Chapter 3: Allocutive marking as complementizer agreement

1. Introduction

The next crosslinguistically rare construction to consider is allocutive agreement, defined as agreement on the verb that shows features, not of the subject or some other argument of the verb, but of the person that the sentence is addressed to. The term comes from Basque linguistics, and Basque is the language in which the phenomenon has been studied most fully, with a generative literature beginning with Oyharçabal (1993). Souletian Basque has four ways to say "Peter worked", depending on who the speaker is talking to, as seen in (1).

(1) Souletian Basque (Oyharçabal 1993)

*a. Pettek lan egin di-**zü**.*
Peter.ERG work do AUX.3.ERG-2SG.H.AL
"Peter worked." (distant, formal)

*b. Pettek lan egin di-**k**.*
Peter.ERG work do AUX.3.ERG-2SG.M.AL
"Peter worked." (to a close male)

*c. Pettek lan egin di-**n**.*
Peter.ERG work do AUX.3.ERG-2SG.F.AL
"Peter worked." (to a close female)

*d. Pettek lan egin d**ü**.*
Peter.ERG work do AUX.3.ERG
"Peter worked." (not allocutive, to a group)

It is clear that this is a form of agreement in Basque, because the vocabulary items that expone it are also used for ordinary agreement with arguments. For example, the same suffixes that show allocutive

agreement in (1) are used for agreement with the second person subject in (2).

(2) Souletian Basque (Oyharçabal 1993)
*Lan egin dü-**zü** / dü-**k** / dü-**n**.*
work do AUX-2SG.H.ERG / AUX-2SG.M.ERG / AUX-2SG.F.ERG
"You (formal/familiar male/familiar female) worked."

Whereas the upward C-agreement studied in the previous chapter is found primarily in one region of the world (Central Africa), allocutive agreement is an uncommon feature found in a smattering of languages scattered around the world. Miyagawa (2012, 2017) argues that the politeness suffix *-mas-* in Japanese is a form of allocutive agreement (see also Yamada 2019), and similar morphology is found in Korean (Portner, Pak et al. 2019). Antonov's (2015) typological study adds four more languages to the list: Pumé (Venezuelan), Nambikwara (Brazilian), Mandan (Siouan), and Baja (Cushitic). Most recently, allocutive marking has turned up in a range of South Asian languages, including Jingpo (Zu 2018), Tamil (McFadden 2020), Punjabi (Kaur 2020), Maithili (Kumari 2022), and Magahi (Alok 2020, Alok 2021).[1]

Allocutive marking in Magahi is of special interest because its system is particularly rich in some respects, and much is known about it from Alok's published work and our research together. Magahi is an Indo-Aryan language of Northeastern India, closely related to Hindi. Like Basque, it has four ways to say 'I am coming', as shown in (3).

(3) Magahi (fieldwork, Deepak Alok)
*a. Ham jaa-it h-i-**au**.*

[1] My working notion of allocutivity here is a little bit narrower than that of, say, Alok & Haddican's (2022) survey of the phenomenon. I focus on those languages which plausibly have agreement on some head with a null pronoun (Ad) in the left periphery. This potentially excludes languages in which Ad itself is an overt pronoun cliticized to the verb (Gallician—and also Basque, on Haddican's view) and languages where what is exponed is not agreement but an applicative-like head that selects Ad (e.g. Japanese, Korean). I mean to leave open exactly how much the latter two kinds of constructions have in common with the first. However, Oyharçabal's work on Basque and Miyagawa's on Japanese are included here as full partners because they are couched in terms of agreement and they have had a founding influence on the allocutivity literature.

I go-PROG be-1SG-NH.AL
"I am going." (nonhonorific, to a peer)

*b. Ham jaa-it h-i-**o**.*
I go-PROG be-1SG-H.AL
"I am going." (honorific, to an elder)

*c. Ham jaa-it h-i-**ain**.*
I go-PROG be-1SG-HH.AL
"I am going." (high honorific, to a community leader)

d. Ham jaa-it h-i.
I go-PROG be-1SG
"I am going." (not allocutive, sayable to anyone)

In (3d), there is no allocutive marking, only ordinary subject agreement, and this version can be said to anyone. The other three versions have an additional suffix that records the social status of the addressee relative to the speaker, according to a three-way system: nonhonorific (NH), for someone of the same social standing as the speaker or lower (one's friend, one's same-age cousin, one's child); honorific (H), for someone of higher social standing than the speaker (one's parent, grandparent or older brother); high honorific (HH), for someone of much higher social standing than the speaker (a teacher, a priest, or a local ruler). Although honorific features like these may seem like strange things to agree with from the English perspective, Alok (2020, 2021) shows that this is normal within Magahi: subject-verb agreement also reflects the subject's person and honorific status using the same three-way distinctions (but not the subject's gender or number; see Alok (2021) for a full paradigm and Alok and Baker (2022) for further discussion of the honorific features in Magahi). This chapter is in part a reflection on Alok's work, considering how new data from Magahi allows us to place the general phenomenon of allocutive agreement within the current overall framework.

Every generative study of allocutive agreement to date assumes that it is the result of a functional head agreeing with a syntactically represented null DP that denotes the addressee. Oyharçabal's (1993) early study calls the null DP e_{ALLOC} and argues that it is a variable, the trace of an operator that moves to Spec CP. Since Speas & Tenny's (2003) introduction of the neo-performative hypothesis, generative linguists have mostly called the null addressee-denoting DP *Hr*

(hearer) or *Ad* (addressee); I use Ad in this work.[2] Indeed, one of Speas & Tenny's initial motivations for positing this ghostly DP comes from a very limited allocutive pattern in the West Chadic language Mupun.[3] This general approach is adopted with some minor differences by Miyagawa, Zu, McFadden, Haddican, and Alok, among others, and I adopt it as well. A starting point for analyzing a sentence like (3a), then, is assigning it a structure like (4).

(4) $[_{HP}$ Ad H $[_{TP}$ I T … $[_{VP}$ go]]]
Agree Agree

This puts allocutive marking in the same conceptual domain as upward C-agreement in the African languages, in that there is agreement with a ghostly DP in the periphery of the clause.

A question is whether this qualifies as *complementizer* agreement, which would bring it one step closer to the African languages. In other words, is H in (4) a C-type head, and if so which one? My answer is yes, although H may not be the same head in the C-space in all languages. In Magahi (and also Tamil), whenever T-agreement and allocutive agreement are distinct morphemes, the allocutive agreement is always farther from the verb root than tense marking and subject agreement, as seen in (3). By Mirror Principle-style reasoning (Baker 1985, Cinque 1999), this suggests that allocutive agreement is on a head above T, putting it within the C-space. Speas & Tenny (2003) claim that Ad is the argument of an sa ("speech act") head, very high in the left periphery of the clause. saP is present in root clauses but if it can be selected at all, it is only by a limited number of verbs of saying. Miyagawa (2012, 2017) derives from this that only 'say' verbs can have allocutive *mas* in their CP complements in Japanese; this is a subset of the verbs that allow embedded root phenomena in a variety of languages (see also Zu (2018) on Jingpo, Portner et al. (2019) on Korean, Alok and Haddican (2022) on Punjabi).

---

[2] I use Ad, following Bill Haddican, because this label is a bit more accurate. People may very well hear a sentence that is not addressed to them, but the gender or status of such people is not reflected in allocutive marking.

[3] The Mupun pattern is limited in that the plural addressee marker *nuwa* is used only in interrogatives, not in root declarative sentences. I do not consider it here.

However, a distinctive feature of Magahi is that allocutive marking is freely available in all kinds of embedded clauses. (5) gives examples of allocutive marking in CP complements, where the embedded verb bears the same allocutive marking as the matrix verb.

(5) Magahi (fieldwork, Deepak Alok)

*a. Santee-aa sochk-**au** ki Bantee-aa bhag ge-l-**au**.*
Santee-FM thought.3.NH.S-NH.AL that Bantee- FM run go-PFV.3.HN.S-NH.AL
"Santee thought that Bantee went to run." (to a peer)

*b. Santee-aa sochk-**o** ki Bantee-aa bhag ge-l-**o**.*
Santee-FM thought.3.NH.S-H.AL that Bantee-FM run go-PFV.3.HN.S-H.AL
"Santee thought that Bantee went to run." (to a parent)

*c. Santee-aa sochk-**ain** ki Bantee-aa bhag ge-l-**ain**.*
Santee-FM thought.3.NH.S-HH.AL that Bantee-FM run go-PFV.3.HN.S-HH.AL
"Santee thought that Bantee went to run." (to a teacher)

Allocutive marking is also possible in embedded clauses to some extent in Tamil (McFadden 2020) and quite freely in some innovative Southern Basque dialects (Antonov 2015, Haddican and Etxeberria 2022). Antonov (2015) observes that it is more common for allocutive marking to be unembeddable—this is also the case in the four less-studied languages that he surveys—but this is not a universal feature of allocutivity. In Magahi, not only is allocutive marking found inside an embedded CP, it is also compatible with the ordinary declarative complementizer *ki*.[4] Indeed, it seems to be hosted on a head that is lower than *ki*, in that it affixes to the verb by a process of head movement or local dislocation that is not blocked by the presence of *ki*. Alok (2020: §3.2) shows that allocutive marking is also compatible with question particles, relative complementizers, purposive complementizers, and all other known material from the high left periphery in Magahi (different from Basque). Alok thus concludes that

[4] As in Hindi, there is some question whether *ki* is really a complementizer in Magahi (see Alok 2020 for some discussion). However, I assume that it is a C-type head (tentatively Force) for simplicity/familiarity.

the head that agrees with Ad in Magahi is Fin—the lowest head in the C-space according to Rizzi (1997), and the head right above T. I adopt this view as well, returning to some refinements and the possibility of crosslinguistic variation in §3.2.1. On this hypothesis, allocutive marking does indeed count as a form of C-agreement.

The fact about Magahi allocutive marking that links it most firmly to upward C-agreement in the African languages is shown in (6). Here again there is allocutive marking on the verb inside a CP complement. But unlike in (5), the embedded allocutive marking in (6) does not repeat the allocutive marking in the matrix clause; rather, it covaries with the goal of the matrix verb 'tell'. When the matrix goal is 'Bantee', the embedded allocutive marking is NH; when the matrix goal is 'grandfather' the embedded allocutive is H; when the matrix goal is 'professor', the embedded allocutive is HH.[5] In each case, this is different from the allocutive marking on the matrix verb.

(6) Magahi (fieldwork, Deepak Alok)

*a. Santee-aa **Bantee-aa-ke** kahk-ain ki Ram-ke Sita-se baat kar-e-ke chah-**au**.*
Santee-FM Bantee-FM-DAT told.3.NH.S-HH.AL that Ram-DAT Sita-INS talk do-INF-DAT should-NH.AL
"Santee told Bantee that Ram should talk to Sita." (to a teacher)

*b. Santee-aa **baabaa-ke** kahk-au ki Ram Sita-ke dekh-l-**o** ha-l.*
Santee-FM gr'father-DAT told.3.NH.S-NH.AL that Ram Sita-ACC see-PFV.3.NH.S-H.AL be-PFV
"Santee told grandfather that Ram saw Sita." (to a peer)

*c. Santee-aa **profesar saaheb-ke** kahk-au ki Ram Sita-ke dekh-l-**ain** ha-l.*
Santee-FM professor HH-DAT told.3.NH.S-NH.AL that Ram Sita-ACC see-PRF.3.NH.S-HH.AL be-PFV

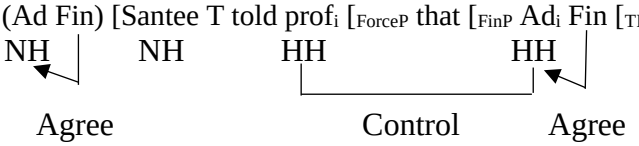
---

[5] More precisely, the honorificity marking on the embedded verb in examples like (6) is determined not just by the goal argument of the matrix verb but by the social relationship of the referent of the goal argument of 'tell' to the referent of the agent argument of 'tell'—which is held constant as Santee in (6). Alok and Baker (2022) derive this from the semantics of the honorific features in Magahi, which is intrinsically relational and sensitive to the closest Sp.

"Santee told the professor that Ram saw Sita." (to a peer)

This looks very much like upward C-agreement, in that a head in the lower clause agrees in honorificity with an argument of the matrix clause. In fact, it looks the most like upward object agreement in Kipsigis, where C shows suffixal agreement with the goal argument of the matrix clause. I thus propose a parallel analysis: not only can a C-type head agree with Ad, but Ad can be controlled by an argument of the matrix clause. This results in a structure like (7) for (6c).

(7) (Ad Fin) [Santee T told prof$_i$ [$_{ForceP}$ that [$_{FinP}$ Ad$_i$ Fin [$_{TP}$ Ram saw...]]].



McFadden (2020) reports similar possibilities in Tamil. In (8a), the allocutive marker *-ŋgæ* is understood as showing the respect of the speaker of the sentence as a whole for the addressee of the sentence as a whole, as *-ain* does in (5c) in Mahagi. In (8b), the same allocutive marker *-ŋgæ* can only be understood as showing the respect that Maya, the referent of the subject of the matrix clause, has for the person she is talking to, as *-ain* does in (6c) in Magahi.[6]

(8) Tamil (McFadden 2020: (17), (19))

*a. Maya [avæ pooʈʈi-le ʤejkkæ-poo-r-aaɭ-**ŋgæ**-nnŭ] so-nn-aa.*
Maya she contest-LOC win-go-PRS-3SG.F.S-H.AL-C say-PST-3SG.F.S
"Maya said that she would win the contest." (the speaker honors their addressee)

b. Maya [taan pooʈʈi-le ʤejkkæ-poo-r-een-ŋgæ-nnŭ] so-nn-aa.
Maya self contest-LOC win-go-PRS-1SG.S-H.AL-C say-

[6] The interpretation of allocutive marking on the embedded verb in (8) is correlated with what devices are used to have the embedded subject refer to the matrix subject. (8a) expresses this with an ordinary third person pronoun, whereas (8b) does it with an anaphoric element *taan* that triggers "monstrous" first person agreement on the embedded verb. This is parallel to the fact that shifted allocutive marking in Magahi goes hand in hand with the use of shifted indexicals to refer to arguments of the matrix verb, as discussed in Chapter 4. On monstrous agreement in the Dravidian languages, see Chapter 6.

PST-3SG.F.S
"Maya said that she would win the contest." (Maya honors her addressee)

The upshot is that all of the key ingredients involved in upward C-agreement are also at work in allocutive marking in some languages. This motivates a unified analysis, which I pursue in this chapter.

It should be kept in mind, however, that the control relationship shown in (7) is not required, either internally to Magahi and Tamil, or across languages. Haddican and Etxebarria (2022) show that in the Basque dialects that allow embedded allocutive marking, that marking matches the allocutivity on the matrix verb. For example, (9a) is possible, with the same masculine singular familiar marking *-k* on both the matrix verb and the embedded verb, showing the features of the addressee of the sentence as a whole (compare (5) in Magahi). However, (9b) is not, where the embedded verb has feminine marking *-na*, which would reflect how Jon would address his female addressee Miren. This contrasts with (6) in Magahi and (8b) in Tamil, showing that embedded Ad cannot be controlled in Basque.

(9) Basque (Haddican and Etxebarria 2022; Etxebarria p.c.)

*a. Jon-ek Imanol-i [etorri-ko du-**k**-ela] esa-n zi-o-**k**.*
Jon-ERG Imanol-DAT come-FUT AUX-2.SG.M.AL-C say-PRF AUX-3.SG.DAT-2.SG.M.AL
"Jon told Imanol that he will come." (to a male friend)

*b. *Jon-ek Miren-i [etorri-ko du-**na**-la] esa-n z-i-o-**k**.*
Jon-ERG Miren-DAT come-FUT AUX-2.SG.F.AL-C say-PRF AUX-3.SG.DAT-2.SG.M.AL
"Jon told Miren that he will come." (to a male friend)

In contrast, SoK in the African languages cannot fail to be controlled by a matrix argument. The possibility of (9a), (8a), and (5) alongside (8b) and (6) raises questions about the obligatoriness of "obligatory control" in this domain. There are thus some differences between upward C-agreement and allocutive marking to be considered as well.

My specific hypothesis in this chapter is that Ad is very similar to SoK in the African languages, in that it participates in the same relationships of licensing, agreement, and control. There are two significant differences. First, it receives a goal-like thematic role from C rather than an agent-like one, which affects which matrix argument

controls it (like OoK in Kipsigis). Second, it has inherent second person features whereas the African ghostly DPs are unspecified for phi-features until they receive them from their controller. I discuss the status of Ad in §3.2, arguing that it is an noninterrogative A-bar specifier and that it has intrinsic second person features. In section §3.3, I extend the framework to claim that there is also a first person ghostly operator Sp, and that in some languages C can agree with Sp, resulting in speaker agreement rather than addressee (allocutive) agreement. In this section, I also discuss why the T/Agree Condition applies to C agreeing with SoK in the Niger-Congo languages but not to C agreeing with Ad or Sp; the difference follows from the fact that Ad and Sp have intrinsic person features. In section §3.4, I give evidence that Ad is controlled by the same principles as OoK and SoK are; in particular, here too control is subject to the GOCS and the condition on thematic matching. At the end of this section (§3.4.5), I argue that the apparent optionality in the control of Ad in Magahi and Tamil follows from the fact that CP complements are extraposed in these languages, meaning that they are associated with both a low, VP-internal position (where control happens) and a high adjunct position (where control is blocked). This adds up to a relatively comprehensive analysis of the allocutive agreement phenomenon.

2. Properties of the Ghostly Ad

2.1. Ad is an A-bar specifier

I start by being more precise about what exactly Ad is. It has many similarities with the SoK of African languages (and even more with OoK) in that it is a necessarily null pronominal element which occupies an A-bar position in the periphery of a finite clause, licensed by a C-type head. There is one primary difference: Ad has intrinsic second person features, whereas SoK (and OoK) have no intrinsic phi-features. Several consequences follow from this primary difference.

I outline the similarities first. That Ad like SoK is pronominal, necessarily null, and a possible target of Agree is obvious by inspection plus hypothesis. As null pronouns, both can be controlled by a nominal in the higher clause and both can bind pronouns in the TP complement of the C that licenses them (see below for more).

Less obvious is the claim that Ad, like SoK, is licensed by a head in the C space. Some indirect evidence for this comes from the fact, mentioned in Chapter 1, that allocutive marking is not possible on

infinitival and gerundival verbs in Magahi, as shown in (10).

(10) Magahi (fieldwork, Deepak Alok)

*a. Santee-aa jaa-yel-(**au/*o/*ain**) chaha h-au/o/ain.*
Santee-FM go-INF-(NH.AL/H.AL/HH.AL) want be.3.NH-NH.AL/H.AL/HH.AL
"Santee wants to go."

*b. Ham okaraa **dhekhe**-se bach-l-i-ain/o/au.*
I 3SG.NH.ACC see.PTCP-INS avoid-PFV-1SG-HH.AL/H.AL/NH.AL
"I avoided seeing him." (no alternative forms of *dhekhe-se*)

This is parallel to the fact that an agreeing complementizer cannot appear with a verb in gerund or infinitival form in Ibibio:

(11) Ibibio (fieldwork, Willie Willie)

*a. Okon a-maa-tre [(*a-bo) u-koot ngwet].*
Okon 3SG-PST-stop 3SG-C GER-read book
"Okon stopped reading the book."

*b. Nditọ e-ke-yem (*e-bo) edi-ta ebot.*
children 3PL-PST-want (3.PL-C) INF-eat goat
"The children want to eat goat."

Nominalized clauses like (10b) and (11a) do not have complementizer heads; if anything, they have a DP-type superstructure. From this plus the hypothesis that Ad like SoK is licensed by a head in the C-space, it follows that allocutive marking is possible on the nonmatrix verbal element in (5) and (6) but not in (10b). Infinitival clauses may not have C-heads either (especially in restructuring type constructions), and if they do, they are different C-heads from the ones in finite clauses (e.g., *for* versus *that* in English). It is thus plausible that the C-head that licenses Ad in (5) and (6) is not there to do so in (10a).

The fact that Ad like SoK is licensed by a C-space head strongly suggests that it is in an A-bar position rather than an A-position. This predicts that Ad like SoK cannot be the antecedent of an anaphor inside the associated clause. Oyharçabal (1993) originally pointed this out for Basque, using examples like (12). Here the allocutive marking on the auxiliary verb shows that Ad is present triggering agreement that surfaces on the verb. Nethertheless, it is not possible for there to be a reflexive anaphor inside the sentence that is bound by this Ad.

(12) Basque (Oyharçabal 1993)

**Hire burua-rekin mintzatzen nau-k.*
you self-SOC speaking AUX.1.ABS-M.AL
("I am speaking to yourself.")

This is also true in Magahi. In (13), shifted allocutive marking shows that 'grandfather' controls Ad in the embedded clause. However, this does not allow reflexive *apan* to refer to the grandfather by virtue of being locally bound by Ad.

(13) Magahi (fieldwork, Deepak Alok)

*a. *Santee-aa baabaa-ke kahk-au ki apan Ram-ke dekh-l-o ha-l.*
Santee-FM gr'father-DAT told.3.NH.S-NH.AL that self Ram-ACC see-PFV.3.NH.S-H.AL be-PFV
("Santee told grandfather$_i$ that himself$_i$ saw Ram.")

b. #Santee-aa baabaa-ke kahk-au ki Ram apan dekh-l-o ha-l.
Santee-FM gr'father-DAT told.3.NH.S-NH.AL that Ram self see-PFV.3.NH.S-H.AL be-PFV
"Santee told grandfather$_i$ that Ram$_k$ saw himself$_{k,*i}$."

The examples in (14) make the analogous point for SoK in Ibibio. The presence of *a-te* or *a-bo* shows that SoK is present in the CP complement, controlled by the matrix subject *Okon*. However, it is still impossible for a local anaphor (lit. 'body his') in the embedded clause to refer to Okon by virtue of being bound by SoK.

(14) Ibibio (Afranaph D1a´´, fieldwork)

*a. *Okon á-ké-bó a-te ké idém ọ́mọ̀ ọ́-sọ́p-idém.*
Okon 3SG-PST-say 3SG-C that body his 3SG-be.fast-body
("Okon$_i$ said that himself$_i$ is smart.")

b. #Okon á-diọ́ngọ (a-bo) ké Edem á-ma idem ọ́mọ̀.
Okon 3SG-know 3SG-C that Edem 3SG-like body his
"Okon$_i$ knows that Edem$_k$ likes himself$_{k,*i}$."

Although these ghostly DPs are similar to each other in this respect, they differ from the logophoric operator which Charnavel (2019, 2020) posits in constructions with long-distance anaphors in many languages; these can bind an anaphor giving the appearance of

genuine long-distance behavior. See Chapter 5 for discussion.[7]

Although Ad like SoK occupies an A-bar position in the CP region of the clause, Ad like SoK does not create an island for extracting a *wh*-phrase out of a CP complement. (15) shows that shifted allocutive marking is possible in a clause from which 'when' has been extracted in Magahi.[8] (The ? here is because 'when' is more easily interpreted as questioning the time of the telling, as in English, but it can also be interpreted as questioning the time of dying.) Similarly, (16) shows that 'what' can be moved out of a CP headed by the agreeing complementizer in Kinande, and (17) shows the same thing for Ibibo.

(15) Magahi (fieldwork, Deepak Alok)
*?Kab Santee-aa baabaa-ke kahk-ai/au ki Ram mar-b-**o**?*
when Santee-FM gr'father-DAT told-3.NH.S/NH.AL that Ram die-FUT.3.NH.S-H.AL
"When did Santee tell grandfather [that Ram will die --]?" (to a friend)

(16) Kinande (fieldwork, Philip Mutaka)
*Eki-hi ky-o Kambale a-bw-ira aba-kali **a-ti** ba-gul-e?*
CL7-what CL7-FOC CL1.Kamable CL1.S-tell-APPL CL2-women CL1-C CL2.S-buy-SBJV
"What did Kambale tell the women that they should buy?"

(17) Ibibio (fieldwork, Willie Willie)
*Nso ke Okon a-ke-dọkkọ Emem (**a-bo/a-te**) ke imọ i-k-i-dep?*
what FOC Okon 3SG-PST-tell Emem 3SG-C/3SG-C that LOG 3.LOG-PST-3.LOG-buy
"What did Okon tell Emem that he bought?"

Thus, ghostly DPs do not create *wh*-islands the way that moved *wh*-phrases in the Spec CP of a complement clause often do. In terms of phases, these null DPs do not clog up the edge of the CP phase so that

---

[7] Another factor in ruling out anaphor binding in (12)-(14) could be that Ad and SoK are outside of the TP, so arguably not within the binding domain of the anaphor.

[8] This also shows that apparently shifted allocutive marking in Magahi is not simply the result of direct quotation. See §4,2 for more discussion in the context of indexical shift.

there is no escape hatch for the moved *wh*-phrases. In relativized minimality terms, the ghostly DPs must count as being in a different equivalence class from ordinary *wh*-phrases. This is not surprising, given that moved *wh*-phrases typically bear focus and interrogative features, and there is no reason to think that SoK or Ad has that kind of feature. Overall, then, both types of ghost are invisible to some of standard syntactic diagnostics, drawn from the theories of movement and binding; this is why I call them ghostly DP (operators) in the first place. But they are certainly not invisible to all syntactic processes. For example, they are not invisible to agreement or control.

Ad and SoK are also unlike *wh*-phrases and certain null operators in that they do not need to bind a variable inside the clause headed by the C that licenses them. Generating random *wh*-phrases in the CP periphery that do not bind a trace or a pronoun violates the ban on vacuous quantification (**I wonder* ***what*** *Sarah will buy a car*). But this is clearly not true for Ad or SoK. There is no 'you' inside the clause bound by Ad in (1) from Basque or (3) from Magahi, for example. Nor is there any pronoun or gap inside the embedded TP that refers to Kambale the way that SoK does in (16) from Kinande. I sometimes refer to Ad and SoK as ghostly DP *operators*, but "operator" here only means that they are DPs in A-bar positions; it does not imply that they must bind variables in A-positions.

Before moving on, can I say any more about the head in the C-space that licenses Ad? Yes, to some extent. The nullness of Ad creates an intrinsic challenge to locating it precisely, as does the possible nullness of some heads in the clausal periphery. But one thing that is clear and worth saying is that the licenser of Ad is not Eval, the head that licenses SoK (and perhaps OoK). The evidence for this is that the presence-versus-absence of allocutive marking does not have the same kind of semantic consequences that the presence-versus-absence of upward C-agreement has in Kinande and other languages, as documented in §2.2. Allocutive marking is always optional in Magahi, in the sense that (3d) is possible alongside (3a-c) in matrix clauses. This optionality is found in embedded clauses too: in (18) the embedded verb can be inflected with *au* showing that the addressee is nonhonorific, or just with *ai*, which is simple third person (nonhonorific) subject agreement.[9] This is independent of whether the

[9] I do not have any interesting account as to why allocutive agreement is optional in Magahi. Languages vary in this respect. Oyharçabal (1993) says that allocutive

matrix verb ‘think’ itself bears an allocutive marker or not.

(18) Magahi (fieldwork, Deepak Alok)
*Santee-aa sochk-ai/au ki Bantee-aa bhag ge-l-**ai/au***.
Santee-FM thought-3.NH.S/NH.AL that Bantee-FM run go-PRF-3.NH.S/NH.AL
“Santee thought that Bantee went to run.” (said to a peer)

This is superficially similar to the fact that in Kinande the complement of ‘think’ can have an agreeing C or not. But in Kinande, the choice makes a difference in who is taken to be the source of the content of the embedded clause and who is committed to the content: is it the subject of the matrix clause, or everyone in the context, or no one? There is no such difference in Magahi: either version of (18) is compatible with situations in which Santee heard a rumor that Bantee went to run, or with those in which the idea arose from his own irrational fears, or with those in which it is common knowledge. Similarly, there are no known verbs in Magahi which require allocutive marking on the verb in the embedded clause, and there are no known verbs that forbid it, the way there are for Agr-C in Kinande and Lubukusu. For example, factive verbs like ‘remember’/’remind’ do not allow upward C-agreement in Kinande, but they do allow

---

marking is obligatory in the variety of Basque that he concentrates on, as does McFadden (2020: (16)) for the Tamil spoken by his primary consultant. In contrast, allocutive agreement is optional in Jingpo (Zu 2018), Japanese, Gallician, and Punjabi (Alok & Haddican 2022), as well as Magahi. The fact that the presence-versus-absence of allocutive marking does not make a semantic difference suggests that it is not a matter of whether Ad is present in the representation of the clause or not. This is patently true if second person pronouns need to be bound by Ad, as I argue in Chapter 4, since even the presence of such a pronoun in the clause does not make allocutive marking obligatory in Magahi. This implies that the optionality is in whether or not a C-type head agrees with the Ad that is always there. Various ways of stipulating the optionality of allocutive agreement in Magahi are possible and have been tried by Alok and/or me, but none of them seems obviously better than the others. Alok & Baker (2018) said that Fin agrees with Ad only if (V+)T moves to Fin, and that movement is optional; see Zu (2018) for a similar view on the optionality of allocutive agreement in Jingpo. However, we had no independent evidence for this optional head movement. Alok (2020, 2021) simply says that it is optional to generate an Agree probe on Fin in Magahi. This is descriptively adequate, but unremarkable from an explanatory perspective. This is what I adopt for concreteness, leaving open the possibility of a deeper account.

allocutive marking in Magahi (see also Alok 2020: 96-97).

(19) Magahi (fieldwork, Deepak Alok)
*Santee-aa baabaa-ke yaad dial-k-**au** ki Ram almira-me paisa chhupail-**o** ha-l.*
Santee-FM gr'father-DAT memory give-PFV.3.NH.S-NH.AL that Ram drawer-in money hid-PFV.3.NH.S-H.AL be-PFV
"Santee reminded Grandfather that Ram hid money in the drawer." (to a friend).

Indeed, it is hard to detect any meaning contribution from the presence of Ad in Magahi at all. This shows that it is not licensed by a C-type head with a distinctive lexical semantics, such as Eval.

Details from McFadden's (2020) study of allocutive marking in Tamil suggest that Ad can actually appear in two different places in the clausal periphery across languages. McFadden shows that allocutive marking can co-occur with overt C-heads in Tamil (like Magahi, but unlike Basque). One such C-head is the question particle *aa* in root clauses. Interestingly, allocutive marking can show up either outside this head or inside it, or even in both places at once, as in (20).[10]

(20) Tamil (McFadden 2020: (23a), (23b), (25b))
*a. Niiŋgæ saap-ʈ-aačč-aa-**ŋgæ**?*
you.PL eat-ASP-RES-Q-AL
"Have you eaten?" (the speaker expects that the addressee has)

*b. Niiŋgæ saap- ʈ-aačč̆-**ŋgæ**ɭ-aa?*
you.PL eat-ASP-RES-AL-Q
"Have you eaten?" (the speaker doesn't know)

*c. Niiŋgæ saap-ʈ-aačč̆-**ŋgæ**ɭ-aa-**ŋgæ**?*
you.PL eat-ASP-RES-AL-Q-AL
"Have you eaten?" (the speaker expects that the addressee has)

[10] McFadden (2020) reports that the difference in morpheme order may go along with some differences in the bias of the question: Q-ALLOC order goes with an expectation that the answer will be yes and is seeking confirmation, whereas ALLOC-Q order is more neutral in this respect. But he treats these observations as tentative and does not include them in his formal analysis. Neither will I.

If C agreement with ghostly DPs is very local, this suggests that a full root clause in Tamil might have two Ad positions, with the higher one binding the lower one.[11] If *aa* is a Force head in Tamil, as suggested by its clause-typing meaning, then one Ad would be higher than Force, in saP, a speech act projection. The other Ad would be lower than Force, say in FinP, the lowest C head right above the T space in Rizzi's (1997) map of the left periphery. This gives a structure like (21) for the examples in (20). I refer to the special Ad in Spec saP of a root clause as Ad*.

(21) $[_{saP}$ Ad^*_i sa $[_{ForceP}$ Q $[_{FinP}$ Ad_i Fin $[_{TP}$ you $_i$ T eat]]]]
+Hon (Agree) +Hon (Agree) (Agree)

Languages can then vary as to whether overt allocutive marking is the spell out of an agreeing sa head or the spell out of an agreeing Fin head or both (see also Alok 2021, Alok and Haddican 2022).[12]

The other important fact that relates to this is whether allocutive marking can happen in embedded clauses in a given language or not.

---

[11] This is similar to McFadden's own proposal, but he suggests only one Ad, in the higher Spec SAP position, which a lower head (which he calls AllAgr) can agree with at a distance. McFadden's version is more parsimonious in having only a single Ad representing the addressee syntactically, but it has a problem with unshifted allocutive marking in the embedded clause in (8a)/(22a). Here the embedded Fin/AllAgr head is too far from the root Ad in SAP to agree with it directly. In contrast, the root Ad* can bind the embedded Ad in Spec FinP without regard for phasal boundaries, as is normal for pronoun binding, and the embedded Fin can agree with Ad in its Spec in a very local manner.

[12] In Magahi, Basque, and Tamil, allocutive agreement is semi-independent from normal agreement on T with the subject, in that both can be present on the same finite verb. Therefore, it makes sense to posit two distinct probing heads and the main question is which head is it that agrees with Ad. This is what is sketched in (21). Somewhat different from this are Jingpo (Zu 2018) and Punjabi (Kaur 2020), in which the finite verb agrees with either the subject or Ad but not both, depending on whether the subject has participant features or not (and, in Jingpo, on discourse factors). Kaur (2020) plausibly analyzes this in terms of Cyclic Agree. A single probing head looks for a [+participant] goal. First it probes downward and agrees with a first or second person subject, if there is one. If there is not, it can probe upward instead and agree with Ad. This results in allocutive marking only if the subject is third person (or in oblique case). (Zu uses T-to-sa movement, rather than Cyclic Agree, to get a similar effect.)

As mentioned above, typological comparison reveals that this is possible in some languages but not others. The first languages in which allocutivity was studied seriously either forbid allocutive marking in embedded clauses (Basque, also Jingpo, Zu 2018) or strongly limit it (Japanese, Miyagawa 2012, 2017). However, more recent research has shown that embedded allocutivity is possible in some languages, and Magahi and Tamil are part of this trend. In the limited data that McFadden (2020) gives about embedded allocutive marking in Tamil, the allocutive marker shows up only once, inside the C head, as in (22) (repeated from (8)). (Note that the morpheme order is the same whether Ad is controlled by the goal or not).

(22) Tamil (McFadden 2020: (17), (19))

*a. Maya [avæ pooʈʈi-le ʤejkkæ-poo-r-aaɭ-**ŋgæ**-nnŭ] so-nn-aa.*
Maya she contest-LOC win-go-PRS-3SG.F.S-H.AL-C say-PST-3SG.F.S
"Maya said that she would win the contest." (speaker honors their addressee)

*b. Maya [taan pooʈʈi-le ʤejkkæ-poo-r-een-**ŋgæ**-nnŭ] so-nn-aa.*
Maya self contest-LOC win-go-PRS-1SG.S-H.AL-C say-PST-3SG.F.S
"Maya said that she would win the contest." (Maya honors her addressee)

This fits well with the structure in (21) together with the widespread view that saP (or cP, in the terms of Portner et al. 2019) either cannot be selected by a verb at all, because of its special speech act semantics, or it can only be the complement of 'say' and a few other verbs that can license embedded root phenomena (Miyagawa 2012, Sundaresan 2012, Miyagawa 2017, Zu 2018, Portner, Pak et al. 2019). The largest phrase that can be routinely selected by verbs is a ForceP. This fits the fact that only the inner allocutive marker is possible in Tamil, not the outer one. Conservative Basque fits in as a language in which allocutive marking is overt agreement on the sa head, so it is impossible in embedded clauses. Miyagawa's (2012, 2017) Japanese is similar, with allocutive *mas* a realization of agreement on sa, which a restricted class of 'say' type verbs can select (but see Yamada (2019) for a different view). Similarly, in Korean, allocutive marking is packaged together with speech-act/clause-typing morphology; as a result, it is strictly the final morpheme in root clauses and is not embeddable (Portner et al. 2019; Alok and Haddican 2022).

Magahi then takes its place as a language in which the agreement on Fin is overt, but the agreement on sa is not. Embedded allocutive marking is freely possible in Magahi because FinP is easily embeddable—as the complement of a wide range of verbs, and also in adjunct clauses, relative clauses, and indeed anywhere a finite clause can be. This is also consistent with where allocutive marking is realized morphologically, in that morphemes like *-au* (NH.AL) show up inside of C, attached to the verb rather than to complementizer like *ki*, in both matrix and embedded clauses.[13] This gets the major variants that we know about at this point.[14]

In summary, Ad is a null pronoun generated in an A-bar position but without focus or interrogative features, which does not need to bind a variable. It is licensed by one of two functional heads: sa or Fin. If it is licensed only by sa in a given language, then allocutive marking is limited to root clauses. If it is licensed also by Fin, then it can appear in embedded clauses as well. Whereas SoK is licensed by Eval, a head with a distinctive semantics, embedded Ad is licensed by Fin, a head with little distinctive semantics. Although (21) is my fullest official structure for a root allocutive example, I often abbreviate it by showing only one Ad in the specifier of an unspecified C head to simplify the representation.

---

[13] There are complications to this simple generalization in constructions that contain auxiliary verbs. In some such constructions, both subject agreement and allocutive agreement show up on the auxiliary verb, as expected; in others, both agreements show up on the main verb (unexpected; the embedded clause of (19) is one of many examples). One might think that this is evidence that allocutive marking must be on a head lower than Fin, but Alok (2020: §3.5.2) argues that it is the result of PF factors influencing where the relevant morphology is realized.

14 In Basque, the position of the allocutive marker does not give information about what head is undergoing Agree with Ad, sa or Fin. This makes some sense if allocutive marking in Basque is really is a kind of clitic that gets attracted to the finite auxiliary and then is ordered by the complex templatic rules of clitic clustering in Basque (Haddican 2018). This would imply that Basque language learners cannot reliably infer from morpheme order where the clitic originated, whether in Spec saP or Spec FinP. This could open it up to be a point of dialectal variation in Basque as to whether allocutive marking is allowed in embedded clauses—no in conservative northern varieties but yes in innovative southern varieties. See Anton (2015), Haddican & Etxeberria (2022).

2.2. Ad is a second person pronoun

In most of the respects discussed so far, Ad is minimally different from SoK (and OoK); it is another instance of the same kind of thing. There is one big difference between them, however: Ad has intrinsic second person features, whereas SoK (and OoK) have no intrinsic phi-features. I hypothesize that this is the only significant difference between Ad and SoK/OoK other than the difference in exactly which head in the C-space licenses them.[15]

The most obvious reason for saying that Ad is second person is that in a simple root clauses it (Ad*) denotes the addressee, just as ordinary second person pronouns do. Allocutive marking is, after all, *addressee* agreement. This sameness of referential possibilities makes sense if there is also a sameness of person features.

A second reason for saying that Ad is second person comes from the morphological details of some of these languages. We have already seen in (1) and (2) that Basque allocutive agreement uses some of the same vocabulary items as subject agreement with second person pronouns. This is also the case in Jingpo (Zu 2018: 4 (6), (7)). Magahi's verbal paradigms are less immediately revealing in that its allocutive affixes are for the most part unique, not homophonous with any affixes seen in the subject agreement paradigm.[16] However, it is

---

[15] Ideally, I would like to take this a step farther, seeking to explain the difference in phi-features in terms of the differences in the functional heads that license the ghostly DPs. We know from Zanuttini (2008) and Zanuttini et al. (2012) that the imperative head imposes second person features on its subject. I would like to say that, in a similar way, sa imposes [+2] on its specifier, by virtue of what sa means. Eval has a different meaning, so it does not require this feature on its licensee SoK—just as T-space heads other than imperative do not. (This seems plausible to me for the Ad* licensed by sa, but it is not clear how it extends to the Ad licensed by Fin in Magahi, given that Fin does not have a special meaning related to speech acts.)

[16] There is one suffix that is used for both subject agreement and allocutive agreement in Magahi: the HH marker *-thi(n)*. This marks both HH subject agreement (seen in (23c)) and HH allocutive agreement (seen on the matrix verb in (27c)). However, *-thi(n)* is also used for third person honorific subjects. Therefore, it is some kind of default HH marker with relatively broad insertion conditions, not one that is unambiguously associated with second person. See Alok (2020, 2021) for full paradigms and discussion. Similarly plural *-ŋgæ* is used with third person as well as second person and allocutive in Tamil.

notable that agreement with ordinary second person subjects makes the same three-way honorificity distinctions among NH, H, and HH that allocutive marking does. This can be seen by comparing the allocutive marking in (3) with the subject agreement in (23).

(23) Magahi (Alok 2020: 45)

a. Tu/tohani dauR-l-eN.
you.SG/you.PL run-PFV-2.NH.S
"You (a peer/peers) ran."

b. Tu/tohani dauR-l-a.
you.SG/you.PL run-PFV-2.H.S
"You (a parent/parents) ran."

c. Apne dauR-la-thi(n).
you.HH.SG/PL run-PFV-2.HH.S
"You (a professor/professors) ran."

This three-way contrast is a characteristic of the second person in Maghai. In contrast, third person elements undergo impoverishment so that they only show a two-way distinction (nonhonorific vs honorific) and first person is always nonhonorific in Magahi; see Alok (2020) and Alok & Baker (2022) for examples, analysis, and discussion.

The strongest evidence that Ad is second person in Magahi comes from pronoun binding, in that a pronoun that is bound by Ad must be second person. To see this, consider again the situation in which there is a CP complement of 'tell', and the embedded verb has allocutive marking determined by the status of the goal of 'tell', as in (6). Suppose that in addition there is a pronoun in the embedded clause that refers to the goal of 'tell'. The striking fact is that such a pronoun must be second person, and cannot be third person, as shown in (24).

(24) Magahi (fieldwork, Deepak Alok)

Santee-aa Bantee-aa-ke kahk-ain ki Ram toraa/#okraa dekh-l-au.
Santee-FM Bantee-FM-DAT told.3.NH.S-HH.AL that Ram you.NH.ACC/#3SG.NH.ACC see-PFV.3.NH.S-NH.AL
"Santee told Bantee$_i$ that Ram saw you$_i$/#him$_i$." (said to a teacher)

This example qualifies as an instance of indexical shift of the second person pronoun, providing me with a bridge to the topic of indexical shift, the theme of the next chapter. For now, what is important is the relationship between allocutive marking and indexical shift in Magahi. When there is allocutive marking of the sort in (24), 'you' must be understood as shifted, and 'him' is not possible with the same antecedent 'Bantee'. In contrast, when the embedded verb does not have overt allocutive marking, using a third person pronoun to refer to the matrix third person goal is perfectly fine, as in (25).[17]

(25) Magahi (fieldwork, Deepak Alok)
Santee-aa Bantee-aa-ke kahk-ai ki Ram okraa dekh-l-ai.
Santee-FM Bantee-FM-DAT told.3.NH.S that Ram 3SG.NH.ACC see-PFV-3.NH.S
"Santee told Bantee$_i$ that Ram saw him$_{i,k}$."

The representation of (24) is (26).

(26) Ad*$_k$ Fin [Santee T told Bantee$_i$ [$_{FinP}$ Ad$_i$ Fin [$_{TP}$ Ram saw pron$_i$]]].
2HH NH NH 2NH 2NH (*3NH)
Agree Control Agree

Ad is present in the embedded clause and controlled by the matrix indirect object 'Bantee', as shown by the allocutive agreement surfacing on the embedded verb (different from the agreement with the root Ad). Given this, a pronoun inside the CP complement that is bound by (c-commanded by and coindexed with) 'Bantee' is also bound by Ad. Indeed, Ad is its closest binder.[18] Now we observe that the pronoun referring to 'Bantee' must be second person, and we

---

[17] There are other possibilities as well; see Chapter 4 for a more complete presentation of indexical shift and its interaction with allocutive marking.

[18] This is a tacit appeal here to Fox's (2000) Rule H: in a structure [… NP – pronoun1 – pronoun2 …], the second pronoun needs to depend on the first one, not directly on the ultimate antecedent, unless it makes a semantic difference. This rules out having the embedded object bound directly by the matrix object, such that it would be third person. See Chapters 4 and 5 for more on Rule H and ghostly DPs.

know that pronouns normally need to share the features of their binder. Therefore, I infer that the ghostly DP Ad in (26) is second person—not third person or underspecified for person. This is the central syntactic difference between allocutive constructions and upward C-agreement constructions in the African languages, I claim. (26) is the same as the structure of upward C-agreement with the object in Kipsigis, except for the phi-features of the DP in CP.

A clear way to see this difference between Ad and SoK/OoK in the African languages is to vary the person of the controller of the ghostly DP. It is notable that although embedded allocutive verb morphology can vary with the social status of the indirect object in Magahi, as in (6), it does not vary with the person of the indirect object. This can be seen in (27), where the same allocutive marking shows up on the embedded verb regardless of the person of the object. The affix *-au* on the embedded verb (distinct from the allocutive marking on the matrix verb) shows that the goal of 'tell' is in a nonhonorific relationship to the agent of 'tell', but that goal can be first, second, or third person.

(27) Magahi (fieldwork, Deepak Alok)

a. Santee-aa ***Bantee-aa-ke*** *kah-l-o ki adamiaa chal ge-l-****au****.*
Santee-FM Bantee-FM-DAT tell-PFV.3.NS.S-H.AL that worker walk go-PFV.3.NS.S-NH.AL
"Santee told Bantee that the worker left." (said to the speaker's grandfather)

b. Santee-aa ***ham-raa*** *kah-l-o ki adamiaa chal gel-****au****.*
Santee-FM me-DAT tell-PFV.3.NS.S-H.AL that worker walk go-PFV.3.NH.S-NH.AL
"Santee told me that the worker left." (to gr'father)

c. Profesor X ***apne-ke*** *kah-l-a<u>thin</u> ki adamiaa chal gel-****au****.*
Professor X you.HH-DAT tell-PFV-3.HH.S:HH.AL that worker walk go-PFV.3.NH.S-NH.AL
"Professor X told you that the worker left." (said to a teacher; NH *au*=professor X's relationship to the addressee, not the speaker's)

In contrast, C-agreement in the African languages does vary with the person of the controlling argument of the matrix verb. (28) gives a

paradigm from Lubukusu.

(28) Lubukusu (Diercks 2013: (20a), (31), (1b))

a. ***N****-a-bol-el-a Nelsoni* ***ndi*** *ba-keni ba-a-cha.*
1SG.S-PST-say-APPL-FV CL1.Nelson 1SG.C CL2-guests CL2.S-PST-go
"I told Nelson that the guests left."

b. ***Ewe*** *w-a-bol-el-a Nelsoni* ***o-li*** *ba-keni ba-rekukha.*
you 2SG.S-PST-say-APPL-FV CL1.Nelson 2SG-C CL2-guests CL2.S-left
"You told Nelson that the guests left."

c. ***Alfredi*** *ka-bol-el-a baba-ndu* ***a-li*** *ba-kha-khil-e.*
CL1.Alfred 3SG.S-say-APPL-FV CL2-people 3SG-C CL2.S-FUT-conquer
"Alfred told the people that they will win."

This is also true for upward C agreement with objects in Kipsigis: the agreement varies in person as well as in number (Diercks and Rao 2019: 382). So SoK and OoK originally lack phi-features of their own, and inherit them from their controllers, whereas Ad has intrinsic second person phi-features and keeps those features even when it is controlled.[19] Ad acquires honorificity features from its controller, but not its person feature(s).

Another consequence of this difference in person features is that C-agreement does not go along with indexical shift in the African languages the way that control of allocutive does in Magahi. For example, in Kipsigis a pronoun referring to the matrix goal argument that controls Agr-C is third person if the ultimate antecedent is itself third person. In (29), 'you' in the embedded clause cannot refer to Kibeet, the goal of the telling event, whereas 'he' in the embedded clause can. This is the opposite of what we see in (24) in Magahi.

---

[19] One might think that intrinsic person features on Ad would prevent it from being controlled by a DP with different features. However, technically the [+2] feature is not semantically interpretable on embedded Ad, so there is no deep incompatibility. See §4.4 for discussion.

(29) Kipsigis (Madaline Bossi, p.c.)

a. Kɔɔ-mwa-ʤi Kibeet a-le-nʤi ii-tiiny-e tɔɔndet kaaron.
PST.1SG-tell-APPL.3.O Kibeet 1SG-C-3.O 2SG.S-have-PROG visit tomorrow
"I told Kibeet$_i$ that you$_{*i,ad*}$ will have a visitor tomorrow."

b. Kɔɔ-mwa-ʤi Kibeet a-len-ʤi Ø-tiiny-e tɔɔndet kaaron.
PST.1SG.S-tell-APPL.3.O Kibeet 1SG-C-3.O 3SG.S-have-PROG visit tomorrow
"I told Kibeet$_i$ that he$_i$ will have a visitor tomorrow."

The upshot is that the DP goal of C-agreement in Magahi is parallel to SoK and especially to OoK in Kipsigis, but they are not identical; the two ghostly DPs differ in intrinsic person features. We see below that the fact that Ad has intrinsic phi-features whereas SoK does not also accounts for some other differences between allocutive agreement and upward C-agreement, including the fact that allocutive agreement is not subject to the T/Agree Condition and the fact that it is possible in clauses which are not in positions that allow for obligatory control.

3. Sp and speaker agreement

3.1. Sp as a first person analog of Ad

According to Speas & Tenny's (2003) neoperformative hypothesis, the addressee-denoting element Ad is paired with an analogous speaker-denoting element Sp. As the agent is the external argument of a verb like 'tell' and the goal is an internal argument, so Sp is an external argument in the C-space and Ad is an internal argument. I adopt this assumption, positing that Sp is present in the same syntactic contexts that Ad is. As Ad is intrinsically second person, Sp is intrinsically first person. As the matrix object controls OoK in Kipsigis and Ad in Magahi, so the matrix subject controls SoK in a variety of African languages and Sp in Magahi. As positing OoK in addition to SoK extends the theory of upward C-agreement with subjects to the rarer phenomenon of upward C-agreement with objects, so positing Sp in addition to Ad extends the theory of allocutive agreement to the rarer phenomenon of speaker agreement.

Like most languages, Magahi does not have speaker agreement. In this language, the key evidence that FinP licenses Sp as well as Ad will be

that indexical shift of first persons happens in the same contexts as second person indexical shift does in sentences like (24)/(26). Like second person indexical shift, first person indexical shift is also sensitive to whether allouctive agreement shows that Ad is controlled. This indexical shift is the topic of Chapter 4, but (30) provides an introductory example. Here in the same syntactic environment in which 'you' in the embedded clause refers to Bantee, the goal of the matrix verb, 'I' in the embedded clause refers to Santee, the subject of the matrix verb. If we accept that this usage of 'you' comes from the pronoun being bound by a second person operator (Ad) that is controlled by the matrix goal, parity of reasoning suggests that this usage of 'I' comes from the pronoun being bound by a first person operator (Sp) that is controlled by the matrix subject.

(30) Magahi (fieldwork, Deepak Alok)
Santee-aa Bantee-aa-ke kah-l-ai ki ham toraa dekh-l-i-au ha-l.
Santee-FM Bantee-FM-DAT tell-PFV.3.NH.S that I you.NH.ACC see-PFV-1SG-NH.AL be-PFV
"Santee told Bantee that I saw you." (if said to a teacher, then you=Bantee and I=Santee only).

This theory of indexical shift is developed in detail in Chapter 4.

The generalized structure that adds in Sp is given in (31), an enrichment of (26). As a matter of implementation, I assume that Fin can be decomposed into a sequence of two heads, Fin1 and Fin2, with Sp the specifier of the higher head Fin1 (as the agent is the specifier of the higher head v/Voice) and Ad the specifier of the lower head Fin2 (as the goal is the specifier of the lower head V or Appl in the thematic domain).[20] This is parallel to my distinction between Eval1,

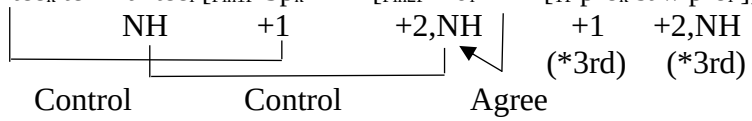
---

[20] Similarly I assume that there is a version of Sp, Sp*, that occurs only in root clauses and designates the speaker for the sentence. I assume for concreteness that Sp* is the specifier of sa1 and Ad* is that specifier of the lower head sa2, following Haegeman & Hill (2013), Miyagawa (2017), Zu (2018), and McFadden (2020) (although terminologies differ some). The more fully elaborated version of (21) that includes both Sps and the heads that license them is given in (i). Also, parallel to fn 15, I want to say that sa1 imposes [+1] on its specifier in virtue of its meaning, similar to the way the promissive head does in Korean (Zanuttini et al. 2012: 1240, 1239).

(i) $[_{sa1P}$ Sp^*_k sa1$[_{sa2P}$ Ad^*_i sa2 $[_{CP}$ C $[_{Fin1P}$ Sp_k Fin1 $[_{Fin2P}$ Ad_i Fin2 $[_{TP}$ you_i

the licenser of SoK, and Eval2, the licenser of OoK in Chapter 2. However, to help with the readability of my structures, I often include only one of these heads, calling it Fin (or just C).

(31) Santee$_k$ tell Bantee$_i$ [$_{Fin1P}$ Sp$_k$ Fin1 [$_{Fin2P}$ Ad$_i$ Fin2 [$_{TP}$ pro$_k$ saw pro$_i$]]]



Fin1 does not agree with Sp in Magahi, the way that Fin2 agrees with Ad. However, if Sp is part of the syntactic representation of certain clauses, just as Ad is, we might naturally expect crosslinguistic variation on this point: there could very well be languages which have "speaker agreement" rather than (or in addition to) allocutive-addressee agreement. A possible case of this is the Caucasian language Sanzhi Dargwa (Forker 2019), pointed out to me by Troy Messick. This language has complementizers derived from a converb form of 'say'.[21] It` is also rich in agreement, particularly agreement in gender and number. One case in point is that the 'say'-complementizer agrees with the matrix subject in gender, a form of upward complementizer agreement. *Ik'ul* 'that' shows masculine

---

saw me$_k$]]]]]]

The alternative to (i) would be to say that Sp and Ad are both specifiers of a single head Fin, Sp* and Ad* are both specifiers of a single head sa, and SoK and OoK are both specifiers of a single head Eval. The choice between these two implementations is largely one of theoretical design: whether one wants to maintain that a head can only have a single specifier versus whether one wants to avoid multiplying abstract functional heads. For the most part, the difference is not crucial to my analyses. I couch my theory in terms of the two-heads-with-one-specifier-each version for two reasons. First, this is the one that most previous users of the neo-performative theory have adopted. Second, it is somewhat easier to state the necessary agreement parameters in these terms. For example, Dargwa has agreement with Sp but not Ad, and Magahi has agreement with Ad and not Sp. In terms of (31), we can easily say that Fin1 but not Fin2 is an agreement probe in Dargwa and Fin2 but not Fin1 is an agreement probe in Magahi. It is less clear how to state this parametric variation if Sp and Ad are two specifiers of a single Fin head in both languages.

[21] Forker (2019) does not take a firm stand on whether the relevant form is truly a complementizer of verbal origin or a subordinate form of the verb 'say' (she glosses it as 'say'). The issues for evaluating this would be similar to those discussed for Ibibio and Kipsigis in Chapter 2.

agreement (Ø) with the matrix subject in (32a) but feminine agreement (*r*) with the corresponding subject in (32b).

(32) Sanzhi Dargwa (Forker 2019: (13b), (5a))

a. ***It-i-l*** *xar b-irʁ-ul ca-b [[ina-d du-l murhe daʕʔaʕn-ne d-arq'-ib=da=jal]* ***Ø****-ik'-ul].*
that-OBL-ERG ask N-ask.IPFV-ICVB be-N where-N.PL 1SG-ERG gold secret-ADVZ N.PL-do.PFV-PRET=1=Q M-C-ICVB
"He asks where I hid the gold."

b. ***Dam*** *han b-ič-ib [[a-b-elk'-un-ne]* ***r****-ik'-ul].*
1SG.DAT seem N-occur.PFV-PRET NEG-N-write.PFV-PRET-CVB F-C-ICVB
"I (fem.) thought that he did not write."

Messick observes that the Cs here agree with the matrix subject even when it bears an oblique case—ergative in (32a) or dative in (32b). This is different from verbs and other elements, which agree only with absolutive DPs in Sanzhi Dargwa. Thus, the matrix verbs in (32) show neuter agreement, agreeing with the neuter-gender predicate of the light verb construction, or perhaps as default agreement, rather than masculine or feminine agreement with the oblique subject. Messick points out that this is evidence that C is really agreeing directly with a DP in the CP periphery (which is presumably caseless, so active for Agree), where that DP is controlled by the matrix subject. This is a version of Diercks's indirect agree hypothesis. My question now is what precisely does C agree with: is it SoK, as in African languages, or Sp? Within my theory, this amounts to the question of whether the ghostly DP associated with C is first person or not. We cannot tell from the form of agreement on C, because like many other heads in Dargwa, C agrees only in gender and number, not person. However, we can tell by looking at the features of pronouns bound by the DP. Dargwa does allow indexical shift in the complement clause. 'I' can get a shifted reading in which it refers to the matrix subject, as in (33).[22]

(33) Sanzhi Dargwa (Forker 2019: (26))

Šaʕban-ni-j *han bi-irk-ul ca-b [[****dam*** *žawab b-alχ-ad]* ***Ø****-ik'-ul].*

[22] Dargwa has another "shifty" phenomenon in which the reflexive pronoun *ca* can trigger first person agreement on the verb in this sort of embedded context only

Shaban-OBL-DAT seem N-be.IPFV-ICVB be-N 1SG.DAT
answer N-know.IPFV-PRS.1 M-C-ICVB
"Shaban$_i$ thinks that he/I$_{i,k}$ know the answer."

Here 'I' is ultimately bound by the matrix subject 'Shaban', the same DP that C appears to agree with. Therefore, it is bound by the ghostly DP licensed by C. And it is first person. We can infer from this that the ghostly DP itself is first person—it is an instance of Sp. This then is a likely case of C (Fin1) agreeing with Sp rather than Ad.

Another language that shows agreement with Sp is the Tibeto-Burman language Jingpo, according to the analysis of Zu (2018). An example is given in (34a). Unlike Dargwa, speaker agreement is clearly first person in Jingpo, in that it uses the same vocabulary item that is otherwise triggered by a first person plural subject, seen in (34b).[23]

(34) Jingpo (Zu 2018: (53a,b), (75a))

*a. Jongma du hkum mas-ai/**saga**-ai.*
student arrive complete 3PL-DECL/1PL-DECL
"The students have all arrived." (1PL implies affection and solidarity between the students and the speaker.)

*b. (Anhte) masum lang hti **saga**-ai.*
we three times read 1PL-DECL
"We have read (it) three times."

I conclude that there are plausible cases of speaker agreement as well as better-known addressee agreement, as expected within the Speas & Tennyan framework. See also Rose (2015) for some South American

(Forker 2019: (21)). This is monstrous agreement in the sense of Messick (2023). See Chapter 6 for discussion of how this fits into my overall framework. (Dargwa might sometimes have 1LogOp rather than Sp in its CP periphery.)

[23] Note that the putative speaker agreement in (34a) must be first person plural, even when only a single person is speaking. Perhaps then the declarative head is really agreeing with both the subject and Sp in Jingpo, and the 3pl and 1sg features sum to 1pl. Because of questions about this, I hesitate to use Jingpo as a paradigm case of agreement with Sp.

languages that may have agreement with Sp.

3.2. The T/Agree Condition revisited

One important property of upward C-agreement that emerged from the discussion in Chapter 2 is that, in the Niger-Congo languages, C can only agree with SoK if the controller of SoK itself triggers agreement on a canonically-agreeing functional head, such as T. I called this the T/Agree Condition, deriving it from the distinction between Agree-Link and Agree-Copy. Now we can revisit this, given that we have evidence that speaker agreement and addressee agreement are not subject to this condition. This is clear in Dargwa: we saw in (32) and (33) that ergative and dative subjects do not themselves trigger agreement on elements in the matrix clause; agreement is only possible with absolutive nominals in this language. Despite this, the Sp that these subjects control can trigger agreement in gender and number on the C-head of the embedded clause. The most minimal comparison is between experiencer objects in Ibibio and the oblique experiencer argument of a verb like 'know' in Dargwa: the former cannot trigger agreement on C (Eval) whereas the latter can (on Fin1). This contrast sheds further light on the T/Agree Condition.

Indeed, it is even clearer that allocutive marking in Magahi does not obey the T/Agree Condition, even in embedded clauses where it is most similar to upward C-agreement. As a good Indo-European language, Magahi has subject agreement but no object agreement (except in the special case of dative subject constructions). This however does not prevent Fin2 from agreeing with the ghostly DP Ad that the object controls, as in (6) and many other examples. In this section, I reformulate my analysis of the T/Agree Condition from §2.5 to account for this.

It would be easy enough to simply stipulate this difference by saying that the relevant C-like heads in Dargwa and Magahi are primary agreers, such that they themselves trigger Agree-Copy as well as Agree-Link, whereas the C-like heads in the African languages are dependent agreers that trigger only Agree-Link. Then the Cs in the African languages need T to trigger the actual copying of phi-features onto all the heads linked to SoK, whereas the Cs in Dargwa and Magahi can do this themselves. I tentatively suggested that Kipsigis was different from the Niger-Congo languages in just this way. But Tamil allocutive agreement is also like Magahi in this respect: C agrees with Ad in an example like (8b) even though nothing agrees

with the (covert) goal argument of ‘say’, the controller of Ad.

Let us suppose, then, that enough examples of embedded and shifted allocutive agreement come to light to make it clear that this difference between speaker/addressee agreement and upward C-agreement is not accidental. I propose that this difference between agreement with Sp/Ad and agreement with SoK can be derived from the basic difference that Sp and Ad have intrinsic phi-features whereas Sok does not. To make this work, I now hypothesize that C is always a primary agreer, just as T is—in the Niger-Congo languages as well as in Magahi and Dargwa. Then we can envision derivations like (35) for a generic example like ‘Mary thinks that Sue came’, comparing C-agreement with Sp in Dargwa with C-agreement with SoK in Kinande, as a near minimal pair.

(35) a. [$_{CP}$ SoK / Sp C [Sue came]]
Ø [+1, F] (Ø / +1, F)
Agree-Link + Agree-Copy

b. [$_{VoiceP}$ Mary Voice [think [$_{CP}$ SoK/Sp C [Sue came]]]]
[3sg, F] [3sgF] [1sgF] (Ø / +1, F)
control

c. T [$_{VoiceP}$ Mary Voice [think [$_{CP}$ SoK/Sp C [Sue came]]]]
[3sg, F] [3sgF] [1sgF] (Ø / +1, F)
Agree-Link

d. T [$_{VoiceP}$ Mary Voice [think [$_{CP}$ SoK/Sp C [Sue came]]]]
[3sg] [3sg] [3sg]/[1sgF] [3sg]/[1sgF]
Agree-Copy

When a C head is first merged with SoK or Sp (or Ad) in the embedded clause, it undergoes Agree-Link and Agree-Copy immediately ((35a)). In Dargwa (or Magahi), C then gets first (or second) person features from Sp (or Ad) immediately, and that is essentially all there is to it.[24] However, in the Niger-Congo languages

[24] However, it is not so obvious how to handle gender agreement on C in Dargwa: does that feature originate on Sp, or is it inherited from the controller, as all

C cannot get phi-features from SoK at this early stage, because SoK does not have phi-features yet. It has no intrinsic phi-features, as we have seen, and control has not happened yet at this point in the bottom-up derivation. When the matrix VoiceP is built, as in (35b), control happens and SoK receives the phi-features of its controller. However, Agree-Copy does not automatically reapply, the initial opportunity for that having already passed. If T is later merged and enters into Agree with the controller of SoK, as in (35c), Agree-Link takes place, and Agree-Copy is triggered again. Recall from Chapter 2 that I formulated Agree-Copy as in (36) to allow for the phenomenon of dependent agreement, found not only with Cs in the Niger-Congo languages but with infinitives in Hindi and participles in Icelandic.

(36) Agree-Copy:
If head H points to DP and H is [+Agree-Copy], then phi(DP) is copied onto *all heads* linked to DP.

Therefore, the embedded C in a Niger-Congo language can get phi-features as a side effect of this second instance of Agree-Copy, triggered by the matrix T, as shown in (35d). In this way, it arrives at essentially the same endpoint that C in Dargwa and Magahi get to in one step. In contrast, if no T ever triggers Agree-Copy with the controller of SoK (say because the matrix VoiceP is embedded in a causative construction and T agrees with the causer, or because the agent is null or oblique in a passive construction), this does not happen. Then Agree-Copy is not triggered again by any head that points to SoK or a controller thereof and C never sees the fuller features put on SoK by obligatory control.

This gives the basic contrasts in the Niger-Congo languages that were discussed in §2.5 in a slightly revised way. In this version, the fact that the T/Agree Condition is relevant to C-agreement with SoK but not to C-agreement with Sp or Ad follows from the fact that Sp and Ad have intrinsic phi-features but SoK (and OoK) do not—the fundamental difference between these ghostly DPs according to the current view.

4. Conditions on the control of Ad

---

features are with SoK? I leave this open. Note also that I assume that a vacuous application of Agree-Copy that does not successfully transfer features from the goal (since the goal does not have features yet) does not cause the pointer created by Agree-Link to be removed.

4.1. Orientation

Now that we have had a good look at the nature of Ad and its place in the structure, we can consider in more detail how it can be controlled. My goal is to show that the same core conditions that regulate the arguments of the matrix verb controlling SoK and OoK, the targets of C agreement in African languages, also regulate the arguments of the matrix verb controlling Ad in Magahi. (Although Sp in Magahi was introduced above, systematic discussion of its control behavior is deferred to the next chapter, on indexical shift.) Two principles from Chapter 2 are relevant to this topic. First, the Generalized Obligatory Control Signature puts conditions both on where the clause containing a ghostly DP needs to be in order for the operator to be controlled and on what can be the controller. The GOCS is repeated in (37).

(37) The Generalized OC Signature: (GOCS)
If a clause with an intrinsically null DP (PRO, SoK, OoK, ***Ad***, ***Sp, …***) at its edge is generated within the XP headed by lexical head X, then the null DP is controlled by an argument of X. Which argument of the X is the controller is determined by the thematic roles of the controller and the controlee.

I also argued for a thematic role matching condition that governs which argument of a matrix verb controls which ghostly DP in the periphery of a clause that depends on the verb. (38) repeats this.

(38) The obligatory controller of X in a CP inside VP is the argument of the verb whose thematic role (best) matches the thematic role of X.

In this section, I argue that the same principles are work in Magahi, with Ad behaving very much like OoK. If the empirical case that I build is convincing, this will help to justify these as the UG basis for a generalized control theory. One might think that a single phenomenon (upward C-agreement) attested in one region of the world is too slender a basis to justify proposing new general principles. If, however, the same principles work for a not-obviously related phenomenon found in different areas of the world, then the plausibility of UG being at work is much greater. This will be a significant step towards a generalized control theory. (The third principle discussed alongside (37) and (38) in Chapter 2 was the T/Agree Condition, but we have just seen that this does not apply to

Sp/Ad constructions, for principled reasons.)

4.2. Thematic role matching and Ad

The signature property of C-agreement in Africa discussed in Chapter 2 was that the superordinate subject controls it and the superordinate object does not, despite the object being structurally closer to SoK and the agreeing C than the subject is. Kipsigis then entered the story, showing (tentatively) that the goal argument of the matrix verb can control a ghostly DP inside the CP after all, although not SoK but only a distinct item, dubbed OoK, which triggers agreement in a different position. The pattern is that agent arguments can control SoK but not OoK, and goal/theme arguments can control OoK but not SoK. This motivated the thematic-role matching condition in (38).

Now we apply this to Ad. Continuing to follow Speas & Tenny's (2003) intuition that Sp and Ad are arguments of C-type heads in a way that is parallel to agent and goal being arguments of a verb like 'tell', Ad gets an object-like thematic role from the C-space head that licenses it (sa and/or Fin). The expected consequence of this given (38) is that the goal of a verb like 'tell' can control Ad, resulting in shifted allocutive marking, but the agent of a verb like 'tell' cannot. The positive side of this generalization has been seen above in (6). (39a) is a slight variant of (6c). Here the HH status of the goal argument 'professor' (or, more accurately, their HH status relative to Santee) is reflected in the HH affix *-ain* on the embedded verb, different from the H allocutive marker *-o* on the matrix verb. The negative side of this generalization can be seen in (39b). Suppose that the subject of the matrix clause could control Ad. In this case, the subject is grandfather, in an H relationship to the speaker of the sentence (and to the indirect object *Santee*, referring to the speaker's cousin). Then we might expect the allocutive marker *-o* H to be possible on the embedded verb, in contrast to the NH marker *-au* on the matrix verb. But his is impossible.

(39) Magahi (fieldwork, Deepak Alok)

a. Santee-aa ***profesar saaheb-ke*** *kahl-o ki Ram Sita-ke dekh-l-****ain*** *ha-l.*

Santee-FM professor HH-DAT told.3.NH.S-H.AL that Ram

Sita-ACC see-PRF.3.NH.S-HH.AL be-PFV
"Santee told the professor that Ram saw Sita." (a peer of Santee speaking to his grandfather) (*…*dekh-l-au* see-PRF.3.NH.S-NH.AL)

*b. *Baabaa Santee-aa-ke kahk-**au** ki Ram Sita-ke dekh-l-**o** ha-l.*
grandfather Santee-FM-DAT tell.PFV.3NH.S-NH.AL that Ram Sita-ACC see-PFV.3.NH.S-H.AL be-PFV
("Grandfather told Santee that Ram saw Sita." said to a friend)

The inability of the agent/subject to control Ad can also be seen by the badness of *dekh-l-au* as the embedded verb in (39a). Here the embedded verb can have shifted HH marking reflecting the status of the goal 'teacher' relative to Santee, but not shifted NH marking *dekh-l-au* reflecting the status of the agent Santee (relative say to the speaker). (Unshifted allocutive marking here would be *-o*, marking the H status of the addressee grandfather relative to the speaker; this is also grammatical.) So there is parallelism between C-agreement and shifted allocutive marking with respect to the thematic role matching condition in (38).

We can confirm that thematic role matching is important, not just grammatical functions, by comparing the matrix verb 'tell' with the matrix verb 'hear'. The verbs denote similar events, but with 'hear' the experiencer/goal argument is realized in the subject position, whereas the agent/source argument is realized (if at all) as an oblique phrase inside the greater verb phrase. (40) shows that the hearer subject can control Ad in a way that is like how the tellee object does in (39a), and different from the agent subject in (39b). The allocutive marking on the embedded verb in (40) can be *-o*, distinct from the HH marking *-ain* on the matrix verb. This is a shifted allocutive H, showing the status of the experiencer subject 'grandfather' relative to the source Bantee.

(40) Magahi (fieldwork, Deepak Alok)
***Baabaa** Bantee-aa-se sun-la-thin ki Ram Siita-ke bajaar-me dekh-l-**o**.*
Grandfather Bantee-FM-INS hear-PFV.3.HS.S-HH.AL that Ram Sita-ACC market-in see-PFV.3.NH.S-H.AL
"Grandfather heard from Bantee that Ram saw Sita in the market."

This again illustrates the thematic role matching in (38). It shows that

it is more accurate to say that the agent/source controls Sp (see Chapter 4) and the experiencer/goal controls Ad than to say that the subject controls Sp and the oblique nominal controls Ad.

These is a sense in which the thematic roles referred to in (38) must be relatively fine-grained ones. The kind of thematic analysis that determines where the arguments of a clause are generated (e.g., by principles like the Uniformity of Theta Assignment Hypothesis) are quite coarse-grained: they often do not distinguish (say) agent from causer and source or experiencer from goal. But control theory does distinguish these second-order roles in some contexts. We can see fresh evidence of this by comparing 'hear' with 'ask' in Magahi. The verbs are superficially comparable: both have nominative subjects and instrumental objects, as seen by comparing (40) with (41).

(41) Magahi (fieldwork, Deepak Alok)

a. Raam ***profesar saaheb-se*** *puchhk-au ki kaa Siita ait-****ain****.*
Ram professor HH-INS ask.PFV.3.NH.S-NH.AL that what Sita come.FUT.3.NH.S-HH.AL
"Ram asked the professor whether Sita will come." (said to a peer)

b. Raam ***jaun-waa-se*** *puchhk-ain ki kaa Siita ait-****au****.*
Ram John-FM-INS ask.PFV.3.NH.S-HH.AL that what Sita come.FUT.3.NH.S-NH.AL
"Ram asked John whether Sita will come." (said to a teacher)

Despite the similarity in case-marking between (40) and (41), what controls Ad is different in the two examples. In (41), it is the social status of the instrumental DP (relative to the asker) that determines shifted allocutive marking, not the social status of the nominative DP (relative to the source) as in (40). This can be attributed to a difference in fine-grained thematic roles. The internal argument of 'ask' the intended source of the new information that the asker seeks. This presumably accounts for why it bears the same semantic case (instrumental) as a canonical source phrase in Magahi. But this argument counts as a goal as well as a source, since the asking event itself is directed to the referent of this argument. In (41b), the question goes to John, even though its answer is expected to come from John. The internal argument of 'ask' is thus a composite goal-source, whereas the internal argument of 'hear' is a pure source. Because the internal argument of 'ask' has goal entailments as well as source

entailments, it matches the Ad argument of C thematically and can control it. I conclude that multilayered thematic roles can shape how OC happens in the allocutive construction.

We can also compare experiencer arguments with goal arguments in Magahi. These are similar thematic roles, as shown by the fact that both types of arguments are marked by the dative postposition *-ke*. (In the theory of Baker (2024), this indicates that both are generated as the higher of two nominals inside VP.) However, they have different behavior when it comes to controlling Ad. Two verbs that take dative experiencer arguments in Magahi are 'seem' and 'remember'.

(42) Magahi (fieldwork, Deepak Alok)

a. Santee-aa-ke laga h-ai ki Ram tej h-ai.
Santee-FM-DAT seem be-3.NH.S that Ram smart be-3.NH.S
"It seems to Santee that Ram is smart."

b. Ram-ke yaad ha-l-ai ki Santee almira-me paisa chhupai-l-ai ha-l.
Ram-DAT memory be-PFV-3.NH.S that Santee drawer-in money hid-PFV-3.NH.S be-PFV
"Ram remembered that Santee hid the money in the drawer."

The question, then, is whether the dative argument of such constructions can control shifted allocutive marking on the embedded verb, the way that the dative argument of a verb like 'tell' can. In fact, this is not possible, as shown by the ungrammaticality of (43a,b).[25] Here the matrix verb is marked with NH allocutive morphology *-au*, making it clear that the whole sentence is addressed to a peer of the speaker. The dative subject is 'grandfather', the referent of which is in an honorific relationship to the speaker. If 'grandfather' could control Ad in the embedded clause, then we would expect the H allocutive marker *-o* to be possible on the embedded verb, in agreement with Ad. But this is not possible.

---

[25] We see in Chapter 4 that these dative subjects can control Sp, resulting in first person indexical shift. This confirms that experiencer arguments match the thematic role of Sp rather than that of Ad. Chapter 2 presented the parallel fact, that an experiencer object in Kipsigis controls SoK rather than OoK, triggering prefixal agreement rather than suffixal agreement on the C-like head *le*.

(43) Magahi (fieldwork, Deepak Alok)

*a. *Baabaa-ke laga h-**au** ki Ram Siita-ke beijati karl-**o**.*
grandfather-DAT seem be.3.NH.S-NH.AL that Ram Sita-GEN insult do-PFV.3.NH.S-H.AL
("It seems to Grandfather that Ram insulted Sita." said to a friend)

*b. *Baabaa-ke yaad ha-l-**au** ki Ram almira-me paisa chhupai-l-**o** ha-l.*
gr'father-DAT memory be-PFV.3.NH.S-NH.AL that Ram drawer-in money hid-PFV.3.NH.S-H.AL be-PFV
("Grandfather remembered that Ram hid money in the drawer." said to a friend).

The 'remember' example is interesting because this predicate can undergo a transitivity alternation: 'remind' is formed by using the same nominal predicate *yaad* 'memory' with the light verb 'give' rather than 'be'. In the transitive version, the dative argument can control shifted allocutive marking in the embedded clause:

(44) Magahi (fieldwork, Deepak Alok)

*Santee-aa baabaa-ke yaad dial-k-**au** ki Ram almira-me paisa chhupai-l-**o** ha-l.*
Santee-FM gr'father-DAT memory give-PFV.3.NH.S-NH.AL that Ram drawer-in money hid-PFV.3.NH.S-H.AL be-PFV
"Santee reminded Grandfather that Ram hid money in the drawer." (to a friend).

I take this as showing that thematic role classification is not only relatively fine-grained, but also somewhat context-dependent: the thematic role of one argument of a verb can depend on what other arguments the verb selects. This suggests that we should state the matching condition on thematic roles using something like macroroles (actor vs undergoer) in the sense of Foley & Van Valin (1984) or proto-roles (proto-agent vs proto-patient) in the sense of Dowty (1991). These works share the idea that an argument whose thematic status is somehow intermediate between that of a prototypical agent and that of a prototypical patient can count as the agent if it is the most agent-like participant in the event denoted by the verb whereas it can count as the patient if there is another participant in the event that is more agent-like (and there is no other participant that is more

patient-like).[26] For current purposes, I do not need the full apparatus of these theories; the special cases in (45) are sufficient for the task at hand. (I use the term "initiator" here following Ramchand (2008), rather than Foley and Van Valin's "actor", which I find too easy to confuse with agent.)

(45) Macroroles of initiator and undergoer:

a. The agent of an event is always its initiator.

b. The patient/theme of an event can be its undergoer.

c. The goal of an event can be its initiator if there is no agent.

d. The goal of an event can be its undergoer if there is an agent.

e. Ad is the undergoer of an event denoted by C (also OoK).

f. SoK is the initiator of an event denoted by C (also Sp).

g. The macroroles of the controller and the controllee must match.

(45g) makes it explicit that matching macroroles is enough to license obligatory control; this was already implicit in Chapter 2, where we saw that a range of finer-grained thematic roles can control SoK, including agent, source, experiencer, and causer. (45a) ensures that canonical agents are initiators and canonical patients are undergoers; this preserves the result that control with a matrix verb like 'tell' or 'ask' works as intended. (45c) and (45d) are the new part motivated by the data we are now considering. (45c) allows the goal/experiencer of verbs like 'seem' or 'remember' to control SoK (and Sp), but it prevents them controlling Ad (or OoK) given that the same argument

---

[26] Although similar in spirit, the two theories differ in their conception. Foley & Van Valin first classify an argument as agent, theme, patient, or whatever using predicate decomposition, and then pick out which is the actor and which is undergoer according to a hierarchy. In contrast, Dowty identifies a list of properties that contribute to being a proto-agent or a proto-theme, and the proto-agent is the argument the largest number of those properties. A difference between these two implementations is that "agent" and "theme" have a status along with "actor"/"proto-agent" and "undergoer"/"proto-patient" for Foley and Van Valin, but not for Dowty. My formulation is closer to Foley & Van Valin's than Dowty's, but this is primarily for expository purposes.

cannot be both an initiator and an undergoer. In contrast, (45d) allows the goal/experiencer of a verb like 'remind' to control Ad, but not SoK or Sp, because the agent will always be the initiator when there is one. This gives us the context sensitivity that we need for the 'remember'/'remind' alternation. (45) also leaves room for a bit of flexibility with a verb like 'hear', which has a goal/experiencer argument and sometimes a source argument. Here I assume that either the goal or the source can be picked as the initiator, since they are roughly equally good proto-agents. When the agentivity of the source is highlighted by picking it as the initiator, then the experiencer/goal can be taken as the undergoer by (45d). This accounts for the possibility of it controlling Ad, as in (40). However, when there is no source, or it is backgrounded, then the experiencer can be taken as the initiator by (45c). This allows it to control SoK, as we saw in some examples in Chapter 2. I consider a goal that is also an initiator to be an experiencer, and will often call it that, whereas a goal that is an undergoer I call simply a goal. (45) then fleshes out what is meant by "(best) match" in (38). In this execution, the roles of initiator and undergoer match perfectly, but there is some complexity and context sensitivity as to what is an initiator and what is an undergoer, as is well-known. This is relevant to how I derive thematic role matching from more fundamental principles in Chapter 8.

The dative subject constructions in (42)-(43) can also be compared to the passive of a ditransitive verb like 'tell' in Magahi. Dative case on the goal argument is not suppressed in passive sentences in this language, similar to Icelandic. As a result, a passive example like (46) looks very much like the dative subject constructions in (43), with the addition that it is possible to express the agent as a kind of *by*-phrase.

(46) Magahi (fieldwork, Deepak Alok)
*Baabaa-ke (Bittuu-aa diya) kahal ge-l-**au** ki Ram ait-**o**.*
grandfather-DAT Bittuu-FM by told go-PFV.3.NH.S-NH.AL that Ram come.3.NH.S-H.AL
"Grandfather was told (by Bittuu) that Ram will come." (said to a friend)

Although (46) looks like (42)-(43), it behaves more like (44) in that the dative argument 'grandfather' can control Ad in the complement clause, resulting in the shifted allocutive marker *-o* (H) on the embedded verb in that clause. The agent is present in (46), overtly or covertly. This causes the tellee to be categorized as an undergoer

rather than as an initiator, by (45d)). This in turn makes it possible for the tellee to control Ad, resulting in shifted allocutive marking.[27]

4.3. The locality of the control of Ad

Next I turn to the implications of the GOCS in (37) for the control of Ad, and hence for the possibility of shifted allocutive marking in Magahi. The GOCS puts two kinds of restrictions on a relationship of obligatory control. First, it constrains the controller of a null DP to be an argument of the verb (or other lexical head) that CP modifies or is the complement of. Second, it requires the clause containing the null DP to appear in the phrase headed by the verb that the controller is an argument of. I consider the first sort of restriction in this subsection and turn to the second in the next subsection.

Recall that upward C-agreement in African languages displays a kind of clause-level locality: C can only agree with the subject or the object of the immediately superordinate verb. This is shown schematically in (47): the second C head 'that' can agree with Z or W, the arguments of the lower verb 'tell', but not with X or Y, the arguments of the higher verb 'ask'. This is a consequence of the GOCS.

(47) X asked Y [whether [Z told W [Z/*X-that-(W/*Y) [so and so happened]]]].

One has to be a bit careful about how one tries to replicate this effect in Magahi. The complication is that control of Ad is in some sense optional in this language: it happens in examples like (6) but not in examples like (5). One example of each kind is repeated in (48).

---

[27] From a crosslinguistic perspective, the agent of a passive construction sometimes behaves more like a source phrase (with 'hear') than like a full/true agent. As such, it can be the initiator, but it need not be, opening the door for the goal argument to be the initiator instead. Thus, the goal subject of a passive can sometimes control SoK in Lubukusu, and it can control LogOp in Japanese. This apparently does not happen in Magahi however, perhaps because it is not clear that the goal argument can actually occupy the subject position.

(48) Magahi (fieldwork, Deepak Alok)

*a. Santee-aa **Bantee-aa-ke** kahk-ain ki Ram-ke Sita-se baat kar-e-ke chah-**au**.*
Santee-FM Bantee- FM-DAT told.3.NH.S-HH.AL that Ram-DAT Sita-INS talk do-INF-DAT should-NH.AL
"Santee told Bantee that Ram should talk to Sita." (said to a teacher)

*b. Santee-aa sochk-**au** ki Bantee-aa bhag ge-l-**au**.*
Santee-FM thought.3.NH.S-NH.AL that Bantee-FM run go-PRF.3.NH.S-NH.AL
"Santee thought that Bantee went to run." (said to a peer)

Descriptively, in (48a) the allocutive marking on the embedded verb is determined by the social status of the matrix goal 'Bantee' (NH), whereas in (48b) the allocutive marking on the embedded verb repeats the allocutive marking on the matrix verb (NH). (48a) is the case where Ad in the embedded clause has the matrix goal as its obligatory controller. In contrast, Ad in the embedded clause seems to be bound by Ad in the matrix clause, getting its honorificity values from that. (I return to how this is possible in §3.4.5.)

Despite this complicating factor, we can see that shifted allocutive marking does display the clause-level locality expected of OC by careful consideration of an example like (49). Here the lowest verb 'go' bears the H marker *-o*. This does not reflect the addressee of the sentence as a whole; that is a peer of the speaker, as shown by the presence of the NH allocutive marker *-au* on 'tell', the predicate of the root clause. Rather, it must be controlled by 'grandfather', the goal argument of 'tell' in the root clause, since this is the only NP in the sentence that refers to someone of higher social status (note that 'Santee' and 'Bantee' are first names that bear the familiarity marker *-aa*). However, the lowest verb can only bear *-o* triggered by the matrix goal in this way if the verb in the intermediate clause, 'be', also bears the H marker *-o* (or has no allocutive marker). It is bad in this case for the middle verb to bear the NH marker *-au* in agreement with allocutive marking on the verb in the root clause (contrast with (48b).

(49) Magahi (fieldwork, Deepak Alok)

*Santee-aa **baabaa-ke** kahk-au ki Bantee-aa socha h-**o/*au** ki Ram parichha paas ho ge-l-**o**.*
Santee-FM grandfather-DAT told.3.NH.S-NH.AL that Bantee-FM think be-H.AL/*NH.AL that Ram exam pass

become go-PFV.3.NH.S-NH.AL
"Santee told grandfather that Bantee thinks that Ram passed the test."

This subtle difference is evidence for the GOCS, which implies that the Ad of the lowest CP cannot be controlled directly by the root clause indirect object 'grandfather', because 'grandfather' is not an argument of 'think', the verb that the clause containing 'pass go' is the complement of. Since 'think' does not have a goal or theme argument, there is no possible obligatory controller for Ad in the lowest clause. Its only option, then, is for it to be bound by the Ad of the middle CP, which can in turn be controlled by the matrix goal (or not). In short, (50a) is not a possible representation in Magahi, with control at a distance. In contrast, (50b) is a possible representation, with local control plus local binding of Ad3 by Ad2.

(50) a. *$Ad1_k$ C S_i told Gr'$father_n$ [$Ad2_k$ that B think [$Ad3_n$ [R pass]]]
*control (+Ad1 binds Ad2)

b. √$Ad1_k$ C S_i told Gr'$father_n$ [$Ad2_n$ that B think [$Ad3_n$ [R. pass]]]
control +Ad2 binds Ad3

This supports the claim that only an argument of the verb that selects CP can control the Ad in the periphery of that CP.[28]

4.4. Allocutive marking in noncomplements

For the African languages, I considered the possibility of C-agreement in CPs that are not complements of verbs but are in some other syntactic environment: sentential subjects, CP adjuncts, relative clauses, and noun complements. We can also investigate the

[28] Another consequence of the GOCS for allocutive marking in Magahi should be that the possessor of (say) the goal argument cannot determine allocutive marking on the embedded clause in an example like "Santee texted(NH.AL) to **Grandfather**'s phone that Bantee is coming(**H.AL**) soon." It is true that second person indexicals cannot shift to refer to 'grandfather' in this environment (see Chapter 4), but I did not test allocutive marking, in part because it is not easy to construct natural-sounding examples.

possibility of controlling Ad in clauses in these positions. Like upward C-agreement, the only one of these constructions that allows shifted allocutive agreement is the CP complements of nouns. However, the verb in any of these clause types can contain an unshifted allocutive agreement, which expresses the relationship of the speaker to their addressee—the same marking that appears (optionally) on the verb of the root clause. The upshot of this is that the conditions on where a clause with Ad has to be in order to undergo obligatory control are the same as for SoK/OoK, but unlike SoK, Ad is also possible in environments in which it does not undergo obligatory control. In uncontrolled environments, Ad gets its value not from an argument of a higher verb, but rather from the next highest Ad, often the Ad of the root clause (but not necessarily, as in (49)/(50b)).

One canonical environment of NOC rather than OC according to Landau (2001, 2013) is CPs in subject position. However, Magahi is like several of the African languages in that bare CPs headed by the complementizer *ki* cannot appear in subject position, perhaps because this complementizer is more verbal than nominal in category. So this case does not arise in Magahi.[29]

Another canonical environment for NOC rather than OC according to the GOCS is adjunct clauses, especially higher adjunct clauses which are generated in positions outside VP. Alok (2020: 97) says that allocutive marking is possible in all kinds of adjunct clauses in Magahi. (51) gives two examples of this.

(51) Magahi (Alok 2020: 97)

*a. Santee-aa ai-l-**au** jab Bantee-aa chal ge-l-**au**.*
Santee-FM come-PFV.3.NH.S-NH.AL when Bantee-FM walk go-PFV.3.NH.S-NH.AL

---

[29] A CP is possible inside an NP headed by a noun like ‘news’ (used with a demonstrative) in Magahi. It is possible that this carrier noun has little effect on the control dynamics of the structure and is present only to give the CP the nominal features that it needs to sit in Spec TP and be the target of T-agreement. However, I do not know of any predicates in Magahi that permit a CP-(like) subject and two additional arguments that could control Sp and Ad, giving allocutive shift a chance of happening in the CP subject. See Chapter 4 for discussion of how this structure behaves with respect to indexical shift.

"Santee came when Bantee left."

*b. Santee-aa ghare ruk-l-**au** taaki Bantee-aa bimaar na paR-**au**.*
Santee-FM home stay-PFV.3.NH.S-NH.AD so.that Bantee-FM sick not fall.PFV.3.NH.S-NH.AD
"Santee stayed home so that Bantee would not get sick."

McFadden (2020: (20)) gives a similar example of allocutive marking inside a temporal adjunct clause in Tamil. Note that in these examples the allocutive marking on the embedded clause is the same as that on the root verb, showing the social relationship of the addressee of the sentence as a whole to its speaker. Adjunct clauses do not allow shifted allocutive marking, where allocutive marking on the embedded verb is different from that on the matrix verb, showing the social relationship of one of the matrix arguments with respect to another. In theoretical terms, Ad is possible in adjunct clauses, bound by a higher Ad, but it cannot be controlled by an argument of the matrix verb. This restriction can be seen most clearly in 'so that' clauses, where conditions are maximally favorable for allowing the control of Ad. As discussed in the next chapter, first-person indexical shift is possible in such clauses, showing that Sp of the adjunct clause can be controlled by the subject of the matrix clause, as in (52). This is parallel to the fact that C-agreement is possible in this type of adjunct clause in Lubukusu and Ibibio.

(52) Magahi (Alok 2020)
Bantee-aa ghare ruk-l-ai taaki ham bimmar na ho jaa-i.
Bantee-FM home stay-PFV.3.NH.S so.that I sick NEG become go-1SG.S
"Bantee$_i$ stayed home so that he/I$_{i,sp}$ would not become sick."

The example in (53) shows again that these 'so-that' clauses can contain allocutive marking, where that marking matches the marking on the matrix verb. (This is less obvious in (53) than in (51b), because in (53) the NH allocutive marking in the root clause forms a portmanteau with the 3.H.S subject marking, resulting in a different allomorph of NH.AL from the one on the verb in the 'so that' clause.)

(53) Magahi (fieldwork, Deepak Alok)
*Baabaa netaa-jore batiai-l-**thu** taaki hamraa kuchh phaidaa hob-**au**.*

grandfather leader-with talk-PFV-3.H.S:NH.AL so.that I.DAT something benefit be.FUT-NH.AL
"Grandfather spoke with/to the leader so that I will get some benefits." (said to a peer)

A matrix verb like 'talk' can also have a second argument, a kind of goal phrase, as in (53). So here we have an adjunct clause that can be controlled into and that has a controllable Ad, as well as a matrix verb that has a potential controller for Ad.[30] However, even under these maximally favorable circumstances, shifted allocutive marking in the embedded clause is not possible, as shown by the ungrammaticality of (54). Here *-thu* on the verb 'talk' shows that the sentence as a whole is addressed to an NH or H person. The HH marking *-ain* on the embedded verb 'be' does not match this matrix addressee, but would reflect the status of the community leader relative to the grandfather (or the speaker). This would be a shifted use, the result of Ad in the 'so that' clause being controlled by 'leader' in the root clause. It is, however, ungrammatical (regardless of whether 'I' shifts or not).

(54) Magahi (fieldwork, Deepak Alok)
Baabaa netaa-jore batiai-l-thu** taaki hamraa kuchh phaidaa hob-**ain**.*
grandfather leader-with talk-PFV-3.H.S:NH.AL so.that I.DAT something benefit be.FUT-HH.AL
("Grandfather$_i$ spoke with/to the leader so that he/I$_{i,sp*}$ will get some benefits.")

Therefore, higher adjuncts, like a 'when'-clause, do not allow the control of either operator, Sp or Ad, much as they do not allow upward C-agreement in Ibibio. But even lower adjuncts, like 'so that' clauses, which do allow the matrix subject to control Sp and SoK, do not allow the internal argument to control Ad. I return to this in Chapters 4 and 8, ultimately claiming that the 'so that' clause is merged with VoiceP rather than VP. As such, the GOCS implies that ghostly DPs in its periphery have to be controlled by arguments of

---

[30] The oblique instrumental argument of *batiai* 'speak' can control Ad when the embedded clause can be parsed as a complement rather than an adjunct, in an example like 'Grandfather spoke with/to Bantee$_i$ (saying) that you$_i$ should not be sad.' This shows that it is not the case of the goal that rules out control in (54).

Voice, not by arguments of V per se. The agent of the root clause is such an argument, but the goal phrase is not.

Next consider relative clauses adjoined to the extended projection of a noun phrase. This is an environment that does not allow OC in English or upward C-agreement in the Niger-Congo languages. The expectation is that in Magahi relative clauses will not allow allocutive marking that is shifted relative to that of the clause of which the NP modified by the relative clause is an argument. This is correct. (55) shows that a relative clause in Magahi can have unshifted allocutive agreement, expressing the relationship of the speaker to the addressee (although allocutive marking happens to be absent in the root clause in (55)).

(55) Magahi (Alok 2020: 98)
*Laikwaa [je uhan khaRaa h-**au**] hamar bhaai h-ai.*
boy REL there stand be.3.NH.S-NH.AL my.GEN brother be-3.NH.S
"The boy who is standing there is my brother." (said to a peer)

This shows that the CP in the relative clause can contain an Ad coordinate; it is not some kind of truncated clause that has no room for such elements. However, (56) shows that Ad in the relative clause cannot be controlled by the arguments of the verb 'tell' in the root clause. Here the goal of 'tell' is 'grandfather', the referent of which is in an honorific relationship to both Santee and the speaker of the whole sentence. If this NP could control Ad, then we would expect that the H marker *-o* to be possible as allocutive marking on the verb in the relative clause, but it is not. (This judgment also holds if the relative clause is extraposed to sentence final position, after 'tell'.)

(56) Magahi (fieldwork, Deepak Alok)
*Santee-aa [ii khabar [je Ram okraa kahk-**au/*o**]]*
***baabaa-ke** kah-k-au.*
Santee-FM this news REL Ram him.ACC tell-NH.AL/*H.AL grandfather-DAT tell-PFV.3.NH.S-NH.AL
"Santee$_i$ told grandfather the message that Ram told him$_i$."

Again, we see that it is possible to have an Ad in a clause even when it cannot be controlled by some other NP.

The fourth and final case to consider is CPs that function as the complement of a noun like 'rumor' or 'news'. In Chapter 2, we saw

that in this sort of structure, the head C can agree out of the NP in the African languages. If the same principles of control are at work in both domains, the expectation for Magahi is that shifted allocutive will be possible in this structure too, as long as the verb selecting the complex NP can select a goal phrase that is a possible controller. Examples relevant to this are (57). In (57a), what is semantically the complement of the object 'rumor' is extraposed to the right edge of the sentence, as is typical. In (57b) it remains inside the object, as is also possible. In either case, control of Ad is possible, resulting in a shifted allocutive marking that is different from that on the root verb.

(57) Magahi (fieldwork, Deepak Alok)

*a. Santee-aa **baabaa-ke** ii khabar kahk-au ki Ram ait-**o.***
Santee-FM grandfather-DAT this news told.3.NH.S-NH.AL that Ram come.3.NH.S-H.AL
"Santee told grandfather the news that Ram will come." (said to a friend)

*b. Santee-aa **baabaa-ke** ii khabar ki Ram ait-**o** kahk-au.*
Santee-FM grandfather-DAT this news that Ram come.3.NH.S-H.AL told.3.NH.S-NH.AL
"Santee told grandfather the news that Ram will come." (said to a friend)

In Chapter 2, my analysis was that nouns like 'news' can have arguments of their own, sometimes expressed overtly by possessors or PPs, but often covert (cf. *My news to you is that you passed the class*). The covert goal argument of the noun can control the Ad in the CP associated with the noun. This covert argument of the noun can in turn be controlled by the goal argument of 'tell'. This results in H marking on the embedded verbs in (57), reflecting the status of the grandfather.

The main goal of this section has been to point out substantive similarities between the control of Ad by an argument of the superordinate verb in Magahi, which results in shifted allocutive marking, and the control of SoK and OoK by the arguments of the superordinate verb in Bantu languages and Kipsigis, which results in upward C-agreement. Both constructions show clear signs of being subject to the GOCS and to a condition on thematic role matching. The positive similarities are seen in the complements of 'tell' type verbs, 'hear' type verbs, passives of 'tell', clausal locality effects, and noun complement clauses. Distinctive Magahi constructions that

follow the same general principles include dative subject constructions and triadic verbs with oblique objects like ‘ask’. The other side of the coin is that upward C-agreement and shifted allocutive marking are both ruled out in NOC contexts, including high adjunct clauses, relative clauses, and matrix clauses (also in principle CP subjects, but Magahi does not have these). These similarities make the two constructions ripe for a unified analysis in terms of a generalized control theory based on (37) and (38).

There is an important difference that has come to light as well. If Ad in Magahi is not controlled, the structure is still acceptable; Ad then denotes the addressee in the speech context (root clauses) or the same person as the next highest Ad (in embedded contexts). This contrasts with SoK, which is ruled out when OC does not apply. This difference can be derived, I claim, from the fundamental difference between Ad and SoK/OoK that I identified in §4.2.2: Ad has intrinsic second person features, whereas SoK has no intrinsic features at all.

First, we need a general principle to rule out uncontrolled SoKs, which have no features to start with and never receive any in the course of the syntactic derivation. It is plausible to attribute this to the Principle of Full Interpretation. A featureless DP like SoK cannot be interpreted at the LF interface, leading to a violation ((58a)). However, since OC adds the features of the controller to the controlled DP, it provides a way of satisfying this constraint ((58b)).

(58) a. A DP with no interpretable features is ruled out at LF.
b. OC assigns interpretable phi-features of the controller to the controlled DP.

Perhaps there are other ways of acquiring features prior to LF as well. But crucially ordinary pronominal binding is not one of them, in that it does not happen until LF and is not restricted by the PIC; I assume that a pronoun must already have some phi-features in order to participate in that. Thus, OC is de facto obligatory for certain kinds of elements, within the range of operations being considered here.

In contrast, the [+2] feature of Ad is sufficient to guide it to an adequate interpretation. For unembedded Ad in root clauses (Ad*), it is stipulated to refer to the addressee of the sentence (see also §4.4 for more discussion). For embedded Ad, the [+2] feature guides it toward finding an antecedent that matches its [+2] feature—such as another Ad. This antecedent can be outside the phase that contains Ad, and even outside a syntactic island, as in the relative clause example in

(55) and the adjunct clause examples in (51). This accounts for why SoK is not possible in relative clauses, high adjunct clauses, and sentential subjects, whereas Ad is possible in all these environments, as well as in ordinary root clauses like (1) and (3).[31]

4.5. The apparent optionality of the control of Ad

I round out this discussion by considering an issue with my hypothesis that shifted allocutive agreement happens in complement clauses when Ad undergoes obligatory control. The issue is that we have seen many times that Ad in Magahi can apparently resist OC, taking a higher Ad as its antecedent rather than an argument of the superordinate verb. This is possible even when Ad is inside a complement clause. As a result, shifted allocutive marking appears to be optional. We see this in single-embedded structures in the fact that embedded allocutive can either resume the allocutive marking on the matrix verb, expressing the addressee's social standing relative to the speaker (see (5)), or it can be shifted to show the matrix goal's rank

---

[31] The expectation coming out this account is that SoK is not possible in root clauses, ruling out agreeing Cs in such clauses. The literature generally takes it for granted that this is true, but it is not so easy to show it given the fact that agreeing Cs are homophonous with verbs of saying in many of the relevant languages, and verbs can undergo ellipsis in the context of an agreeing C in the Lunda/Luvale/Luchazi cluster of languages (Kawasha 2007: 186-187). In fact, Spadine (2020) argues that agreeing Cs are possible in matrix clauses in Tigrinya, but crucially they can take overt nominals in Spec CP, as in (i).

(i) Tigrinya (Spadine 2020: 16 (6))
[**Kidane** [Almaz mäṣḥaf ʔanbib-a] ʔil-**u**].
Kidane.M Almax.F book read-3SG.F C-3SG.M
"Kidane [says/thinks…] that Hiwet read a book."

There is no violation of (58a) here: the DP in Spec CP has interpretable features, so Full Interpretation does not require it to be controlled. The structure is then usable as a root sentence. Spadine argues that what many have taken to be sentences with defective verbs meaning 'say' or with null verbs meaning 'say' taking a CP complement are really unembedded CPs, not only in Tigrinya, but also in languages like Ewe and Malayalam. This might well be true for the Luvale cluster too, for *-te* in Ibibio, and beyond. Although Spadine's analysis is compatible my account, it does bring forward the unanswered question of why C-heads allow full overt DPs as their specifiers in some cases but not others.

relative to the matrix subject (see (6)). A minimal pair is given in (59); (59a) has the unshifted embedded allocutive marking that results from the binding of Ad by the root clause Ad, whereas (59b) has the shifted embedded allocutive marking that results from the control of Ad by the matrix goal *Bantee*. Both outcomes are possible with the very same matrix clause, headed by the verb 'tell'. See also (8) for the analogous two possibilities in Tamil.

(59) Magahi (fieldwork, Deepak Alok)

*a. Santee-aa Bantee-aa-ke kahk-**ain** ki Ram-ke Sita-se baat kareke chah-**ain**.*
Santee-FM Bantee-FM-DAT told.3.NH.S-HH.AL that Ram-DAT Sita-INS talk do.INF should.3.NH.S-HH.AL
"Santee told Bantee that Ram should talk to Sita." (to a teacher)

*b. Santee-aa **Bantee-aa-ke** kahk-ain ki Ram-ke Sita-se baat kareke chah-**au**.*
Santee-FM Bantee-FM-DAT told.3.NH.S-HH.AL that Ram-DAT Sita-INS talk do.INF should-3.NH.S-NH.AL
"Santee told Bantee that Ram should talk to Sita." (to a teacher)

Ad in a singly embedded clause thus has two choices: it can be controlled by the matrix goal, or it can have the same value as Ad*, the matrix Ad. It seems, then, that "obligatory control" is not really obligatory in this construction, even though Ad in (59) is in the context described by the GOCS: it is in the periphery of a CP generated inside the VP headed by 'tell'.

This observation can be refined and extended by looking at doubly embedded clauses, as in (60) from Magahi. Here the matrix Ad* is NH; the speaker is addressing a peer. The Ad in the middle clause can also be the NH form *-au*, as in (60b), or it can be the H form *-o*, if Ad in the middle clause is controlled by 'grandfather', the goal of the highest clause, as in (60a). These are the same two options as we see in (59), not taking into account yet that the embedded verb 'think' also selects a CP complement. The new observation is that the Ad in the lowest clause must match the Ad in the middle clause: it must be *-o* not *-au* in (60a) and it must be *-au* not *-o* in (60b). This pattern is expected in part: there is no goal in the middle clause built around 'think' that could control the Ad in the lowest clause. But there is new information in that *-au* on the lowest verb cannot match *-au* on the highest verb rather than *-o* on the middle verb in (60a).

(60) Magahi (fieldwork, Deepak Alok)

*a. Santee-aa <u>baabaa-ke</u> kahk-**au** ki Bantee-aa socha h-**o** ki Ram parichha paas ho ge-l-**o**/*ge-l-**au**.*
Santee-FM gr'father-DAT told.3.NH.S-NH.AL that Bantee-FM think be.3.NH.S-H.AL that Ram exam pass become go-PFV.3.NH.S-H.AL/*go-PFV.3.NH.S-NH.AL
"Santee told grandfather that Bantee thinks that Ram passed the test." (to a peer)

*b. Santee-aa <u>baabaa-ke</u> kahk-**au** ki Bantee-aa socha h-**au** ki Ram parichha paas ho ge-l-**au**/*ge-l-**o**.*
Santee-FM gr'father-DAT told.3.NH.S-NH.AL that Bantee-FM think be.3.NH.S-NH.AL that Ram exam pass become go-PFV.3.NH.S-NH.AL/*go-PFV.3.NH.S-H.AL
"Santee told grandfather that Bantee thinks that Ram passed the test." (to a peer)

The structure of the crucial example in (60a) is given in (61). We see is that Ad3 cannot be bound by Ad* past Ad2.[32]

(61) [Ad*$_n$ C S told gr'father$_i$ [Ad2$_i$ that B think [Ad3$_{i,*n}$ that R passed]]]
NH H H H,*NH

This shows that Ad cannot automatically denote the addressee regardless of its position in a syntactic structure. Only the matrix Ad, Ad*, automatically denotes the addressee; other Ads do so only by being bound (directly or indirectly) by Ad*. This binding of one Ad by another is subject to the locality condition stated in (62).

(62) If (Sp and) Ad are not controlled by an argument of the immediately superordinate verb, then they must be bound by the closest c-commanding (Sp and) Ad.

This asserts that in a representation like (61), Ad3 can be bound by Ad2 (controlled by 'grandfather'), but it cannot be bound by Ad* or

---

[32] The analog of this for indexical shift has been known since Anand & Nevins (2004), often discussed under the label "No Intervening Binder." See Chapter 4 for discussion.

left free. ((62) is stated for both Sp and Ad; the extension to Sp is justified by patterns of first person indexical shift in Chapter 4.)

In contrast, the obligatory control of SoK in African languages really is obligatory. There is no similar operator-binding-operator option for SoK in complement clauses in these constructions. Suppose that SoK could avoid being controlled by the matching argument of the superordinate verb and be bound by the higher SoK instead, parallel to (59a) in Magahi. Then we should be able to see C agreeing not with the closest higher subject, but with a further away subject—the one that controls the higher SoK, as in (63) with the structure in (64). But this is impossible.

(63) Lubukusu (Diercks 2013: 371)
Alfredi ka-a-loma ***a****-li* ***baba-andu*** *ba-mwekesia* ***ba-li/*a-li*** *omu-keni k-ola.*
CL1.Alfred CL1.S-PST-say CL1-C CL2-people CL2.S-revealed CL2-C/*CL1-C CL1-guest CL1.S-arrive
"Alfred said that people revealed that the guest arrived."

(64) [A_i said [SoK1_i that B_n revealed [SoK$2_{n,*i}$ that G arrived]]]
CL1 CL1 CL2 CL2/*CL1

Diercks and Rao (2019) show that the same kind of clause-level locality holds for upward C agreement with objects in Kipsigis, as discussed in Chapter 2. So Ad can be bound by the next closest Ad, but SoK (and OoK) cannot be bound by the next closest SoK (or OoK). This is one further difference between the two kinds of ghostly DP operators.

My analysis is that this difference between Ad and SoK reduces to the one we have already seen in adjunct clauses, discussed in the previous subsection. The connecting idea is CP extraposition. Suppose that complement clauses in Magahi can extrapose from complement position, such that they adjoin to (say) the TP projection of the clause they are generated in. This is a very common process across languages, and indeed it fits with the fact that finite CPs appear after the matrix verb (and all associated auxiliaries) in this otherwise head-final, S-XP-V language. After extraposition, the CP complement is in the same structural position as a high adjunct clause. Then we can simply assume that the optionality that underlies (59) and (60) is not an optionality in obligatory control per se (which would be

paradoxical) but a choice as to which copy of the complement clause is interpreted for the purposes of control theory. If the first-merge position inside VP is the chosen one, then obligatory control applies obligatorily, forcing Ad to be controlled by the goal of the complement-taking verb 'tell'. This results in (59b). However, if CP's derived position outside VP is the chosen one, then obligatory control does not apply: the clause is not merged with a projection of the lexical head. This is OK because OC is allowed to fail in Ad constructions in Magahi, since Ad has an interpretable [+2] feature. As such, it is not at risk of violating Full Interpretation when control does not happen. Rather, it can take the closest Ad as its binder, in accordance with (62).[33] The uncontrolled complement case in (59a) thus reduces to the high adjunct case in (51). I assume that the same analysis can be given for the optional allocutive shift in Tamil in (8), although the presence of CP extraposition is less obvious in this language.[34] This analysis is discussed further in §4.5.2, in the context of why indexical shift is optional in some languages but not others.

Now consider (63)/(64), the SoK side of the comparison. CP extraposition is presumably possible in the African languages too. It may be what derives the V-NP/PP-CP order (not V-CP-XP order) that we see in these languages, as in English. The CP can be interpreted for control in its first-merge position. Then SoK is subject to OC, so that it is controlled by a suitable argument of the matrix verb, as in the grammatical version of (63)/(64). However, the CP cannot be interpreted in the extraposed position in this case, given the principles we already have in place. Extraposition bleeds obligatory control, but SoK needs obligatory control; otherwise it fails Full Interpretation, as in (58a). The upshot is that extraposition does not create new possibilities for SoK for the same reason that SoK is not possible in

---

[33] It is important to assume that CP extraposition is clause-bounded: it can only adjoin CP to the TP of the clause that CP originated in (the Right Roof constraint of Ross (1967)). If the clause could move higher, then (62) would not be enough to explain the restriction in (60a).

[34] Embedded CPs are not often extraposed rightward in Tamil overtly. However, CP S V order is common in Dravidian languages, as an alternative to S CP V order. Thus, I conjecture that CPs extrapose leftward in Tamil. CP either adjoins above or below the subject, or the subject can optionally move still higher, to the left of CP. On this assumption, the analysis given for Magahi works for Tamil also. (I thank Sreekar Matam for discussion.)

clauses that are first-merged in high adjunct positions. In this way, this further difference between upward C-agreement and embedded allocutive constructions can also be explained in terms of the fundamental difference that SoK in C-agreement constructions lacks intrinsic phi-features, whereas Ad has them.

I note briefly that there is at least one language, innovative Southern Basque, which allows allocutive marking in embedded clauses but does not allow shifted allocutive marking, where it is determined by the relationship between the arguments of the matrix clause. This was seen in (9), repeated here as (65).

(65) Basque (Haddican and Etxebarria 2022; Extebarria p.c.)

*a. Jon-ek Imanol-i [etorri-ko du-**k**-ela] esa-n zi-o-**k**.*
Jon-ERG Imanol-DAT come-FUT AUX-2.SG.M.AL-C say-PFV AUX-3.SG.DAT-2.SG.M.AL
"Jon told Imanol that he will come." (to a male friend)

*b. *Jon-ek Miren-i [etorri-ko du-**na**-la] esa-n z-i-o-**k**.*
Jon-ERG Miren-DAT come-FUT AUX-2.SG.F.AL-C say-PFV AUX-3.SG.DAT-2.SG.M.AL
("Jon told Miren that he will come." to a male friend)

A conceivable analysis of this within my framework is stipulating that CP extraposition is obligatory and not subject to reconstruction in this language. But this is not a particularly attractive view, both because the complement clause is not visibly extraposed in Basque the way it is in Magahi, and because it is not clear what plausible parameter could block the reconstruction. A more promising analysis is to posit that there is a nominal layer over the CP in Basque, which is not present in Magahi or Tamil, and this additional structure disrupts the obligatory control of Ad. (See §4.5.3 for discussion of the fact that nominalization blocks the control of Sp and Ad and hence indexical shift in Turkic languages.) This conjecture is consistent with the fact that the matrix subject in (65a) bears ergative case; if this is a dependent case (Marantz 1991, Baker 2015), it shows that the clause counts as a nominal object.[35] I leave further investigation of this to

[35] However, Basque ergative case marking is not the most straightforward example of dependent case, and might be taken to be an inherent case more reasonably then in some other languages. See Baker & Bobaljik (2017) for discussion.

future research by those who know more about Basque syntax than I.

One loose end in this account is (62), the statement that when Ad is bound rather than controlled, it must be bound by the closest superordinate Ad. I want to treat this relationship as an instance of the bound variable anaphora that is typical of pronouns, because it is not subject to absolute syntactic locality conditions like the PIC (see again (51) and (55)). However, bound variable anaphora is not generally subject to a relativized minimality-like constraint like (62) either (cf. the possibility of examples like *Every girl$_i$ thinks that every mother$_k$ should acknowledge that she$_i$ is talented*). My proposal about this is that, while it is true that third person pronouns are not subject to such a condition, there is independent reason to think that first and second person (+participant) pronouns are. First and second person pronouns are known to be subject to some extra conditions that simple third person pronouns are not. Kratzer (2009) and others have shown that participant pronouns can participate in bound variable anaphora. Baker (2008: 126) proposed the principle in (66) for such pronouns, calling it the Person Licensing Condition (PLC).[36]

(66) a. A first person pronoun must be locally bound by the closest c-commanding element that is [+1] (an instance of Sp or another first person pronoun).
b. A second person pronoun must be locally bound by the closest c-commanding element that is [+2] (an instance of Ad or another second person pronoun).

This is motivated for the ordinary pronouns ‘I’ and ‘you’ in argument positions in Magahi and some other indexical shift languages.[37] A

---

[36] This term has been used by others in different ways—for example, by Béjar and Rezac (2003) in their theory of the Person Case Constraint. There is some connection between the two PLCs, but it is a rather distant one, Bejar and Rezac’s version insisting that first and second person pronouns trigger agreement on a suitable head, rather than that they be bound by a particular operator.

[37] There are superficial examples of this sort in languages like Amharic and Mishar Tatar, but Deal (2020) treats them as having indexiphors, not true indexicals. I discuss this in Chapter 6.

vivid illustration is the badness of examples like (67) in Magahi.

(67) Magahi (fieldwork, Deepak Alok)

*a. *Santee soch-l-ai ki (ham) hamraa dekh-l-i.*
Santee think-PFV-3.NH.S that I me.ACC see-PFV-1.S
("Santee thinks that I saw me.")
Sp_i* $Santee_k$ thinks [Sp_k that [$I_{i,k}$ saw $me_{k,i}$]].

*b. *Santee-aa Bantee-aa-ke kah-l-ai ki (tu) toraa dekh-l-eN.*
Santee-FM Bantee-FM-DAT tell- PFV-3.NH.S that (you.NH) you.NH.ACC see-PFV-2.NH.S
("Santee told Bantee that you saw you.")
Ad_i* Santee told $Bantee_k$ [Ad_k that [$you_{i,k}$ saw $you_{k,i}$]].

Without a condition like the PLC, one might expect these examples to be possible with meanings like 'Santee thinks that he saw me' (or 'Santee thinks that I saw him') and 'Santee told Bantee that he saw you' (or 'Santee told Bantee that you saw him'). This would be the result of one of the participant pronouns in the complement CP being bound by Sp*/Ad* and the other one being bound by the controlled Sp/Ad in the complement CP. But this is impossible, a kind of Shift Together violation. (66) plays a role in explaining why.[38] This forces the two first person pronouns in (66a) to corefer with *Santee*, the controller of the closest c-commanding Sp, and the two second person pronouns in (66b) to corefer with *Bantee*, the controller of the closest c-commanding Ad. These interpretations then violate Condition B of the Binding theory, since a nonanaphoric pronoun is locally bound by a coreferential DP in the same clause, on a par with **I scratched me* and **You scratched you* in English. So there is reason to think that the binding properties of [+participant] pronouns are more restricted than those of [-participant] pronouns in the way described by the PLC.

I did not originally envision the PLC applying to Sp and Ad as bindees rather than as binders of ordinary pronouns. However, given the current hypothesis that Ad is itself second person (and Sp is first person), it naturally falls under this principle. This gives us what we need to derive the descriptive generalization in (62) and thus strengthen the analysis of (60)/(61). As it is a kind of second person

[38] See Chapter 4 for some discussion of alternatives, in particular, the context overwriting analysis of Anand (2006) and Deal (2020).

pronoun, (66b) implies that Ad must be bound by the closest higher [+2] Ad, the immediately superordinate one. It is a bit paradoxical that binding Ad is relatively unconstrained—it can reach across phases and into islands—but not totally unconstrained—it is subject to a kind of relativized locality condition as described by the PLC. The obligatory control of SoK is more tightly constrained, whereas normal (third person) pronoun binding is less constrained. But this is the behavior that local pronouns are already known to have.

5. Conclusion

This chapter has investigated allocutive agreement constructions, with an emphasis on rich new data from the Magahi language, but with an eye also on Basque, Tamil, and other languages. I have adopted the generative tradition of saying that this is the result of a functional head in the CP space agreeing with a null DP that refers to the addressee of the sentence—with Ad (also called Hr) in the sense of Speas & Tenny (2003). What is particularly interesting about Maghai is that this Ad (and hence allocutive marking) is possible in embedded clauses as well as in matrix clauses, and when it appears in embedded clauses it can be controlled by a suitable argument of the matrix verb. On the one hand, allocutive marking in these clauses turns out to be essentially a variant of upward C-agreement, making a connection to the material discussed in Chapter 2. In particular, the same principles of control theory apply to both constructions, including the GOCS and the thematic role matching principle. On the other hand, allocutive marking in this context becomes a kind of shifted allocutive marking, which is both conceptually parallel to ordinary indexical shift and empirically intertwined with it. This makes a connection to the material discussed in Chapter 4.

Whereas licensing, control, and agreement work in recognizably the same way in both upward C-agreement constructions and allocutive constructions, there are nontrivial differences between them as well. I argued that these differences are the result of Ad intrinsically bearing phi-features, notably [+2], whereas SoK is devoid of phi-features, which forces it to get them by undergoing OC. From this, the general outlines of a typology of ghostly DP operators begins to come into view, in which ghostly DPs vary as to whether they have phi-features at all, and if so, which ones.