}}$ t$_j$ nikomu$_{\text{[uNEG]}}$]]]]

Zeijlstra (2004) assumes that n-words must be licensed by or agree with an element bearing an [iNEG] feature. This element could be the negative marker *ne* (the first analysis in (9b)) or a null negative operator *Op* (the second analysis in (9b)), the latter of which bears an [iNEG] feature and is projected at a position higher than TP. If the first analysis in (9b) is adopted, Upward Agree is needed to value the unvalued feature on the object *nikomu*, as the object is c-commanded by, not c-commands, *ne* (or *nevolá*), which bears an [iNEG] feature. On the other hand, if we adopt the second analysis in (9b), all the syntactic objects bearing the [uNEG] are c-commanded by the null negative operator with [iNEG]. Classical probing-based Agree as defined in (1) cannot be applied to check these [uNEG] features, thus calling for an alternative analysis, e.g., Upward Agree.

Zeijlstra (2012) defines Upward Agree as in (10).

(10) *Definition of Upward Agree* (Zeijlstra 2012: p. 514)
α agrees with β iff:

a. α carries at least one uninterpretable feature and β carries a matching interpretable feature;
b. β c-commands α;
c. β is the closest goal to α.

I would like to briefly note that according to (10c), in the configuration where there are two c-commanding syntactic objects carrying matching interpretable features, α will agree with the

[8] In example (9), I have replaced interpretable features in the original representations, e.g., [iNEG], with corresponding valued features, e.g., [vNEG], to match the definition of MS in (3); nothing important hinges on these changes. In addition, to simply the discussion, I have abstracted away from the actual valued and unvalued features involved in (11) and (12).

one closer to it, causing an intervention/blocking effect to the structurally higher candidate. I will return to this point in Section 5 and show that this leads to incorrect empirical predictions.

Inflection doubling (parasitic participles) also calls for an analysis along the lines of Upward Agree. Wurmbrand (2012a) encodes morphological selection (e.g., a modal selects an infinitive verb) as feature agreement. She calls this type of feature agreement "Reverse Agree," which shares with Zeijlstra's (2012) Upward Agree analysis in its core assumption of upward probing direction. In the following example from Swedish, the parasitic participles are in boldface, and the auxiliaries that license the parasitic participles are underlined.

(11) *Inflection doubling in Swedish* (Wiklund 2001: p. 200)

a. Jag <u>hade</u> velat [**last** / läsa boken].
I had want.PART **read.PART** / read.INF book.DEF
'I would have liked to read the book.'

b. [$_{CP}$ … [$_{AuxP}$ <u>had</u>$_{[vF]}$ want.PART$_{[uF]}$ [$_{VP}$ **read.PART**$_{[uF]}$ …]]]

Wurmbrand (2012a) assumes that semantically vacuous parasitic participles carry an uninterpretable feature. Since a c-commanding auxiliary can function as a licenser, Wurmbrand (2012a) suggests that the auxiliary carries a corresponding interpretable feature that agrees with the parasitic participle it licenses. This is hence another case to which Reverse Agree/Upward Agree does provide an appealing analysis.

Let us now consider another case which has been previously suggested to be an instance of Upward Agree: multiple case licensing in Japanese. (12) illustrates that in Japanese raising-to-subject constructions, multiple DPs can be marked as nominative within an infinitival embedded clause.

(12) *Multiple case licensing in Japanese* (Hiraiwa 2001: p. 76)

a. John-ga [yosouijouni nihonjin-ga eigo-ga hidoku] kanji-ta.
John.NOM than-expected the-Japanese.NOM English.NOM bad.INF think.PST
'It seemed to John that the Japanese are worse at speaking English than he had expected.'

b. [$_{CP}$ … DP1.NOM$_{[uF]}$ [… DP2.NOM$_{[uF]}$ … DP3.NOM$_{[uF]}$ … V.INF] T$_{[vF]}$]

If we assume that (i) nominative case assignment from the T head to the DPs in (12) is accomplished via agreement, (ii) T bears a valued [vCase] feature that is responsible for case assignment, and (iii) all DPs carrying unvalued [uCase] features must be valued by agreeing with T, (12) presents a problem to the standard probing-based Agree. The problem is that the DPs with the unvalued [uCase] feature do not c-command the T head which bears the valued [vCase] feature. To address this problem, one must stipulate that there is some unvalued feature on T, other than the case feature, that triggers the downward search aiming for DPs bearing corresponding valued features, and then case assignment is just a free rider of the agreement relation that has been established. One such choice is to assume that the unvalued ϕ-features on T initiate the search for DPs with corresponding valued ϕ-features, and agreement and case are just two sides of the same coin (e.g., Chomsky 1995). However, this analysis encounters a different problem. If Agree is defined as in (1), we would expect that probing will be terminated as soon as a DP with matching valued ϕ-features is found (due to the "closest c-command" requirement). The consequence is that T in (12) will probe for and agree with the subject DP only (as in the instances where only the subject acquires the nominative case), and therefore can assign case to that DP only. In other words, an intervention effect caused by the subject DP is predicted. Contrary to this prediction, no such intervention effects are observed in (12), because all the DPs are assigned case by the T head. As a result, Hiraiwa (2001) proposes the operation MULTIPLE AGREE, as defined in (13), to account for the absence of intervention effects in (12).

(13) *Multiple Agree* (Hiraiwa 2001: p. 69)
Multiple Agree with a single probe is a single simultaneous syntactic operation; Agree applies to all the matched goals at the same derivational point derivationally simultaneously.

The Multiple Agree analysis assumes that the T head in multiple nominative case assignment bears a [+multiple] feature, and that this feature enables the probe to look for multiple matching goals that are locally available (see also Chomsky 2008: p. 142). Such an analysis therefore requires the T head to be ambiguous between either bearing a [+multiple] feature or not, depending on the construction.

To some extent, this approach, while empirically successful, requires ad hoc assumptions and cannot explain why Multiple Agree occurs. In addition, Multiple Agree introduces the concept of "simultaneity" in Agree, which is fundamentally incompatible with the idea of Minimal Search or probing that respect "closest c-command." In contrast, an Upward Agree analysis of (12) does not need to assume simultaneity in Agree or to stipulate a T that bears a [+multiple] feature. This is because each DP carrying an unvalued case feature can independently search upward and agree with the valued case feature on T. No intervention effects are expected, in accordance with the data. Therefore, an Upward Agree seems simpler and less stipulative than the Multiple Agree analysis, as far as the multiple case-assignment in (12) is concerned.

In sum, we see that Upward Agree solves significant empirical problems faced by the classical probing-based Agree analysis. It seems, Agree, at least in the cases discussed, should probe upward.

Aiming for an unified analysis of both upward and downward agreement, Bjorkman & Zeijlstra (2019) revise and extend this approach to regular long-distance downward agreement. The main modification is to split agreement to two components, checking via Upward Agree and feature valuation. Checking is exactly Upward Agree as defined previously by Zeijlstra (2012) shown in

(10) above. It is to establish an accessibility relation between a c-commanding syntactic object with interpretable features and a c-commanded syntactic object with uninterpretable features. Accessibility established by Upward Agree is then a precondition for feature valuation, according to what I will call "Accessibility Condition":

(14) *Accessibility Condition* (Bjorkman & Zeijlstra 2019: p. 536)
α and β are accessible to each other iff an uninterpretable feature on β has been checked (via feature checking) by a corresponding interpretable feature on α.

The Accessibility Condition in (14) suggests that as long as *any* uninterpretable feature of a syntactic object checks with a corresponding interpretable feature of another syntactic object, which puts these two syntactic objects in a checking relation, *all* the features on them are made accessible to each other for valuation purposes. This applies to all interpretable and uninterpretable features (as long as they have already been checked; see below).

However, the Accessibility Condition needs to work together with another hypothesis[9]:

(15) Only checked features can be valued.

Consequently, when a syntactic object α bearing uninterpretable feature uF1 checks with β bearing corresponding interpretable feature iF1, all the other features on α, for example, an interpretable feature iF2, could value corresponding uninterpretable features, uF2, on β, as long as uF2 have been checked. Bjorkman & Zeijlstra (2019) further allow the valuation relations to be disassociated from the checking relations: if a uF on α fails to be valued by an iF on a c-commanding syntactic object β, it can be valued instead by a uF on another syntactic object γ that is c-commanded by α, as long as (14) and (15) are met. This theoretical design indicates that Bjorkman & Zeijlstra's (2019) approach allows both downward and upward valuation (Preminger & Polinsky 2015). The mechanism that unifies various types of agreement, as claimed by Bjorkman & Zeijlstra (2019), is therefore not the valuation component, but the checking (via upward probing) component.

As noted, Bjorkman & Zeijlstra's (2019) analysis has yielded attractive analyses of various phenomena and is certain to inspire valuable additional work. However, Preminger (2013) and Preminger & Polinsky (2015) have discussed at length that this approach as a hybrid approach that resorts to both downward valuation and upward valuation is not successful in accounting for ϕ-agreement in general. Recently, Rudnev (2021) also provides empirical arguments against various predictions of Bjorkman & Zeijlstra's (2019) analysis of ϕ-agreement in ergative-absolutive languages. I will not restate these counter-arguments here. The focus of this paper is instead to

[9] Bjorkman & Zeijlstra (2019) assume that (15) is in particular a restriction only on uninterpretable features and does not apply to unvalued interpretable features, as shown by their analysis of the long-distance agreement in Tsez (pp. 556–558). Note that this is different from the Accessibility Condition in (14) which applies to all types of features.

discuss a critical component of Bjorkman & Zeijlstra's (2019) analysis that was rarely put under careful scrutiny: the probing direction. It is also worth noting that the direction of probing/checking is probably even more crucial than the direction of valuation, from the perspective of the system itself. This is because as long as checking relations could be established, valuation is simply making use of these checking relations, whatever direction of "feature transmission" it is. In fact, feature valuation is just feature copying between parts of a syntactic structure: feature valuation can function without the involvement of any concept of syntactic structural relations such as c-command relations. This is also consistent with the idea that feature valuation could be a non-syntactic operation or could occur not in syntax; instead, it could be in the postsyntactic morphology, a proposal by Arregi & Nevins (2012) that Bjorkman & Zeijlstra (2019) parallel their approach to in terms of the valuation component. Indeed, the original empirical data that have motivated Bjorkman & Zeijlstra's (2019) system is exactly the checking/Upward Agree component, not the valuation component, as previously reviewed.

3.2 MS-Agree analysis of apparent upward agreement

For inflection doubling and multiple case licensing, as indicated by the syntactic representations in the (b) part of examples in (11) and (12), these attested cases of Upward Agree are all phase-constrained. That is, the same minimal phase (the CPs in the examples) dominates both the valuee carrying the unvalued features and the valuer bearing the corresponding valued features. This gives us a hint on how to model these cases of Upward Agree with MS-Agree. As a reminder, the ST and SD for MS-Agree can be independently assigned. The ST will be the feature attributes of any features that trigger MS; for example, they could be ϕ-features for ϕ-agreement, Q-feature for questions, and Neg-feature for negative concord. The SD is usually the sister/co-member of the trigger. What are other possible SDs that can be determined in a principled way? Due to the fact that all the Upward Agree cases listed in (9) to (12) are generally phase-constrained, if MS can have the CP phase as the SD and searches downward from it, then we obtain the desired results.

For the negative concord in Czech (9a), with the first alternative analysis in (9b) provided by Zeijlstra (2012), only the object *nikomu* is involved in upward agreement. Depending on whether NegP is a part of the *v*P phase, the reanalysis of this case of upward agreement with MS-Agree will be slightly different. If NegP belongs to the *v*P phase (cf. Akahane 2006), the phase immediately above the object is NegP. To value the [uNeg] feature on the object *nikomu*, MS-Agree will take NegP as the SD and search down from there. The negative marker *ne*, bearing a matching [vNeg] feature, will be correctly returned. On the other hand, if NegP is assumed to be part of the CP phase (cf. Butler 2004), MS-Agree can be initiated when the *v*P phase is built but it fails to return a target. The derivation may proceed until the CP phase is built, which is allowed under Chomsky's (2001) theory of Transfer: Transfer of the current phase occurs after a higher phase is built. MS for Agree purposes can be initiated as long as unvalued features are not transferred, and the timing of Transfer is determined by Chomsky's (2001) theory of Transfer (see also Footnote 13). Then another instance of MS with the CP as its SD will be initiated, returning the negative marker *ne*, as desired.[10]

With the second alternative analysis in (9b), the [uNeg] feature on the subject and the negative marker *ne* can be valued by conducting two independent searches probing downward from the CP phase. For the [uNeg] feature on the object, an analysis along the same lines as that for the object under the first alternative analysis can be provided: an MS taking the CP phase as the SD will find the negative operator below CP and successfully establish an agreement relation between the object and the negative operator.[11]

More generally, I assume that in the above cases of Upward Agree, the SD is not the sister/co-member of the valuee. By contrast, if we still want to find the valuer by downward search, the SD in cases of Upward Agree must be sets that include the NegP/TP (in 9), the AuxP (in 11), and the TP (in 12). Therefore, as long as these examples are concerned, it could be either the *v*P or CP that is assigned as the SD of MS for these cases.[12] MS then searches downward until the head with the ST is found.

We have seen that taking *v*P and CP as the SD can give us a sound empirical account for the attested upward agreement phenomena reviewed above. An immediate question is whether there is a theoretical reason why *v*P and CP, rather than other syntactic objects, should be assigned as the SD of MS here. A reasonable conjecture is that generally, the phase status of *v*P and CP may be relevant. Again, I essentially adopt Chomsky's (2001) theory of phases with a slight revision: *v*P and CP are phases, and the complement of the lower phase head is transferred when a higher phase is built.[13] As a consequence of this theory of phases, if a valuee remains unvalued in the phase that is transferred, like in the apparent upward agreement cases, the valuee will be

[10] Therefore, assuming NegP to be part of the CP phase, not of the *v*P phase, will open the possibility that MS can be initiated two times if the trigger is located internal to the lower phase, as long as the constraints from the Phase Theory (in particular cyclic Transfer) are observed. I leave an exploration of the consequences of this possibility for future research.

[11] As we have just seen, both of the analyses given by Zeijlstra (2012) may involve agreement across a *v*P phase boundary: the lowest n-word *nikomu* bearing [uNeg] is in the *v*P phase, whereas the negative marker or the negative operator which bears [vNeg] is located in the CP phase. If he adopts Chomsky's (2000) theory of Transfer, according to which the complement of *v* is transferred after *v* enters in the derivation, Zeijlstra (personal communication) has to assume that *v*P is not a phase (see also Keine & Zeijlstra 2021). If *v*P is not a phase, MS-Agree can still serve as a feasible analysis of negative concord in Czech, because MS will search downward directly from the CP phase. But if we adhere to the standard view that *v*P is a phase, the Upward Agree approach will also need to adopt instead Chomsky's (2001) theory of Transfer or the like to allow the object to be valued by a syntactic object in a higher phase.

[12] TP is less relevant as (12) shows that infinitive TP cannot be the domain that constrains Upward Agree.

[13] That is, I assume Transfer occurs when the full higher phase is built, instead of when the higher phase head enters the derivation. See Ke (to appear[b]) for discussion of a variant of Transfer Theory that is compatible with this paper.

rendered inaccessible for future MS-Agree. This will either cause the derivation to crash or lead to agreement failures (Preminger 2014). Note that MS is not obligatorily initiated to save the unvalued features from being transferred. Instead, the idea is that if the unvalued features are to be valued, the point when they are to be transferred is the last chance for MS to be initiated to value them. To put it another way, in order to value unvalued features (which is not obligatory), MS must be applied to find a target in the currently built root structure, which is a phase. The [uNEG], the unvalued feature associated with the participle morpheme, and the unvalued feature connected to the nominative case morpheme are all uninterpretable to the interfaces. Therefore, MS takes *the next higher phase* dominating the valuee and looks for a valuer. This is the last chance for the unvalued features to be valued by a valuer if the valuee is inside the Transfer domain, i.e., the lower phase head complement. What if an unvalued feature is located in the edge of the lower phase, and thus it will not be transferred until the next two higher phases (instead of the next higher phase) are built? The assumption is, as a general mechanism, MS should be triggered when the next higher phase is built. The advantage of this approach is that MS then can be blind to, and not be complicated by, the specific syntactic position of its trigger.[14]

Another question that is relevant here is: why not initiate MS after each application of Merge above the syntactic object with unvalued features? This will be in effect making the analysis similar to Upward Agree. There are two reasons speaking against this option. Firstly, such approach will always return a c-commanding target closer to the trigger, and thus predict an intervention effect caused by a closer target to all other competing targets higher up. This prediction concerning intervention effects is the same as that of Upward Agree. Empirical counter-evidence will be discussed in the next two sections. Secondly, from the theoretical perspective, such an approach will result in an inefficient mechanism, as it requires MS to be initiated frequently, and most of the searches will fail without returning a target.

If the phase status of *v*P and CP is relevant here, given that the heads with unvalued features are in the c-command domain of the valuers, we can formalize the locality constraint in upward agreement by assuming that the next built phase as the root structure is taken as the SD for the relevant MS, with additional constraints imposed by Chomsky's (2001) theory of phases:

(16) *The assignment of SD in apparent upward agreement*
In upward agreement, Minimal Search takes as its SD the next built phase.[15]

---

[14] In addition, compared to any other set that is contained in the next phase, the next built phase is the set that contains the most head members before Transfer applies, and therefore it is the SD within which the MS is most likely to find a valuer.

[15] "The next built phase," which is used to determine the SD (and the timing) of MS in upward agreement, refers to the phase that is built right above the trigger as the root structure at a particular derivational step. For instance, if the trigger is internal to a phase below the phase head, MS will be triggered when this local phase is completed.

With the assumption in (16), the cases of upward agreement can be perfectly captured by MS-Agree, as discussed above. Likewise, in (11), the unvalued feature on the verb bearing the parasitic participle can be valued by the auxiliary *hade* 'had' by downward searching from the CP. For multiple case assignment in Japanese (12), the unvalued case feature on the DPs can also be valued by the T head through multiple independent searches with SD = CP. Note that under the current analysis, multiple case assignment in Japanese does not require a [+multiple] feature or resort to a simultaneity approach in probing; instead, each instance of case assignment will be implemented by an MS searching downward from the next higher phase.

Therefore, MS-Agree which employs a downward search algorithm provides neat solutions to apparent cases of Upward Agree, with the assumption in (16).[16] This gives rise to an important question: since these cases of upward agreement are not standard ϕ-features, should they be treated as non-syntactic concord and thus be treated differently from regular ϕ-feature agreement? If this is the case, then we should not try to account for upward agreement and downward agreement with a single mechanism, and it is even misleading to do so (Preminger & Polinsky 2015). Bjorkman & Zeijlstra (2019) address this question by pointing out that treating concord phenomena such as negative concord and inflection doubling as non-syntactic would fail to explain their locality constraints. If these phenomena are syntactic in nature, then there is no reason in principle why these phenomena of concord should be treated separately from ϕ-feature agreement.

To provide an even stronger argument, however, I will discuss in the next two sections cases where upward ϕ-feature agreement is detected. The discussion of cyclic agreement will help us refine the current MS-Agree analysis by specifying the timing of search into the sister of the trigger or the next higher phase. A principled mechanism will be developed to determine when the sister of the trigger or the next higher phase must be taken as the SD: MS will always start from taking the sister of the trigger as the SD; only when the search fails, MS will take the

[16] A reviewer asks if MS-Agree can be extended to account for verbal inflection in clauses with multiple auxiliaries such as English "they would have written a note." This is an interesting question because if the answer is yes, then a downward search algorithm may not make a correct empirical prediction (cf. Wurmbrand 2012a for a Reverse/Upward Agree analysis). While this requires further investigation, the current answer is probably negative. I consider verbal inflection more like an instance of c-selection, which occurs between a head X and another head that X's sister/co-member immediately dominates. Such relation cannot be long-distance as it cannot be established when skipping any intervening c-commanding head (disregarding adverbs), different from what we have seen in Agree. I thus believe verbal inflection is in principle not implemented by Agree. Note that inflection doubling is different in that it is a relation between two heads that are separated by another head. In addition, an Agree-based analysis of verbal inflection must make some construction-based assumptions. For example, it assumes an unvalued feature on *have* in "they would have written a note" for it to agree with a corresponding valued feature on *would*, whereas this unvalued feature is not present on *have* in "they have written a note"; otherwise, the unvalued feature would be left unvalued. Yet such optionality does not apply to *written*.

next built phase above the trigger as its SD and search into it (see Section 4). This will address the question as to when and why MS searches into the next built phase, further validating the assumption in (16). Additionally, the discussion of complementizer agreement in Lubukusu will shed light on a critical empirical diagnostic that can tease apart the predictions of MS-Agree and Upward Agree. Overall, I argue that MS-Agree with the assumption in (16) also offers a neat and simple analysis to upward ϕ-feature agreement, in addition to instances of concord. These upward ϕ-feature agreement phenomena, however, cannot be fully captured by the Upward Agree analysis.

4 Cyclic agreement as MS-Agree

In this section, I use the well-attested phenomenon, cyclic agreement, as an illustration of MS-Agree also being able to provide a neat analysis of upward ϕ-feature agreement. Another benefit of discussing cyclic agreement is that it gives us a clue about when and why MS takes as its SD the next built phase (containing the valuee) in cases of upward ϕ-agreement.

Béjar & Rezac (2003), Rezac (2004), and Béjar & Rezac (2009) investigate an interesting agreement pattern in Georgian: agreement occurs in two cycles. In the first cycle, a head with unvalued features (the trigger) probes downward to find a head matching the unvalued features. If the search in the first cycle is not successful, the unsatisfied trigger is further projected higher, enabling it to probe into its specifier which it originally does not c-command. Let us see some examples.

(17) *Georgian* (Rezac 2004: p. 72)

a. m-xedav-t
1-see-PL
'You(.PL) see(.PL) me.'

b. g-xedav-t
2-see-PL
'I see(.PL) you(.PL).'

Béjar (2003) and Béjar & Rezac (2003) assume that person and number features are represented separately in Georgian: *v* bears a [uPerson] feature and T bears a [uNum] feature. (17a) shows that, when the subject is plural, $T_{[uNum]}$ agrees with the subject (not the object) and $v_{[uPerson]}$ agrees with the object (not the subject). By contrast, (17b) exemplifies that, when the subject is singular, $v_{[uPerson]}$ still agrees with the object, but $T_{[uNum]}$ now agrees with the object instead of the subject. Béjar and Rezac suggest that the singular number feature on the subject is underspecified[17] or not

[17] Let us ignore the issue regarding what it exactly means for a head to be underspecified for a certain feature as the question is orthogonal to the focus of the current discussion (see Preminger 2011 for an alternative analysis).

present in syntactic representations in Georgian, so downward searching from $T_{[uNum]}$ skips the subject and expands the search down to the object. The presence and then disappearance of the intervention effect in number agreement conditioned by the subject's (under)specification status in (17a) and (17b), strongly support the idea that this is a case of downward search.

As discussed above, the examples in (17) illustrate the first cycle Agree, which is an instance of downward Agree. The following examples show the second cycle Agree, which is more important for our purposes:

(18) *Georgian* (Rezac 2004: p. 73)

a. v-k'vd-eb-i
   1-die-X-IND.PRS
   'I am dying.' (originally from Hewitt 1995: p. 118)

b. mi-v-c'er-e
   X-1-write-AOR
   'I wrote it to her.' (originally from Hewitt 1995: p. 190)

(18a) indicates that when the object is absent, $v_{[uPerson]}$, which usually agrees with the object, now agrees with the subject. Similarly, in (18b), where the object is an underspecified third person pronoun, $v_{[uPerson]}$ also agrees with the subject instead of the object. Béjar & Rezac (2009) suggest the reason that $v_{[uPerson]}$ can agree with the subject when the object is not available for agreement is that Agree can change its direction when necessary. That is, in the second cycle, given that nothing can be returned in the first cycle downward search, upward search is initiated and the subject is returned. As shown in **Figure 2**, in cycle one, $v_{[uPerson]}$ looks for the object/internal argument (IA). If IA's person feature is underspecified (indicated by the question mark) or if the IA is absent (indicated by parentheses), the second cycle Agree is initiated. In this second cycle, $v_{[uPerson]}$ searches upward to agree with the subject/external argument (EA).

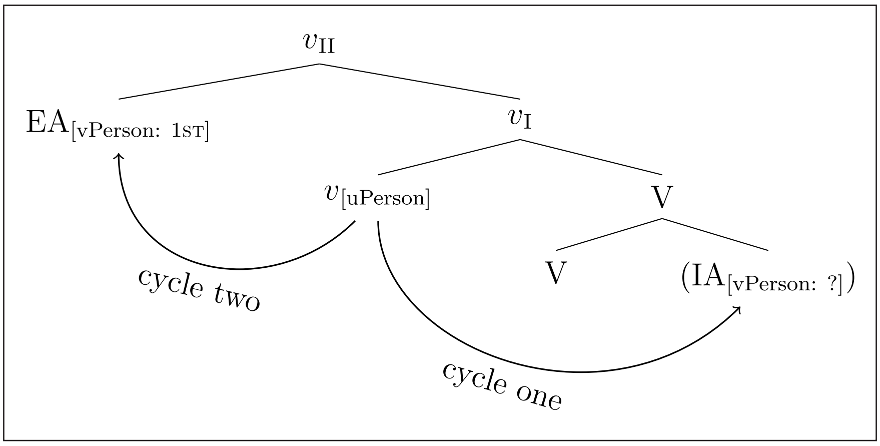


Figure 2: Béjar & Rezac's (2009) analysis of cyclic agreement.

Crucially, Béjar & Rezac (2009) propose that the second cycle Agree could be regular downward search if we assume that the label of v_{I} and v_{II} are both the v head. That is, $v_{[\mathrm{uPerson}]}$ can probe down from the v_{II} projection as v_{II} is just $v_{[\mathrm{uPerson}]}$.

This analysis builds upon the Labeling Theory proposed by Chomsky (1995, 2000), where Merge of two syntactic objects α and β produces a label, and the label can be either α or β, depending on which one syntactically selects the other (setting adjunction/Pair Merge aside). Béjar & Rezac (2009) move a step further: the label, like a regular head bearing an unvalued feature, can initiate another search for Agree, because the unvalued feature is projected to the label. This analysis thus identifies regular heads with labels in the computation of agreement.

This is a clever use of the Labeling Theory proposed by Chomsky (1995, 2000) with non-trivial empirical success. However, as pointed out by Ke (2019, to appear[a]), such an approach removes the distinction between heads and sets once labels have been computed for sets, whereas such a distinction is fundamental for Agree and Labeling to work in the first place. For instance, the label of the head–complement construction, the so-called {X, YP} construction, is computed based on the head–set distinction (Chomsky 2013). If labels are created upon Merge, and if the label of YP, i.e., the head Y, is fed to the Labeling of {X, YP}, this set would be treated the same as {X, Y}, causing a Labeling issue.

In addition, Béjar & Rezac (2009) assumes a special variant of the Labeling Theory, under which (c-)selection is critical to Labeling: the selecting head projects as the label upon Merge. This complicates the Labeling operation, as now Labeling must be much more than MS: Labeling must access to c-selection information in order to identify the label for a set. However, such selection information is not directly relevant for MS in Agree. Since labels are computed upon Merge, this analysis also excludes the possibility that labels are computed top-down right before the relevant syntactic objects are transferred for interpretation at the interfaces (Chomsky 2013).

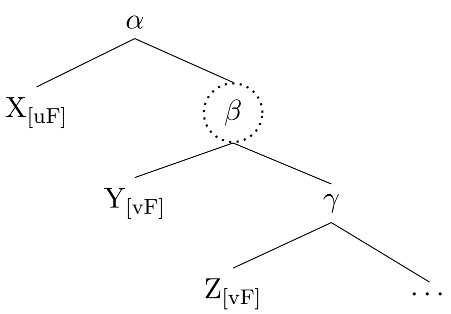
Furthermore, on the one hand, Béjar & Rezac's (2009) analysis treats the label of a set just like a head when using it as a trigger for Agree such that it can probe. On the other hand, this analysis still presumes the head-label distinction in Agree. For instance, in the first cycle Agree, v_{I} probes into the {V, IA}. Note that, at this derivational step, {V, IA} already obtains its label, which is the head V. Crucially, Béjar & Rezac (2009) must assume that v_{I} is able to search into the labeled set {V, IA} as a regular set, not as a label which is simply equal to a head. In addition, the IA must be already labeled upon its Merge into the structure, so probing into the IA implies that it is the set not the label of the set that is taken as the search domain. Similarly, in the second cycle Agree, when v_{II} is projected as the label of {EA, v_{I}}, the EA should also have a label upon its entering into the derivation. Since now v_{II} needs to probe into the EA, the EA must be a set rather than a label. That is, according to Béjar & Rezac's (2009) analysis, a set can probe like a head once it is labeled; however, a labeled set can also be probed into just as a regular set, not like a head. These problems seem to be originated from Chomsky's (1995, 2000) Labeling Theory, which Chomsky (2013, 2015) have acknowledged and discussed extensively.

Béjar & Rezac's (2009) analysis faces other difficulties if we adopt the theory of Labeling in Chomsky (2013, 2015) (see also Ott 2012 and EKS 2014 (Epstein, Kitahara & Seely 2014)), as the whole *v*P (v_{II}), being an instance of {XP, YP}, cannot be labeled as *v* until the EA moves to a higher position, e.g., SpecTP. This implies that it is only after the TP is built that we can come back to the v_{II} to probe for the lower copy of the EA. Note, however, that after movement, the lower copy of the EA is standardly assumed to be invisible for Agree. I will argue that MS-Agree is free of the above theoretical issues. Of course, if we adopt Chomsky's (2013, 2015) Labeling Theory, labels by definition are only for interpretation purposes, and they are not expected to be involved in any syntactic operation such as probing for agreement purposes.

Now we are ready to offer our re-analysis of cyclic agreement with MS-Agree as defined in (4). I adopt the idea that Agree occurs in two cycles. The single, crucial difference between these two cycles is the SD (and correspondingly, the timing) of MS. In the first cycle, the SD is the sister/co-member of the trigger (the head bearing unvalued features that initiate the search), as shown in (19), with a tree diagram in (20). The dotted circle indicates the syntactic object that is taken as the SD.

(19) *Cycle One MS-Agree*
SD = $X_{[\text{uF}]}$'s sister = β[18]
Timing: when $X_{[\text{uF}]}$ enters in the derivation

(20) *Cycle One MS-Agree: SD =* β



This corresponds to the first cycle Agree in Béjar & Rezac (2009), where normally $Y_{[\text{vF}]}$, which c-commands $Z_{[\text{vF}]}$, causes an intervention to Z. That is, only Y will agree with X. But in the cases where the relevant feature on Y is underspecified, MS can skip Y and find Z instead. Downward agreement in general is then modeled by Cycle One MS-Agree.

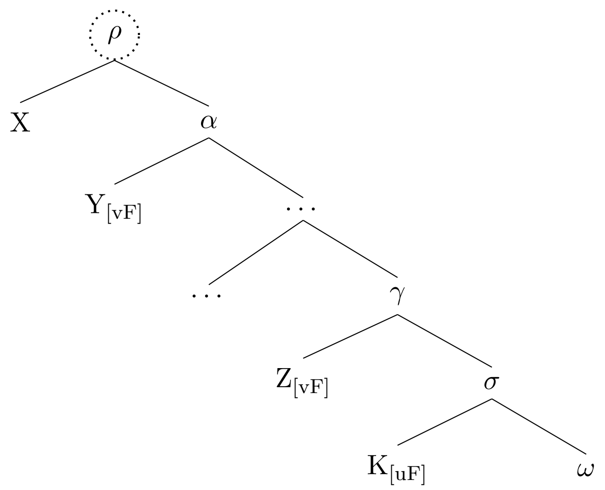
On the other hand, in the second cycle, the SD should be a set containing both the trigger and the search target. We have mentioned in Section 3 that all the attested cases of upward

[18] A reviewer asked if the SD of MS-Agree can be restricted to phases only even in the first cycle. Although this modification will simplify MS-Agree and assign phases as a unique initial SD for all cases of MS, it cannot capture the first cycle agreement as discussed in this section. This is because MS will always be delayed until the current phase is built, returning first the c-commanding targets even if there is another target inside the trigger's sister. Such a simplification misses, for example, the Internal Argument > External Argument preference in person agreement in Georgian (Béjar & Rezac 2009).

agreement can be accounted for if we assume the SD is the next higher phase to be built or completed. This solution can be exactly carried over to cyclic agreement discussed in this section, as shown in (21), with a tree diagram in (22).[19] Again, the syntactic object that is taken as the SD for the relevant MS is indicated by the dotted circle.

(21) *Cycle Two MS-Agree*
SD = the next phase ρ
Timing: when the phase ρ is completed

(22) *Cycle Two MS-Agree: SD = ρ*



Here we further specify the timing of the second cycle MS-Agree, namely, at the time when the next higher phase ρ is completed; more specifically, it is the time when the lexical array for the next phase is exhausted.[20] It is important to note that (22) predicts that $Y_{[vF]}$ will be returned by downward search, and thus it will agree with $K_{[uF]}$, although it is further away from K compared to $Z_{[vF]}$. That is, Y will cause an intervention effect to Z, not vice versa.

[19] There are potentially two ways to find and handle the triggers for the second cycle MS-Agree. One approach is to assume a (memory) stack to store (heads with) unvalued features. Fong & Ginsburg (2019) provide a computational model that implements a variant of this idea. Another approach is what an anonymous reviewer suggested: before Transfer, the syntax system can apply MS to return the heads with unvalued features. At this point, it is unclear which approach is superior empirically, although the later idea sounds more promising theoretically since it does not assume an additional memory stack.

[20] This will capture the contrast between A-movement and Ā-movement in cyclic agreement in Hindi-Urdu (Keine & Dash 2017, 2018), which I cannot discuss due to space limitations: A-movement but not Ā-movement feeds the second cycle Agree. That is, if a DP A-moves from a lower phase to the matrix SpecTP, since the C head has not yet entered into the derivation, the DP can agree with the matrix T; by contrast, if a DP Ā-moves to SpecCP, since the lexical array for the matrix CP is exhausted after the C head merges in, the DP cannot agree with T. Under the current analysis, for the second cycle MS-Agree, MS will look down from C′ after A-movement but before Ā-movement occurs, allowing the A-moved DP to be returned but excluding the Ā-moved DP. Please be referred to Ke (2019, ch. 2) for relevant discussion.

Note that the definition of Cycle Two MS-Agree in (21) also captures the locality constraint of upward agreement in Georgian. Since the *v* head as a trigger for the second cycle MS-Agree initiates an MS after the local *v*P is completed, Spec*v*P will be returned as long as such a Spec*v*P exists. We will see a case where the trigger is a phase head and it does not have a specifier; in that case, non-local agreement is predicted, a prediction that will be confirmed by complementizer agreement in Lubukusu, which will be discussed in the next section.

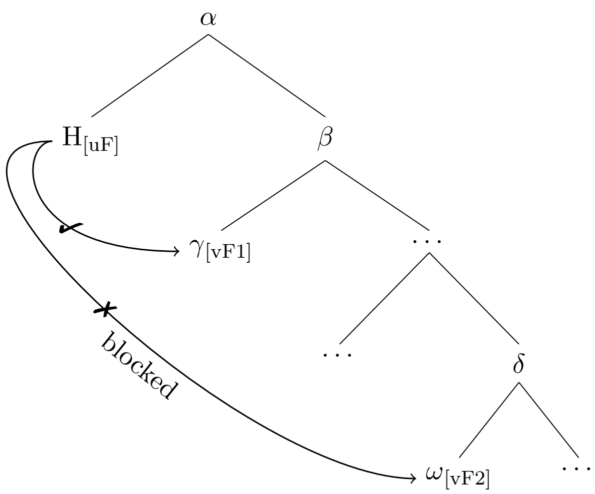
5 Intervention effects

An important type of evidence that can be used to argue for or against Upward Agree or downward Agree is intervention effects. Baker (2008) puts forward the intervention condition, together with the c-command condition, to account for expected intervention effects in both Upward Agree and downward Agree:

(23) *The c-command condition and the intervention condition* (Baker 2008:45,47)
F agrees with XP only if:
a. F c-commands XP or XP c-commands F (the c-command condition).
b. There is no YP such that YP comes between XP and F and YP has ϕ-features (the intervention condition).

The c-command condition requires the valuer to be either c-commanded by the valuee or c-command the valuee. This indicates that Baker (2008) allows both Upward Agree and downward Agree. The intervention condition can be interpreted from two perspectives. First, with regard to downward Agree, the intervener is c-commanded by the trigger and it blocks the trigger from reaching a competing target which is in turn c-commanded by the intervener. As shown in (24), the intervener is γ, and it prevents H from agreeing with *ω*.

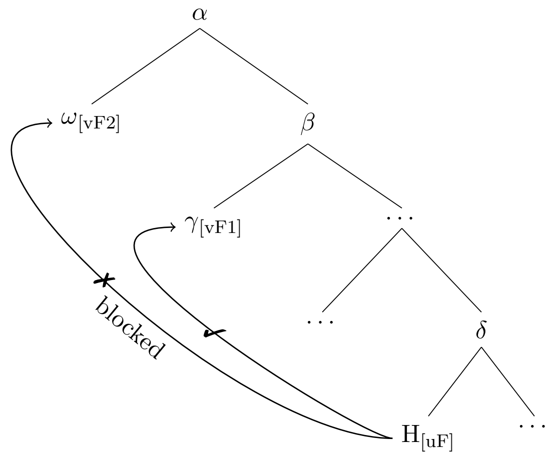
(24) *Intervention in downward Agree*



Such intervention effects have been reported in various studies (e.g., Baker 2008, Béjar & Rezac 2009, Keine 2016, Keine & Dash 2018, Polinsky & Potsdam 2001) in support of downward agreement.

On the other hand, with regard to upward agreement, a different intervention pattern is predicted:

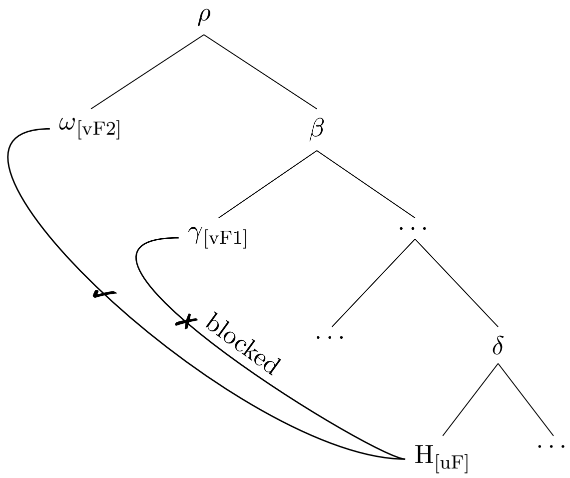
(25) *Intervention in Upward Agree*



(25) indicates that the intervener γ c-commands the trigger H and is c-commanded by a competing target *ω*.

MS-Agree makes the same predictions with Baker's (2008) hybrid analysis as far as downward agreement is concerned. Importantly, however, MS-Agree may make a different prediction if *ω* and γ in (25) are located in the same minimal phase with H or in the next higher phase above H, e.g., ρ. This is because MS will search down from the top and thus reach *ω* before γ. When the SD is equal to ρ, the prediction of MS-Agree is as shown in (26):

(26) *MS-Agree's prediction of intervention if SD* = ρ



Therefore, intervention effects provide us one of the best diagnostics for MS-Agree versus Upward Agree. The prediction of intervention effects also distinguishes the MS-Agree analysis from Carstens's (2016) Delayed Valuation analysis, as I will argue below in this section. The Delayed Valuation analysis allows a trigger to be delayed in its valuation and then agrees with a c-commanding target at the time the next phase head enters the derivation. The Delayed Valuation analysis thus permits any c-commanding target inside the next built phase to agree with the trigger, that is, it does not predict an intervention effect with upward agreement cases.

However, in order to make effective use of this diagnostic, we must have an intervener which shares the same feature attributes but with different feature values compared to those on the competing target. Otherwise, if the intervener bears the same feature attributes and feature values with the competing target, we will never know whether the intervener or the competing target agrees with the trigger: these two options result in identical agreement patterns on the surface. The Upward Agree analysis could simply assume that there is an agreement link that is shared by multiple targets (see Bjorkman & Zeijlstra 2019 for such an analysis of long-distance agreement in Tsez).

Unfortunately, for negative concord, inflection doubling, and multiple case assignment, the features involved are very specific, and there are usually only two options for each feature attribute: the feature is either present or not. For instance, the [vNEG] feature does not make a distinction between different subtypes/values of the negative feature. A syntactic object can either have the feature or not have the feature. As a consequence, there is no way to find an intervener in the classical instances of upward agreement to prove or disprove Upward Agree. This is the same for the MS-Agree analysis of these phenomena due to the same reason. Therefore, Upward Agree and MS-Agree can be considered equally valid analyses as far as the above three types of upward agreement phenomena are concerned.

Φ-feature agreement is different in this respect in a non-trivial way. ϕ-features may have two-way (e.g., singular vs. plural) or even three-way distinctions (e.g., 1st, 2nd, and 3rd person) in feature values for each feature attribute ([NUM] and [PERSON]). Hence, an intervener could be, for example, singular, and its competing target could be plural. By looking at the agreeing number feature at the trigger, we can tell right away whether the intervening syntactic object has created an intervention effect to the competing target or not. Therefore, below I focus on potential intervention effects in agreement with ϕ-features.

5.1 Subject–complementizer agreement in Lubukusu

A phenomenon that is critical to the current purpose is the subject–complementizer (S–C) ϕ-feature agreement in Lubukusu, a Bantu language. S–C agreement in Lubukusu is pertinent to the current discussion because it is, at least appears to be, an instance of long-distance upward agreement.

As Preminger & Polinsky (2015) point out, the most convincing evidence for upward (ϕ-feature) Agree is long-distance Upward Agree[21] that crosses a finite TP (or a CP) clause boundary, which generally excludes covert movement; namely, a head with unvalued features triggers agreement with a target in a higher TP where the target receives its case or θ-role. This is because, due to derivational opacity (the derivational history cannot always be known), we can neither justify nor falsify Upward Agree if at least in some circumstances there is covert movement of the trigger to a higher position c-commanding the target; it is also quite unlikely that we can completely exclude the possibility that the target is in fact raised from a lower position c-commanded by the trigger in an earlier derivation stage. In both of these cases, the trigger c-commands the target at certain point of the derivation, an agreement pattern that can be handled with downward Agree. However, with long-distance Upward Agree, covert movement of the trigger that crosses a finite TP, as an instance of head movement, is widely considered unlikely. Raising of the target to a position in a higher TP where it receives its case or θ-role, as an instance of A-movement, cannot be easily justified either. Therefore, long-distance Upward Agree cannot be readily formalized as instances of downward Agree, and thus it is strongly in favor of Upward Agree.

More importantly, as we have mentioned in the discussion of cyclic agreement, local Upward Agree can usually be captured by the Spec–head relation, which can in turn be formalized as downward Agree if Chomsky's (1995, 2000) Labeling Theory or a theory of projection is assumed, although we have argued that these assumptions complicate and are not well compatible with the more recent formalization of the Labeling Theory (Chomsky 2013, 2015).

Unfortunately, examples of long-distance Upward Agree are extremely rare in the literature. Preminger & Polinsky (2015) thus take it as an unattested agreement pattern in human languages.

However, at least at the surface form, S–C agreement in Lubukusu is an instance of long-distance upward agreement, as the embedded C head agrees with the immediately superordinate subject. Preminger & Polinsky (2015) are aware of this type of upward agreement, but following some previous analyses of S–C agreement in Lubukusu, e.g., Diercks (2013), they assume that S–C agreement in Lubukusu can be successfully explained by an analysis that does not resort to Upward Agree. I argue below that such an non-Agree analysis encounters many difficulties and is not tenable; on the other hand, MS-Agree gives us a much simpler analysis that is free of these problems. More important, S–C agreement provides a testing ground for MS-Agree versus Upward Agree.

Diercks (2013) argues convincingly that in Lubukusu an embedded C agrees only with the subject, not the object, of the immediately superordinate clause in full ϕ-agreement, in terms

[21] I use "long-distance Agree" for long-distance downward Agree, like one can usually find in the literature. Now I am using "long-distance Upward Agree" to distinguish it from long-distance downward Agree.

of noun class agreement. Exemplified in (27), the embedded C head *-li* agrees with the matrix subject (*ba-ba-)ndu*, rather than with the object *Alfredi*, although the latter is structurally closer to the embedded C.

(27) *Subject–complementizer agreement in Lubukusu* (Diercks 2013: p. 358)
Ba-ba-ndu ba-bol-el-a Alfredi ba-li a-kha-khil-e.
2-2-people 2SA-said-APL-FV 1Alfred 2-that 1SA-FUT-conquer
'The people told Alfred that he will win.'

Further evidence shows that (27) is not a surface reflection of an agreement between the logophoric center of the embedded CP (i.e., the "source," see Sells 1987) and the lower C. (28a) indicates that a non-logophor subject can agree with the lower C head; on the other hand, (28b) suggests that the non-subject logophor, *Sammy*, cannot trigger agreement with the lower C. Token together, it is clear that the subject of the higher clause rather than the logophor of the embedded CP agrees with the C head in the lower CP.

(28) *The subject agrees with the complementizer* (Diercks 2013: p. 368)
a. Sammy ka-bol-el-wa a-li ba-keni b-ola.
1Sammy 1SA-say-APL-PASS 1-that 2-guests 2SA-arrived
'Sammy was told that the guests arrived.'
b. Ba-sasi ba-bol-el-wa nende Sammy mbo/*a-li ba-keni ba-a-rekukha.
2-parents 2SA-say-APL-PASS by 1Sammy that/*1-that 2-guests 2SA-PST-leave
'The parents were told by Sammy that the guests left.'

It should also be noted that S–C agreement is subject to locality constraints. A C head must agree with a subject that is right in the immediately superordinate clause, not with subjects in any higher clause, as illustrated by (29).

(29) *The locality constraint* (Diercks 2013: p. 371)
Alfredi ka-a-loma a-li ba-ba-andu ba-mwekesia bali/*ali o-mu-keni
1Alfred 1SA-PST-say 1-that 2-2-people 2SA-revealed 2-that/*1-that 1-1-guest
k-ola.
1SA-arrived
'Alfred said people revealed that the guest arrived.'

Diercks (2013) claims that such subject-orientation in Agree in examples such as (27) can be explained if we assume a null operator in SpecCP which agrees with C *-li* through a Spec–head relation. This null operator is a subject-oriented anaphor, akin to the reflexive clitic *se* in French (Safir 2004). This reflexive raises to adjoin to a higher T where it is bound by the matrix subject (cf. Chomsky 1986), as shown below:

(30) [$_{CP}$ C [$_{TP}$ **ba-ba-ndu**$_{[v\phi]j}$ REFL$_j$-T [$_{vP}$ t_j v [$_{VP}$ V [$_{DP}$ Alfredi] [$_{CP}$ REFL$_j$ **ba-li**$_{[u\phi]}$ …]]]]]

This analysis correctly captures the lack of intervention effects with regard to the object *Alfredi*. However, a potential drawback of Diercks's analysis is that a null reflexive is positioned in every embedded CP that agrees, yet they make no semantic (or phonological) contribution to their local CP. As acknowledged by Diercks (2010), these reflexive operators are not logophors or elements that are related to any thematic roles, and they receive no identifiable thematic interpretation themselves. To make their status even more dubious, the null reflexives are required to undergo covert movement to the T head in the superordinate clause, which is a crucial aspect of the analysis to account for the subject-orientation effects in S–C agreement, but this assumption is essentially unfalsifiable, as it is a null element going through a covert movement.

There are still other drawbacks of Diercks's (2013) analysis. Diercks has to assume two distinct agreement relations, a subject–reflexive and a null reflexive–C dependency. The subject–reflexive binding dependency has been argued *not* to be a primitive operation, but one that is reducible to Agree (e.g., Hicks 2009 and Reuland 2011; see Ke 2019 ch. 3 for a review of various approaches to reduce reflexive binding to Agree). An alternative analysis that can directly analyze S–C agreement as a case of Agree is thus simpler (and preferable). In addition, the null reflexive-C agreement requires the reflexive to be involved in ϕ-feature agreement, an agreement that has been shown to be impossible or unattested (so called "anaphor (anti-) agreement effects," cf. Preminger 2018, Rizzi 1990; Woolford 1999; but see Rudnev 2020 for arguments of an exceptional case of agreeing anaphors).

The above problems can be avoided if we assume a direct agreement relation between C and the superordinate subject. With a direct Agree analysis, we do not need to assume a null reflexive, a hypothesis that does not have empirical basis. The analysis is therefore significantly simplified. In fact, Diercks (2010) considered such a possibility. Nevertheless, Diercks (2010) ends up not adopting a direct Agree analysis, mainly due to the lack of intervention effects when there is an indirect object intervening between the agreeing subject and C. (27), repeated below in (31a), shows that the indirect object 'Alfred' cannot agree with the embedded C. (31b) is a sketch of the relevant aspect of the structure of (31a).

(31) a. Ba-ba-ndu ba-bol-el-a Alfredi ba-li a-kha-khil-e.
2-2-people 2SA-said-APL-FV 1Alfred 2-that 1SA-FUT-conquer
'The people told Alfred that he will win.'

b. [$_{vP}$ Subject ... [Object ... [$_{CP}$ C ...]]]

This is a problem if we assume Upward Agree (e.g., Zeijlstra 2012; Bjorkman & Zeijlstra 2019), which Diercks (2010) does. An Upward Agree analysis will incorrectly predict that in (31), the embedded C must agree with the object, which is structurally closer than the subject to the embedded C. However, if we adopt MS-Agree instead of the Upward Agree approach, this major obstacle for the direct Agree analysis is naturally removed. With the MS-Agree analysis, we no longer predict an intervention effect caused by the indirect object, as MS starts its search from the next higher phase, namely, the *v*P in (31b), and the *v*P-internal subject rather than the indirect object is returned; that is, the search is top–down in the sense of tree representations.

In fact, I argue that “subject-orientation” in Lubukusu can in fact be derived by MS-Agree defined previously in (4) with an independently assigned SD, as in the second cycle of cyclic agreement (21). Instead of being mediated by a null reflexive, I assume that the embedded C can directly agree with the matrix subject if the next built phase is taken as the SD.

To make more concrete the MS-Agree account of subject-orientation in S–C agreement in Lubukusu, consider again the structure under discussion:

(32) $[_{vP}$ **ba-ba-ndu**$_{[v\phi]}$ v $[_{VP}$ V $[_{DP}$ Alfredi$]$ $[_{CP}$ **ba-li**$_{[u\phi]}$ … $]]]$

As a reminder, I assume that MS observes Chomsky’s (2001) Phase Impenetrability Condition; namely, a phase head complement is transferred and is not accessible to MS after a higher phase head enters the derivation. Therefore, when the matrix *v* head merges in, the complement of the embedded C will be transferred, indicated by the dots in (32). The lexical array for the matrix *v*P phase includes the *v*P-internal subject. Now the crucial step is: an MS is conducted with ST = [uϕ] and SD = the phase that is built so far, i.e., the matrix *v*P. This MS first finds the *v*P-internal subject *ba-ba-ndu* (more accurately its D head), which causes intervention to the object *Alfredi*. The subject-orientation phenomenon in Lubukusu is therefore explained: it is also a natural consequence of downward search from *v*P.

Taking the next built phase as the SD can derive all the cases of upward agreement we have discussed, including the subject-orientation effect in Lubukusu, and this is consistent with Carstens’s (2016) Delayed Valuation analysis for Lubukusu S–C agreement. In order to account for Diercks’s (2010, 2013) observation regarding S–C agreement in Lubukusu, Carstens (2016) proposes that the valuation of C can be delayed up to the point of cyclic Transfer, for example, when a higher phase head merges in. Delayed Valuation is initiated only when no target exists in the c-command domain of the trigger. The empirical examination of downward and upward agreement phenomena across languages in this paper supports this overall proposal.

However, a side effect of the Delayed Valuation approach is that it allows the indirect object in the superordinate clause above C to agree with C in Lubukusu S—C agreement, and thus would miss the subject-orientation effect if without an additional mechanism to preclude the indirect object from agreeing with C. Carstens (2016) therefore resorts to the activation condition to prevent the indirect object from agreeing with C: because the indirect object receives an inherent dative case, which renders it inactive and inert to agreement. Such an analysis then needs to explain that, in Lubukusu, (i) why dative DPs can be involved in an agreement relation in some other cases, and (ii) why a subject DP is not rendered inactive after it has been involved in an agreement relation with the C head in the subordinate clause and can still be involved in a ϕ-feature agreement relation with its local T head. Similarly, in West Germanic, the subject can agree with both the unvalued ϕ-features on its local T and C heads: an argument needs to be made to explain why subjects, which agree with T already, are not turned to be inactive but can further agree with C. Carstens's (2016) solutions to both of these exceptional cases, although intriguing and having the potential to account for some other important phenomena, lead to considerable theoretical complications to the operation Agree (see Carsten 2016 Section 3.6 for relevant discussion). In principle, if the Delayed Valuation approach is adopted, whenever an intervention effect is observed, a reason has to be found in order to exclude from agreement a competing target which is either in the c-command domain of the probing head or in a higher phase but does not agree with it.

MS-Agree, however, makes a prediction different from Carsten's Delayed Valuation approach. Subject-orientation in Lubukusu, which is due to the lack of intervention effect from the dative object, is a direct consequence of downward search: the higher target (the subject) is found by downward MS before the lower target (the object) is reached, not the other way around. In addition, MS-Agree accounts for intervention effects in downward and upward agreement with the same downward search algorithm, whereas the Delayed Valuation approach, on the one hand, will probably need to use downward search to derive intervention effects in downward agreement, and on the other hand, may resort to activation condition or other mechanisms to exclude any intervening potential candidates in upward agreement.

The differences between the MS-Agree analysis and another prominent analysis of S–C agreement in Lubukusu is also worth noting. Diercks, van Koppen & Putnam (2020) proposes that the agreeing C head, which bears interpretable but unvalued anaphoric features, moves to adjoin to a higher *v*P (see also Carstens 2016). This higher copy of C head (specifically Force^0) will then probe for the subject in the superordinate *v*P. The lower copy of the C head, rather than the higher copy, is then spelled out (the valued features hence must pass down from the higher copy to the lower copy before the spelling-out). MS-Agree, on the other hand, distinguishes itself by not making such an assumption regarding the movement of the C head as an adjunct to the *v*P.

Note that in Lubukusu, only the second cycle MS-Agree, not the first cycle MS-Agree, is relevant. One may wonder why this is the case. A conjecture is, following Carstens (2016), the CP periphery in Lubukusu has the following structure, where Force, Interrogative (Int), and Finite (Fin) are phase heads (Rizzi 1997, 2001).

(33) $[_{vP}$ **ba-ba-ndu**$_{[v\phi]}$ v $[_{VP}$ V $[_{DP}$ Alfredi$]$ $[_{ForceP}$ **ba-li**$_{[u\phi]}$ $[_{IntP}$ Int $[_{FinP}$ Fin $[_{TP}$ DP$_{subject}$ … $]]]]]]$

When the agreeing C_{Force} *-li* enters the derivation, FinP has been transferred, rendering the lower subject inaccessible for C_{Force}. Due to space limitations, interested readers are referred to Ke (2019: pp. 85–87) for relevant discussion on details of this hypothesized analysis and its potential alternatives.

To summarize, the current MS-Agree analysis of S–C agreement in Lubukusu does not assume (i) the existence of a null reflexive, nor (ii) covert movement of the null reflexive, nor (iii) agreement between the reflexive and the embedded C head. In short, it does not run into any of the problems that we identified above for Diercks's (2013) indirect Agree analysis. In addition, compared to Diercks's (2013) analysis, the current MS-Agree analysis is simpler and more intuitive, because now S–C agreement is established directly by MS-Agree, whereas Diercks needs to assume two distinct agreement relations, subject-reflexive and reflexive-C agreement.

5.2 Potential intervention effects in upward ϕ-feature agreement

We have seen that the MS-Agree analysis of S–C agreement in Lubukusu provides critical evidence for downward MS for Agree, which accurately capture the intervention effects caused by the structurally higher subject to the object. In this subsection, I will discuss two pieces of evidence for Upward Agree offered by Baker (2008), who proposes that both downward Agree and upward Agree exist in natural languages. I argue that closer scrutiny reveals a contradiction in the interpretation of the data regarding upward agreement.

Baker (2008) supported an Upward Agree analysis besides downward Agree based on ϕ-feature agreement on adjectives in Italian (examples are taken from Cinque 1990). The following example indicates that the structurally lower DP *i sostenitori* 'the supporters,' rather than the structurally higher DP, the null matrix subject, agrees with the adjective *infelici* 'unhappy.' The critical aspect of the structure is represented in (34b), where F_A is a functional agreeing head for predicative adjectives. In terms of (34b), F_A agrees with NP2 rather than NP1.

(34) *Adjective agreement in Italian* (Baker 2008: p. 47; originally from Cinque 1990: p. 28)

a. Ne ha resi infelici (*infelice) i sostenitori.
of.it has.3s rendered unhappy.M.PL (*unhappy.M.SG) the supporters.M.PL
'He has rendered its supporters unhappy.'

b. [NP1 verb [NP2 Pred [F_A [AdjP]]]]

Note that Baker (2008) does not assume that NP2 is originally generated inside the AdjP in (34b) due to independent reasons. Therefore, F_A can only probe upward to agree with NP2. Apparently, NP2 seems to block agreement between NP1 and F_A because it intervenes between them, consistent with the prediction of intervention effects by an Upward Agree analysis.

In Footnote 21, Baker (2008) mentions another case in Icelandic to support this type of intervention effects. "[I]n the Icelandic version of 'The women(NOM) consider Mary(DAT) to be cold,’ ‘cold’ must be default masculine singular, not feminine plural in agreement with ‘the women’ (Sigurðsson, personal communication).” (p. 48) This intervention effect even occurs when “the subject of its predication has quirky case and hence does not trigger agreement on the adjective itself.” (ibid.) It must be noted that what Icelandic reveals is not as predicted by Upward Agree. When the local subject of the adjective predication has quirky case, it should be inaccessible for agreement. Consequently, if Upward Agree is the operation involved. the quirky subject can no longer be an intervener; and this should render the higher, matrix (*v*P-internal) subject, ‘the women,’ available for agreement, unless the matrix subject is prevented from agreeing with F_A due to other reasons.

Indeed, Phase Theory may be the underlying mechanism that prohibits the matrix subject from agreeing with F_A. The idea is that F_A is a phase head, and it is transferred when a higher phase head, the matrix F_V, enters in the derivation. As a consequence, the matrix subject is no longer accessible to F_A as the latter has been transferred when the former enters the derivation. The same reasoning can also explain why in the Italian example (34), the matrix subject is not available for agreement with the embedded F_A. In sum, upon closer examination, previously acknowledged as intervention effects in support of Upward Agree may be a by-product of independent restrictions due to Phase Theory. Therefore, they do not constitute counterexamples to MS-Agree: MS-Agree, also constrained by the theory of phases, makes the same predictions as Upward Agree as far as these data are concerned.

6 Comparing with the Upward Agree analysis

In this section, I will make a direct comparison between MS-Agree and Upward Agree. The main differences are on the following three aspects.

(i) The MS-Agree analysis employs a search algorithm (MS) that systematically searches downward, and thus avoids all complications that Bjorkman & Zeijlstra (2019) need to assume to deal with attested cases of long-distance downward search, as pointed out by Preminger (2013). In addition, the search algorithm, which is considered a third factor, is independently developed for Agree as well as Labeling. Therefore, theoretically, MS-Agree comes as at least partially free.

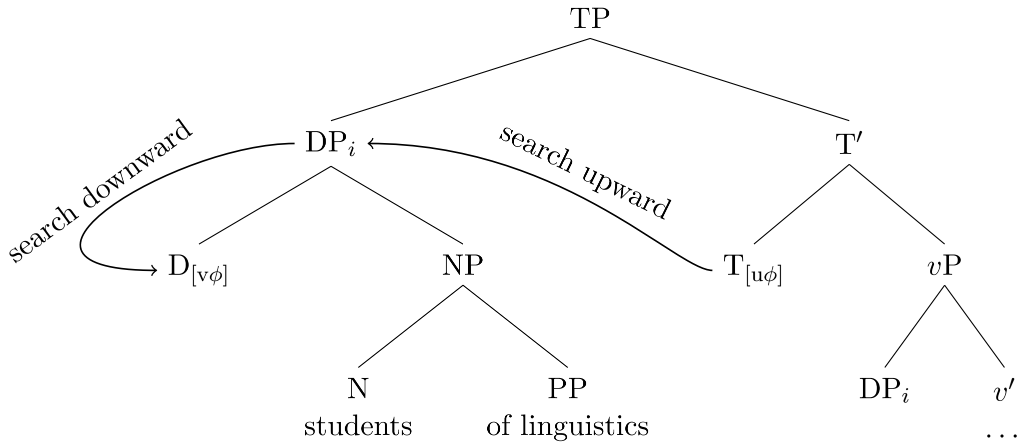
(ii) The MS-Agree analysis makes empirically distinct predictions regarding potential intervention effects: a downward search from the next higher phase predicts that a head matching the ST would cause intervention effects only if it is structurally higher than the target, not if it is at a lower position than the target. By contrast, the Upward Agree analysis predicts the opposite. Since probing/search is assumed to be upward, the structurally lower item that matches the target will always be found before a structurally higher target, leading to an intervention effect opposite to what MS-Agree predicts.

We have seen that S–C agreement in Lubukusu provides us a testing ground exactly for this prediction, and the fact is consistent with the prediction of MS-Agree.

(iii) It is also worth noting that Bjorkman & Zeijlstra's (2019) upward probing is in fact a mixture of upward search and downward search, if we do not assume projection or percolation. Consider the example in (35). In order to value the uninterpretable/ unvalued ϕ-features ([uϕ]) on T, Bjorkman & Zeijlstra (2019) require T to check up with the subject DP in SpecTP with regard to their ϕ-features. As shown in the structure of (35), (36), this process is upward probing.

(35) [$_{\text{DP}}$ the students of linguistics] are meeting with the department chair.

(36)



However, this is not the end of the story if we do not assume percolation or projection (which Bjorkman & Zeijlstra (2019) assume) as the DP is a set which should not bear features. The search algorithm then needs to probe downward into the DP and finds its D head, which bears interpretable valued ϕ-features that match the probe, as indicated in (36). In other words, the Upward Agree analysis will need to be a hybrid of downward and upward probing if projection or percolation is not assumed. By contrast, MS-Agree always search downward, and thus has the advantage of being a simpler and unified search algorithm.

7 Summary

In this paper, we started off by introducing a specific challenge to Chomsky's (2000, 2001) theory of Agree brought by empirical data pertinent to the direction of probing: in various cases of apparent upward agreement, probing is argued to be directed upward. The paper then argued, with the definitions of MS and MS-based Agree (MS-Agree) in Ke (2019, to appear[a]) developed for independent reasons and purposes, that all the apparent upward agreement instances can be reanalyzed as downward search, along the same lines with other attested downward agreement cases well-discussed in the literature.

This paper takes advantage of a unique feature of the definition of MS-Agree, that is, the SD and ST are separately assigned by Agree (and Labeling). This allows us to assign the next higher phase as the SD for cases where apparent upward agreement is observed. Such an MS-Agree analysis can be extended to the analysis of cyclic agreement without the need to assume that labels can probe. Critically, the MS-Agree analysis of cyclic agreement hints that we should consider MS from the next higher phase a second cyclic Agree, which occurs only after first cycle Agree cannot value the probe. Finally, I provide a direct MS-Agree analysis of the S–C agreement in Lubukusu that makes right predictions about the observed non-intervention effects of the object to the subject; by contrast, the Upward Agree analysis incorrectly predicts the opposite. Meanwhile, this direct MS-Agree analysis is free of the undesirable assumptions that characterize the indirect analysis: we do not need to project a null reflexive at SpecCP and assume it moves to SpecTP to agree with the subordinate subject, thus mediating the S–C agreement. MS-Agree thus provides a unified analysis of both instances of upward agreement and downward agreement, without running into conceptual and empirical problems of the Upward Agree analysis.

Abbreviations

1 = first person, 2 = second person, AOR = aorist, APL = applicative, DEF = definite, FV = final vowel, FUT = future, IND = indicative, INF = infinitive, M = masculine, NEG = negative, NOM = nominative, PASS = passive, PL = plural, PART = participle, PST = past, PRS = present, SA = subject agreement, SG = singular, X = gloss unclear/irrelevant.

Acknowledgements

I would like to thank Samuel Epstein, Richard Lewis, Acrisio Pires, Jonathan Brennan, Andrew McInnerney, Yushi Sugimoto, Hisatsugu Kitahara, Daniel Seely, Ezra Keshet, and the SynSem group at University of Michigan for helpful discussion and comments. Thanks also go to Michael Yoshitaka Erlewine and the anonymous reviewers of Glossa for their insightful comments and suggestions. The remaining errors are mine.

Competing Interests

The author has no competing interests to declare.

References

Adger, David. 2010. A minimalist theory of feature structure. In Anna Kibort & Greville Corbett (eds.), *Features,* 185–218. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199577743.003.0008

Akahane, Hitoshi. 2006. Inner islands: a minimalist account. *English Linguistics* 23(2). 315–346. DOI: https://doi.org/10.9793/elsj1984.23.315

Arregi, Karlos & Nevins, Andrew. 2012. *Morphotactics: Basque auxiliaries and the structure of spellout.* Vol. 86. Springer. DOI: https://doi.org/10.1007/978-94-007-3889-8

Baker, Mark C. 2008. *The syntax of agreement and concord.* Vol. 115. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511619830

Béjar, Susana. 2003. *Phi-syntax: a theory of agreement.* University of Toronto Dissertation.

Béjar, Susana & Rezac, Milan. 2003. Person licensing and the derivation of PCC effects. In Ana Teresa Pérez-Leroux & Yves Roberge (eds.), *Romance linguistics: theory and acquisition: selected papers from the 32nd Linguistic Symposium on Romance Languages (LSRL),* vol. 4 (Amsterdam Studies in the Theory and History of Linguistic Science), 49–62. Amsterdam/Philadelphia: John Benjamins. DOI: https://doi.org/10.1075/cilt.244.07bej

Béjar, Susana & Rezac, Milan. 2009. Cyclic Agree. *Linguistic Inquiry* 40(1). 35–73. DOI: https://doi.org/10.1162/ling.2009.40.1.35

Bhatt, Rajesh &Walkow, Martin. 2013. Locating agreement in grammar: an argument from agreement in conjunctions. *Natural Language & Linguistic Theory* 31(4). 951–1013. DOI: https://doi.org/10.1007/s11049-013-9203-y

Bjorkman, Bronwyn M & Zeijlstra, Hedde. 2019. Checking up on (ϕ-)agree. *Linguistic Inquiry* 50(3). 527–569. DOI: https://doi.org/10.1162/ling_a_00319

Bobaljik, Jonathan David & Wurmbrand, Susi. 2005. The domain of agreement. *Natural Language & Linguistic Theory* 23(4). 809–865. http://www.jstor.org.proxy.lib.umich.edu/stable/4048116. DOI: https://doi.org/10.1007/s11049-004-3792-4

Butler, Jonny. 2004. *Phase structure, phrase structure, and quantification*. University of York York, Great Britain dissertation.

Carstens, Vicki. 2016. Delayed valuation: a reanalysis of goal features, "upward" complementizer agreement, and the mechanics of case. *Syntax* 19(1). 1–42. DOI: https://doi.org/10.1111/synt.12116

Chomsky, Noam. 1986. *Knowledge of language: its nature, origin, and use.* New York: Praeger Publishers.

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. 2004. Beyond explanatory adequacy. In Adriana Belletti (ed.), *Structures and beyond,* vol. 3, 104–131. Oxford: Oxford University Press.

Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos P. Otero & Maria Luisa Zubizarreta (eds.), *Foundational issues in linguistic theory: essays in honor of Jean-Roger Vergnaud*, 133–166. Cambridge, MA: MIT Press. DOI: https://doi.org/10.7551/mitpress/9780262062787.003.0007

Chomsky, Noam. 2013. Problems of projection. *Lingua* 130. 33–49. DOI: https://doi.org/10.1016/j.lingua.2012.12.003

Chomsky, Noam. 2015. Problems of projection: extensions. In Elisa Di Domenico, Cornelia Hamann & Simona Matteini (eds.), *Structures, strategies and beyond: studies in honour of Adriana Belletti,* vol. 223, chap. 1, 3–16. Amsterdam/Philadelphia: John Benjamins. DOI: https://doi.org/10.1075/la.223.01cho

Cinque, Guglielmo. 1990. *Types of A-bar dependencies*. Vol. 17 (Linguistic Inquiry Monographs). Cambridge, MA: MIT Press.

Deal, Amy Rose. 2015. Interaction and satisfaction in ϕ-agreement. In Thuy Bui & Deniz Ozyildiz (eds.), *Proceedings of NELS* 45, 179–192. Amherst, MA: University of Massachusetts Graduate Linguistics Student Association.

Diercks, Michael. 2010. *Agreement with subjects in Lubukusu.* Georgetown University Dissertation.

Diercks, Michael. 2013. Indirect agree in Lubukusu complementizer agreement. *Natural Language & Linguistic Theory* 31(2). 357–407. DOI: https://doi.org/10.1007/s11049-013-9187-7

Diercks, Michael, van Koppen, Marjo & Putnam, Michael. 2020. Agree probes down: anaphoric feature valuation and phase reference. In Utrecht University Pomona College & Penn State University (eds.), *Agree to agree: agreement in the minimalist programme,* vol. Open Generative Syntax, chap. 11. Berlin: Language Science Press. DOI: https://doi.org/10.5281/zenodo.3528036

Epstein, Samuel David, Groat, Erich M., Kawashima, Ruriko & Kitahara, Hisatsugu. 1998. *A derivational approach to syntactic relations.* New York: Oxford University Press.

Epstein, Samuel David, Kitahara, Hisatsugu & Seely, T Daniel. 2014. Labeling by minimal search: implications for successive-cyclic a-movement and the conception of the postulate "phase". *Linguistic Inquiry* 45(3). 463–481. DOI: https://doi.org/10.1162/LING_a_00163

Fong, Sandiway & Ginsburg, Jason. 2019. Towards a minimalist machine: combining linguistic theories. In Robert C. Berwick & Edward P. Stabler (eds.), *Minimalist parsing,* chap. 2, 16–38. New York: Oxford University Press. DOI: https://doi.org/10.1093/oso/9780198795087.003.0002

Hewitt, Brian George. 1995. *Georgian: a structural reference grammar.* Vol. 2. Amsterdam/Philadelphia: John Benjamins. DOI: https://doi.org/10.1075/loall.2

Hicks, Glyn. 2009. *The derivation of anaphoric relations.* Vol. 139 (Linguistics Today). Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/la.139

Hiraiwa, Ken. 2001. Multiple agree and the defective intervention constraint in Japanese. In Ora Matushansky, Albert Costa, Javier Martin-Gonzalez, Lance Nathan & Adam Szczegielniak (eds.), *Proceedings of the First HUMIT Student Conference in Language Research (HUMIT 2000)*, vol. 40, 67–80. Cambridge, MA: MITWorking Papers in Linguistics.

Ke, Alan Hezao. 2019. *The syntax, semantics and processing of agreement and binding grammatical illusions.* University of Michigan, Ann Arbor Dissertation.

Ke, Alan Hezao. to appear(a). Can Agree and Labeling be reduced to Minimal Search. *Linguistic Inquiry.* DOI: https://doi.org/10.1162/ling_a_00481

Ke, Alan Hezao. to appear(b). The domain of transfer. In Kleanthes K. Grohmann & Evelina Leivada (eds.), *The Cambridge handbook of minimalism.* Cambridge: Cambridge University Press. https://ling.auf.net/lingbuzz/006103.

Keine, Stefan. 2016. *Probes and their horizons.* University of Massachusetts Amherst dissertation.

Keine, Stefan & Dash, Bhamati. 2017. Agreement reversals and the cyclicity of agree. In *Proceedings of the 48th Annual Meeting of the North East Linguistic Society (NELS 48)*. University of Iceland.

Keine, Stefan & Dash, Bhamati. 2018. The ups and downs of agreement. Manuscript. http://www-bcf.usc.edu/~keine/papers/Up_down.pdf.

Keine, Stefan & Zeijlstra, Hedde. 2021. Morphology of extraction: reassessing vp phasehood. Manuscript. https://stefankeine.com/papers/Keine-ZeijlstravP.pdf.

Kučerová, Ivona. 2018. ϕ-features at the syntax-semantics interface: evidence from nominal inflection. *Linguistic Inquiry,* 813–845. DOI: https://doi.org/10.1162/ling_a_00290

Ott, Dennis. 2012. *Local instability*. Berlin, Germany: de Gruyter. DOI: https://doi.org/10.1515/9783110290950

Polinsky, Maria & Potsdam, Eric. 2001. Long-distance agreement and topic in Tsez. *Natural Language & Linguistic Theory* 19(3). 583–646. DOI: https://doi.org/10.1023/A:1010757806504

Preminger, Omer. 2011. Asymmetries between person and number in syntax: a commentary on Baker's SCOPA. *Natural Language & Linguistic Theory* 29(4). 917–937. DOI: https://doi.org/10.1007/s11049-011-9155-z

Preminger, Omer. 2013. That's not how you agree: a reply to Zeijlstra. *The Linguistic Review* 30(3). 491–500. DOI: https://doi.org/10.1515/tlr-2013-0015

Preminger, Omer. 2014. *Agreement and its failures.* Cambridge, MA: MIT Press. DOI: https://doi.org/10.7551/mitpress/9780262027403.001.0001

Preminger, Omer. 2018. The anaphor agreement effect: further evidence against anaphoraas-agreement. Manuscript. https://ling.auf.net/lingbuzz/004401

Preminger, Omer & Polinsky, Maria. 2015. Agreement and semantic concord: a spurious unification. Manuscript. https://ling.auf.net/lingbuzz/002363.

Reuland, Eric. 2001. Primitives of binding. *Linguistic Inquiry* 32(3). 439–492. DOI: https://doi.org/10.1162/002438901750372522

Reuland, Eric. 2011. *Anaphora and language design* (Linguistic Inquiry Monographs). Cambridge, MA: MIT Press.

Rezac, Milan. 2004. *Elements of cyclic syntax: agree and merge.* University of Toronto Dissertation.

Rizzi, Luigi. 1990. On the anaphor-agreement effect. *Rivista di Linguistica* 2(1). 27–42.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of grammar: a handbook of generative syntax,* 281–337. Dordrecht: Kluwer. DOI: https://doi.org/10.1007/978-94-011-5420-8_7

Rizzi, Luigi. 2001. On the position "int(errogative)" in the left periphery of the clause. In Guglielmo Cinque & Giampaolo Salvi (eds.), *Current studies in Italian syntax: essays offered to Lorenzo Renzi,* 287–296. Amsterdam: Elsevier. DOI: https://doi.org/10.1163/9780585473949_016

Rudnev, Pavel. 2020. The anaphor agreement effect is not about featural deficiency: evidence from Avar. *Glossa: a journal of general linguistics* 5(1). DOI: https://doi.org/10.5334/gjgl.907

Rudnev, Pavel. 2021. Against upwards agree. *The Linguistic Review* 38(1). 65–99. DOI: https://doi.org/10.1515/tlr-2021-2059

Safir, Ken. 2004. *The syntax of anaphora.* New York: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780195166132.001.0001

Sells, Peter. 1987. Aspects of logophoricity. *Linguistic Inquiry* 18(3). 445–479. http://www.jstor.org/stable/4178550.

Smith, Peter W. 2017. The syntax of semantic agreement in English. *Journal of Linguistics* 53(4). 823–863. DOI: https://doi.org/10.1017/S0022226716000360

Wiklund, Anna-Lena. 2001. Dressing up for vocabulary insertion: the parasitic supine. *Natural Language & Linguistic Theory* 19(1). 199–228. DOI: https://doi.org/10.1023/A:1006417208332

Woolford, Ellen. 1999. More on the anaphor agreement effect. *Linguistic Inquiry* 30(2). 257–287. DOI: https://doi.org/10.1162/002438999554057

Wurmbrand, Susi. 2012a. Parasitic participles in Germanic: evidence for the theory of verb clusters. *Taal en Tongval* 64(1). 129–156. DOI: https://doi.org/10.5117/TET2012.1.WURM

Wurmbrand, Susi. 2012b. The syntax of valuation in auxiliary-participle constructions. In E. Jaehoon Choi, Alan Hogue, Jeffrey Punske, Deniz Tat, Jessamyn Schertz & Alex Trueman (eds.), *Coyote working papers: proceedings of the 29th West Coast Conference on Formal Linguistics (WCCFL 29),* 154–162. Cascadilla Proceedings Project.

Zeijlstra, Hedde. 2004. *Sentential negation and negative concord.* University of Amsterdam Dissertation.

Zeijlstra, Hedde. 2012. There is only one way to agree. *The Linguistic Review* 29(3). 491–539. DOI: https://doi.org/10.1515/tlr-2012-0017