Chapter 4: Indexical Shift as the Control of Ghostly Nominals

1. Introduction

My exploration of rare constructions in which complementizers relate to the nominal phrases around them began with upward C-agreement in some African languages and moved from there to allocutive (addressee) agreement in a range of languages, but especially Magahi where allocutive marking is robust in embedded clauses. The next step is to extend this theory to indexical shift constructions. Recall that these are constructions in which first and second person pronouns inside an embedded clause are understood as referring to the subject or object of a superordinate clause, rather than to the speaker or the addressee of the sentence as a whole. This phenomenon is found in a wider range of languages than either upward C-agreement or full allocutivity, although still apparently only in a minority of languages. The list includes Amharic (Leslau 1995, Schlenker 1999, Schlenker 2003, Anand 2006), Zazaki (Anand and Nevins 2004, Anand 2006), Slave (Rice 1989), Uygur (Sudo 2012, Shklovsky and Sudo 2014, Major 2021, Major 2022) and other Turkic languages, and Nez Perce (Deal 2020) as well as Magahi (Alok and Baker 2018, Alok 2020) and others (see Deal 2020 for comprehensive references). Some speakers of Korean (Park 2016) and Japanese (Shiori Ikawa, p.c.) allow it as well, although others do not. There has also been a larger theoretical literature on this topic, especially in semantics.

My hypothesis is that these constructions involve the same kind of licensing and control of ghostly DP operators as upward C-agreement and allocutive agreement do. Indeed, I argue that the ghostly DPs involved in indexical shift are the very same ones that are involved in allocutive marking and analogous speaker agreement: the intrinsically second person element Ad and the intrinsically first person element Sp. The only difference is that the possibility of C-space heads agreeing with these ghostly DPs now fades into the background, while the possibility of the ghostly DPs binding pronouns inside the clause selected by the C-head that licenses them comes to the fore. Magahi

plays a special role in this discussion because it has both allocutive agreement and indexical shift, making it possible to observe that the two phenomena are deeply intertwined, as originally argued by Alok & Baker (2018). However, I argue that the same theory of indexical shift can be used in other languages in which the C heads happen not to be probes for Agree, with the result that they do not have allocutive marking, but only indexical shift. Indeed, the interaction with allocutive marking is close to the only thing that is empirically remarkable about indexical shift in Magahi; otherwise, its indexical shift patterns fall comfortably within the bounds of what is known about indexical shift in other languages. Therefore, an effective theory of the Magahi phenomenon is well on its way to being an effective theory of indexical shift more generally.[1]

Since the ghostly DP operators that are involved in indexical shift constructions are the same ones that are involved in allocutive marking, there is relatively little new theoretical machinery that needs to be introduced in this chapter to get started. I thus begin by recapping the discussion from Chapter 3 about how the standard generative theory of allocutive agreement leads very naturally to a (nonstandard) theory of indexical shift in Magahi.

The story starts with the observation that in some languages allocutive marking is a form of agreement, parallel to ordinary T-agreement with the subject. Given standard assumptions about agreement, this implies that there must be a DP in the representation of the sentence that the relevant head is agreeing with—a DP that, in ordinary root clauses, refers to the addressee of the sentence (Speas and Tenny 2003). However, in some embedded clauses, this agreed-with DP does not express features of the addressee of the sentence as a whole, but rather those of the goal argument of the matrix clause. This covaluation relationship has properties that can be attributed to the theory of obligatory control, particularly given its similarities with the behavior of the proximal target of upward C-agreement in African languages, and (at more remove) with the behavior of controlled PRO in European languages as expressed by the Obligatory Control Signature (Landau 2013). At the same time, there are both formal and semantic

---

[1] The other somewhat distinctive thing about Magahi indexical shift is that overt first person pronouns and null ones (pro) behave a bit differently—a fact that I take to shed some new light on Shift Together phenomena; see §4.5.4 and §6.7.

reasons to say that the crucial DP has a second person feature. Formally, it shows the three-way honorification distinctions seen only with second person elements in Magahi. Semantically, it refers to the addressee in simple sentences.

Up to this point, this is simply an implementation of the standard generative theory of allocutive agreement. The turn toward indexical shift is the idea that the null second person DP in the periphery of the clause (Ad) can, like any other DP, bind a pronoun inside its c-command domain. The bound pronoun matches the DP that binds it in phi-features; therefore, it too is second person. But when Ad is controlled by a matrix argument, the bound pronoun refers to the matrix argument, not the addressee of the sentence as a whole. Voila, a theory of second-person indexical shift (which I abbreviate *u-shift*). This is sketched in (1b) for the canonical Magahi example in (1a).

(1) Magahi (fieldwork, Deepak Alok)

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

b. Ad_k Fin [S tell B_i [Ad_i Fin [R saw pronoun$_i$]]
2HH — Agree; 2NH — Agree; 2NH (*3NH)
control; binding

There are of course other theories of u-shift, which have different starting points from this one. For example, there is the context-shifting operator theory championed by Anand (2006) and Deal (2020), among others. I compare the theory just sketched with the shifty operator theory at various points in this chapter, as they come up, and then somewhat more systematically in its final section, §4.6. But the theory sketched in (1) is one that emerges quite naturally from assumptions that we need anyway, so it is well-worth serious consideration. Moreover, this is a contender not only for Magahi but for other languages that have indexical shift without allocutive marking. The analysis in (1b) could very well hold for them too, with the simple difference that Fin happens not to be a probe for agreement in such

languages, as it is not in the standard IE languages.

This account can also be generalized to first person indexical shift (henceforth abbreviated *i-shift*), where pronouns like 'I' and 'me' are interpreted as referring to a second or third person nominal in the matrix clause. Magahi does not have "speaker agreement" on a par with its overt allocutive agreement, although other languages like Dargwa arguably do (see §3.3.1). However, Magahi does have first-person indexical shift. In fact, this happens under the same verbs that allow second person indexical shift (among others), as in (2).[2]

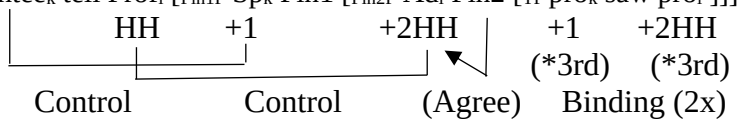
(2) Magahi (fieldwork, Deepak Alok)
Santee-aa profesar saaheb-ke kahk-ai ki ham apne-ke dekh-l-i-ain ha-l.
Santee-FM professor HH-DAT told-3.NH.S that I you.HH.ACC see-PFV-1.S-HH.AL be-PFV
"Santee$_i$ told the professor$_k$ that he/I$_{i,*sp*}$ saw him/you$_{k,*ad*}$."

This is parallel to the u-shift in (1a). Indeed, i-shift is entangled with u-shift, as in the classic Shift Together effect of Anand & Nevins (2004) and Anand (2006): 'I' shifts if and only if 'you' shifts, which in turn shifts if and only if allocutive marking does (if any). It is appropriate, then, to generalize the account of u-shift in (1) to give a parallel account of i-shift. The Fin head selects a second ghostly DP, Sp, which has intrinisic first person features and which denotes the speaker in matrix clauses. (More precisely, Fin1 selects Sp and Fin2 selects Ad; see §3.3.1.) In some embedded clauses, this Sp is controlled by a suitable argument of the matrix verb, typically its agent argument, in accordance with the principles of obligatory control. This Sp can (but need not) also bind a pronoun inside the CP headed by the Fin that licenses it. Such a pronoun will match Sp in phi-features, so it will be first person. In matrix clauses, where Sp denotes the speaker, so does the first person pronoun that it binds. But in embedded clauses where Sp is controlled by the matrix agent, the

[2] For indexical shift examples, I gloss a participant pronoun with a third person antecedent with a slashed translation like 'he/I' or 'her/you'. The third person translation gives a sense of what the sentence can mean, while the first or second person pronoun reflects its form better. This reinforces the information present in the indices. I use sp* as a special index for reference to the speaker and ad* as a special index for reference to the addressee of the root sentence.

first person pronoun that it binds is also bound by that agent. The result is i-shift, as shown in (3). It happens that no C-head in Magahi agrees with Sp, but apart from that the analysis of i-shift is point-by-point parallel to the emergent analysis of u-shift presented in (1).

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



This analysis was first presented in Chapter 3 in a supporting role, as evidence that the operators involved in allocutive agreement have intrinsic [+2] features and those involved in speaker agreement have [+1] features, unlike the ghostly DPs involved in upward C-agreement in the African languages. But now it takes center stage, as an important topic in its own right. Indeed, indexical shift seems to be more widely distributed in the languages of the world than allocutive marking is, and it has been a weightier topic for linguistic theory since Schlenker (1999), especially in the semantics literature.

The discussion unfolds as follows. §4.2 briefly supports my claim that Magahi has true indexical shift, rather than just direct quotation. §4.3 then shows that Sp and Ad are controlled by the same principles of control theory as SoK and OoK are—the Generalized Obligatory Control Signature and the thematic role matching condition (but not the T/Agree Condition). For Ad, this was already shown using evidence from allocutive marking in Chapter 3; here I replicate the result using converging evidence from u-shift. For Sp, this is a new result, which can only be established using evidence from i-shift. Indeed, i-shift is possible with a wider range of matrix verbs than u-shift is, since it happens with dyadic matrix verbs like ‘think’ as well as with triadic verbs like ‘tell’. Furthermore, the control of Sp is more obviously parallel to the control of SoK in the African languages, since the thematic roles involved are the same. Thus, the control paradigm that I can present in this chapter is significantly richer than what I could show in Chapter 3, making the case that obligatory control is at work significantly stronger. §4.4 then goes on to considers the new issues that are raised by the pronominal binding relationships in (1b) and (3), this being the newer theoretical ingredient in this analysis. It reintroduces and explores the consequences of the Person Licensing Condition, which was introduced at the end of Chapter 3 but is more prominent here where

the focus is on bound pronouns. §4.5 faces the issue of optionality in indexical shift constructions, including the Shift Together phenomenon, reconciling it with the obligatoriness of obligatory control. I claim that control into true CP complements is in principle obligatory, and we see this obligatoriness on the surface in some languages, like Uyghur. However, this OC can be bled by either CP-extraposition, which we know happens in Maghai, or by nominalization—both processes that affect whole clauses. Finally, §4.6 surveys briefly the other ways that indexical shift is known to vary across languages and compares my analysis further to the shifty operator analysis of Anand and Deal.

2. Indexical Shift not direct quotation

In this preliminary section, I briefly present three arguments that Magahi has genuine indexical shift, not just direct quotation, drawing from previous work (Alok and Baker 2018, Alok 2020). I also reconfirm a result from Chapter 3, that having Sp and Ad operators present and controlled in a clause does not have the kinds of semantic effects that having a controlled SoK does in languages like Kinande.

Empirically, one needs to show that examples like (1a) and (2) contain true indexical shift rather than direct quotations, as comparison with English might suggest. There are standard tests for doing this, which Alok and Baker (2018) and Alok (2020) apply to Magahi. One such test is whether a question word associated with the embedded clause can take scope over the matrix clause to form a direct question. With English direct quotation this is impossible, but in Magahi and other languages with indexical shift, it is possible, as seen in (4).[3]

(4) Magahi (fieldwork, Deepak Alok)

a. Kab Ram soc-l-ai [ki ham mar-b-ai]?
when Ram think-PFV-3.NH.S that I die-FUT-3.NH.S

[3] This test works best with extracting adjunct *wh*-phrases. Questioning an argument of the verb in the embedded clause tends to use an indirect dependency/scope marking construction in Magahi, and this interferes with indexical shift; see Alok (2020) for some discussion. Also, whether wh-in-situ is possible or not with matrix scope seems to be somewhat variable. However, when using adjunct extraction one needs to be careful to make sure that the adjunct ('when' in (4a,b)) is interpreted with respect to the embedded clause.

"When does Ram_i think that he/$I_{i,sp*}$ will die?" (time of dying questioned)

b. Kab Santee-aa Raam-ke kah-l-ai ki tu mar-b-a?
when Santee-FM Ram-DAT tell- PFV-3.NH.S that you.NH die-FUT-2.NH.S
"When did Santee tell Ram_i that he/$you_{i,Ad*}$ will die?" (time of dying questioned)

(4a,b) also imply that the Sp and Ad in Spec FinP do not create wh-islands the way that question phrases in Spec CP do in languages like English; see §3.2.1 for some discussion.

A second standard test for true indexical shift is based on the licensing of negative polarity items (NPIs). It involves putting an NPI in the embedded clause which is licensed by negation in the matrix clause. In this case, the embedded clause would not be grammatical on its own as a sentence which could be quoted. (5) shows that when this is done in Magahi shifted readings for the indexical pronouns are still possible, whether 'I' in (5a) or 'you' in (5b).

(5) Magahi (fieldwork, Deepak Alok)

*a. Bantee-aa-ke **na** laga h-ai ki hamraa **kuchhbhii** almaari me milt-ai.*
Bantee-FM-DAT NEG seem be-3.NS.S that I.DAT anything closet in find-3.NH.S
"It doesn't seem to $Bantee_i$ that he/$I_{i,sp*}$ will find anything in the closet."

*b. Santee-aa Banteeaa-ke **na** kah-kai ki toraa **koi kitaab** paRhe-ke chah-ai.*
Santee-FM Bantee-FM-DAT NEG tell-3.NH.S that you.DAT any book read-ACC should-3.NH.S
"Santee didn't tell $Bantee_i$ that he/$you_{i,ad*}$ should read any book.'

In addition to these two tests which have been applied to many languages, Alok and Baker (2018) also present a more Magahi-specific test for direct quotation that involves allocutive marking. It turns out that there is a negative interaction between indexical shift and allocutive agreement in the complement of a dyadic verb like 'think' or 'say' in Magahi. As a result, 'I' in the embedded clause of a sentence like (6) can refer to the same person as the subject of the

matrix clause only if allocutive marking is absent on the embedded verb—so yes in (6a) but no in (6b).

(6) Magahi (fieldwork, Deepak Alok)

*a. John kahk-au ki **ham** tej h-i.*
John said-NH.AL that I smart be-1.S
"John$_i$ said that he/I$_{i,sp*}$ am smart." (said to a peer)

*b. John kahk-au ki **ham** tej h-i-**au**.*
John said-NH.AL that I smart be-1.S-NH.AL
"John$_i$ said that I$_{sp*, *i}$ am smart." (said to a peer)

This restriction is unexpected if there is no true indexical shift in Magahi. Then 'I am smart" would have to be a direct quotation in the relevant reading of (6a). Then the fact that allocutive agreement is ungrammatical with this reading, as seen in (6b), is inexplicable, given that *Ham tej hiau* is a fine thing for someone to say in Magahi. It is a very normal way to say 'I am smart' (if talking to a friend). So no explanation of the restriction in (6) would be forthcoming if Magahi allows for direct quotation but not indexical shift. In contrast, on the view that Magahi has true indexical shift, the pattern in (6) will be explained as a kind of defective addressee effect (see §4.5.4).

Indeed, this suggests a stronger conclusion of practical importance to this study: that examples with the form of (6) *cannot* easily be read as containing direct quotations. In this, *ki* clauses in Magahi are rather like clauses introduced with *that* in English, which also cannot be direct quotations (*Chris said (*that) "I am tired."*).[4] Given this, I assume that confusing a direct quote with a syntactic complement with indexical shift is not a very serious danger in Magahi. Therefore, I do not include wide scope adjunct questions or negative polarity items in most of my examples, except when special care might be needed for theoretical reasons. This is a good thing, since doing these tests is far from cost free. Both long distance adjunct extraction and long distance negative polarity licensing bring with them a considerable additional

[4] It is probably too strong to say that a clause introduced by the C-like element *ki* can never be interpreted as a direct quotation in Magahi. It seems that it can be, but this probably requires special intonational marking that does not immediately occur to speakers when presented with sentences like these in neutral contexts. Alok and I have not investigated the intonational factors involved in this.

processing burden in Magahi, making it harder to judge these already complex and potentially ambiguous sentences. It can be done when necessary, but it is good not to have to do so for every example.

The other preliminary to take care of here is the semantics of indexical shift—or lack thereof. In §3.2.1, I showed that including optional allocutive agreement on an embedded clause in Magahi does not have the same semantic consequences as including optional C-agreement on an embedded clause in Kinande and other African languages (§2.2), which was attributable to the presence of SoK and the Eval head. This conclusion is reinforced by data from indexical shift. Some Kinande verbs such as 'want' require an agreeing complementizer, the reflex of a controlled SoK, but no known Magahi verb requires first person indexical shift, the reflex of a controlled Sp. For example, 'want' does not, as seen in (7); rather i-shift is optional with this verb.

(7) Magahi (fieldwork, Deepak Alok)
Santee-aa chaaha h-ai ki (pro) parichha paas ho jaa-i.
Santee-FM want be-3.NH.S that (I) exam pass become go-1.S
"Santee$_i$ wants that he/I $_{i,sp*}$ pass the test."

Conversely, some Kinande verbs such as factives forbid an agreeing complementizer, the reflex of a controlled SoK, but no known Magahi verb forbids first person indexical shift, the reflex of a controlled Sp. For example, 'be angry' ('resent') does not, as seen in (8).

(8) Magahi (fieldwork, Deepak Alok)
Santee-aa gossaayel h-ai ki ham parichha na paas ho-l-i.
Santee-FM angry.PFV be-3.NH.S that I exam not pass become-PFV-1.S
"Santee$_i$ is angry that he/I $_{i,sp*}$ did not pass the test."

Finally, some Kinande verbs with very general meanings optionally allow an agreeing C in their complement, but it has implications for who is taken to be the source of the content of the clause and who is committed to the truth of that content. In Magahi, verbs with general meanings optionally allow i-shift in their complements, but this has no implications for who is committed to the content of the clause. For example, i-shift is optional in (6a), but the choice does not affect whether the information "I am smart" comes from Santee's own deluded imagination, or from hearsay, or is common knowledge. The subject controlling SoK which results in C-agreement goes with an

interpretation in which the subject has a special responsibility for the content of the CP in the African languages, but the subject controlling Sp which results in i-shift does not go with this interpretation in Magahi—or other indexical shift languages, as far as has been described. The same holds true for u-shift: this is possible with any CP-selecting verb that also takes a goal argument, but it is never required, and there is no detectable meaning difference that goes with the optionality comparable to (say) the effect of verum focus that goes along with C-agreement with the object in Kipsigis according to Diercks and Rao (2019). This reinforces the conclusion that Sp and Ad are licensed by different C-like heads, with different semantics from SoK. Sok (and perhaps OoK) are licensed by Eval, whereas Sp and Ad are licensed either by sa (speech act) heads in the root clause, which have special performative meanings, or by Fin heads in embedded clauses, which have little or no detectable meaning.

3. Control of Sp and Ad as seen from indexical shift

3.1. Introduction

I already discussed what Sp and Ad are intrinsically, and how they are licensed by particular C-type heads in Chapter 3, partly reprised above. With this in place, there are two grammatical relationships between Sp and Ad and other elements to explore, given the schematic structure in (3): the control relationships between Sp and Ad and suitable arguments of the matrix verb, and the binding relationships between Sp and Ad and participant pronouns within their scope. Of the two, the binding relationship is the newer ingredient within my analytical framework, not intrinsically part of upward C-agreement constructions or allocutive constructions. I take up this topic second, in §4.4. I start in this section with the more familiar control relationship, which has already been discussed at length for the other ghostly operator constructions. I have shown how a distinctive pattern of facts follows from two relatively simple principles of (generalized) control theory. First, the GOCS states that obligatory control happens only between a null/controllable DP generated at the edge of a clause merged with the projection of some lexical head X (usually a verb) and some argument of X. Second, a matching condition states that the thematic roles of the controller and the controlled DP must be (nearly) the same. More specifically, we saw in §3.4.2 that the controller and

the controllee must match in their macroroles: both must be initiators (proto-agents) or both must be undergoers (proto-patients).[5] We have already seen that these results hold for the control of Ad as much as was feasible using data from allocutive agreement. Now, I show that they are confirmed and extended by the richer set of data that comes from i-shift and u-shift. Since i-shift happens with a wider range of matrix predicates than u-shift does, and since Sp is more comparable to SoK than Ad is, the comparison becomes fuller and more compelling in support of the hypothesis that indexical shift involves the obligatory control of syntactically represented null DPs. If Sp and Ad show evidence of being controlled by the same principles as SoK, which agreement shows to be syntactically realized, there is advantage to saying that Sp and Ad are syntactically realized too, subject to the same principles as syntactic elements like SoK and ordinary PRO.

This result is theoretically significant in two respects. First, it is another major step forward in justifying the GOCS and the thematic matching condition as pillars of a robust generalized control theory, making it worthwhile to rethink the theory of control from this perspective, as I do in Chapter 8. Second, it should go long way toward motivating/solidifying/confirming the theory of indexical shift proposed in Baker and Alok (2018) and Alok (2020). My analysis of C agreement is a variant of one of the standard views in that literature (Diercks's 2013). My analysis of allocutive agreement is a standard view as well—essentially the only existing generative view. But the analysis of indexical shift outlined in §4.1 is not standard. It is rather different from the influential Anand/Deal analysis, in which indexical shift is accomplished by a context shifting C-like head, without pronoun binding, control, or a CP-peripheral DP being involved. But if African C-agreement needs to involve control of one DP by another and cannot be the result of purely semantic operators, and African C-agreement is demonstrably like Magahi indexical shift such that the two should fall under the same theory, then we can infer that there should be a more control-like theory of indexical shift too. This is a higher-level argument that I am developing here.

---

[5] In this section, we also see evidence that the control of Sp is not subject to the T/Agree Condition: the controller of Sp does not need to trigger agreement on T in the matrix clause. This is expected given the analysis in §3.3.2.

3.2. Thematic matching

Recall that a signature property of C-agreement in Africa is that the superordinate subject triggers it and the superordinate object does not—despite the object being structurally closer to the C than the subject is. An analogous pattern holds for i-shift in Magahi: the matrix subject of a verb like ‘tell’ can control Sp, as seen again in (9), but the object of ‘tell’ cannot. If it could, we would expect ‘I’ to shift to corefer with *Bantee*, the goal of the matrix verb, but this is impossible.

(9) Magahi (fieldwork, Deepak Alok)
Santee-aa *Bantee-aa-ke kahl-ai ki* ***ham*** *tej h-i.*
Santee-FM Bantee-FM-DAT told-3.NH.S that I intelligent be-1.S
“Santee$_i$ told Bantee$_k$ that he/I$_{i,sp*,*k}$ am intelligent.”

In contrast, the goal can control the Ad of its CP complement, such that ‘you’ in the embedded clause bound by Ad shifts to the goal Bantee, as in (10). This is like upward C-agreement with the goal in Kipsigis, assuming that to be a real phenomenon.

(10) Magahi (fieldwork Deepak Alok)
Santee-aa ***Bantee-aa-ke*** *kahl-ai ki* ***tu*** *tej h-eN.*
Santee-FM Bantee-FM-DAT told-3.NH.S that you.NH intelligent be-2.NH.S
“Santee$_i$ told Bantee$_k$ that he/you$_{k,ad*,*i}$ are intelligent.”

Conversely, the matrix subject cannot be the understood antecedent of a shifted ‘you’, as can also be seen in (10). This is like the fact that the subject of ‘tell’ cannot control suffixal C-agreement in Kipsigis, but only prefixal C-agreement. So the high-level parallel between C-agreement and indexical shift holds over these central data. This is in line with the thematic matching condition, under the assumption that Sp gets an agent-(like) thematic role from Fin(1) and Ad gets a goal-(like) role from Fin(2).

The idea that it is an agent that controls Sp and a goal that controls Ad should not be interpreted too narrowly. The subject of (what can be used as) a stative, nonvolitional verb like ‘think’ is like the agent of ‘tell’ in controlling Sp resulting in i-shift, as shown in (11).

(11) Magahi (fieldwork Deepak Alok)

Santee-aa soch-l-ai ki ham tej h-i.

Santee-FM think-PFV-3.NH.S that I intelligent be-1.S

"Santee$_i$ thought that he/I$_{i,sp*}$ am intelligent."

There are narrower senses of the term "agent" in which the subject of (11) does not qualify; rather terms like "holder" are sometimes used for these cases. But those narrow senses are not the ones that we want for control of Sp in Magahi (or SoK in the African languages). Similarly, the object of the matrix verb 'convince' is probably a patient argument rather than a goal argument, given that its referent necessarily undergoes a change of (mental) state as part of the convincing event, whereas the object of 'tell' does not.[6] However, the theme of 'convince' is like the goal of 'tell' in that it can control Ad in the complement clause resulting in u-shift, observable in (12).

(12) Magahi (fieldwork, Deepak Alok)

Bittu-aa Chhotu-aa-ke soch-wal-k-ai ki (pro) toraa dekhl-i.

Bittu-FM Chhotu-FM-ACC think-CAUS-PFV-3.NH.S that (I) you.ACC see.PFV-1.S

"Bittu$_n$ convinced Chhotu$_i$ that he/I$_n$ saw him/you$_i$."

We thus need to define matching over more coarse-grained thematic roles like initiator and undergoer (Foley and Van Valin 1984, Ramchand 2008) or proto-agent and proto-patient (Dowty 1991), as mentioned already in §3.4.2. Agents and holders are both subtypes of initiators, and patients and goals (in the context of agents) are both subtypes of undergoers.

One classic way to see that controller choice is determined more by the thematic role(s) of the potential controller than by its syntactic position/grammatical function is to consider examples in which the matrix verb is passive. For example, in Kinande the passivized goal argument cannot control SoK, hence agreement on C, because the covert agent or the *by*-phrase is a better thematic match. Similarly, in Magahi the goal argument of a passive is not able to control Sp, hence it cannot be the antecedent of a shifted 'I'. This is seen in (13). Here

---

[6] Note that case marking is not very useful for distinguishing goals from patients in Magahi, since the same postposition *-ke* that marks dative case is used as a differential object marker with definite/human direct objects, as in Hindi.

Chhotu cannot be the antecedent for shifted ‘I’ in the CP complement for essentially the same reason that *Bantee* cannot be in (9).

(13) Magahi (fieldwork, Deepak Alok)
Chhotu-aa-ke kahal ge-l-ai ki (ham) Ram-ke madad kar-bai.
Chhotu-FM-DAT told go-PFV-3.NH.S that (I) Ram-ACC help do-FUT.3.NH.S
“Chhotu$_i$ was told that I$_{*i,sp*}$ will help Ram.”

This is not as striking in Magahi as it is in the African languages, because the goal argument in (13) retains dative case and (therefore) does not trigger person agreement on the verb. In these respects, it is not a fully canonical subject. However, this is still a similarity with African languages like Kinande rather than a difference. Moreover, the dative experiencer argument of verbs like ‘seem’ and ‘remember’ can control i-shift, as shown in (14).

(14) Magahi (fieldwork Deepak Alok)
a. Santee-aa-ke laga h-ai ki ham tej h-i.
Santee-FM-DAT seem be-3.NH.S that I smart be-1.S.
“It seems to Santee$_i$ that he/I$_{i,sp*}$ is/am smart.”

b. Ram-ke yaad ha-l-ai ki ham almira-me paisa chhupai-l-i ha-l.
Ram-DAT memory be-PFV-3.NH.S that I drawer-in money hide-PFV-1.S be-PFV
“Ram remembered that he/I$_{i,sp*}$ hid the money in the drawer.”

Thus-i-shift in (13) cannot be ruled out purely on superficial morphosyntactic grounds involving the case of the antecedent of ‘I’. Rather, it is plausible to say that the covert agent plays a role in blocking i-shift controlled by ‘Chhotu’ in (13).

What the goal argument of passivized ‘tell’ can do is control Ad, just as it does in the active version in (10). As a result, it can be the ultimate antecedent of shifted ‘you’ in the complement, as in (15).

(15) Magahi (fieldwork, Deepak Alok)
Chhotu-aa-ke kahal gel-ai ki (tu) Ram-ke madad kar-beN.
Chhotu-FM-DAT told go.PFV-3.NH.S that (you.NH) Ram-ACC help do-FUT.2.NH.S
"Chhotu$_{i}$ was told that he/you$_{i,ad*}$ will help Ram."

Even more strikingly, the *by*-phrase in a Magahi passive can control Sp, making it the antecedent of i-shift, as in (16).[7]

(16) Magahi (fieldwork, Deepak Alok)
Chhotu-aa-ke Bittu-aa diyaa ?kah-al/kah-waa-wal gel-ai ki (pro) toraa dekhl-i.
Chhotu-FM-DAT Bittu-FM by tell-PASS/tell-CAUS-PASS go-3.NH.S that (I) you.ACC saw-1.S
"Chhotu$_{i}$ was told by Bittu$_{k}$ that he/I$_{k,sp*}$ saw him/you$_{i,ad*}$."

The control/indexical-shift pattern in these passives is essentially the same as in versions with the active matrix verb 'tell': the agent controls i-shift and the goal controls u-shift, regardless of their surface grammatical functions. This supports the claim that control is determined thematically.

I note that these data also show clearly that in Magahi control of Sp is not subject to a T/Agree Condition, the way that C-agreement in the Niger Congo languages is. Neither the *by*-phrase agent of the passive in (16) nor the dative experiencer of 'seem' and 'remember' in (14) can trigger agreement on the matrix verb in Magahi. They are prevented from doing so by the fact that they are marked by postpositions or oblique case markers, given that agreement in Magahi is case sensitive (cf. Bobaljik 2008, Baker 2015). Nevertheless, both can control Sp and be the antecedent of i-shift (as can source and causee arguments in some cases; see below).

The dative subject verbs of 'seem' and 'remember' provide an

---

[7] Like many South Asian languages, the true simple personal passive is not common in the spoken language, especially with a *by*-phrase. The version in which the passive verb bears causative morphology as well as periphrastic passive morphology (the main verb in perfective particle form *-l* together with intransitive auxiliary 'go') is more natural. This morphological variation goes not affect the syntactic points being made, as far as I know.

instructive minimal comparison with 'tell- PASS' that illustrates some of the details of the theory of thematic roles that I adopted in Chapter 3. There I adopted the view, from sources like Foley & Van Valin (1984) and Dowty (1991), that thematic role classifications are not only coarse-grained, but they are to a degree context-dependent. My assumptions about this are repeated in (17).

(17) Thematic roles and macroroles:

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 (and OoK) are undergoers of an event denoted by C.

f. Sp and SoK are initiators of an event denoted by C.

g. The macroroles of the controller and the controllee must be the same.

The difference is that there is a covert agent with 'tell-PASS' but not with 'seem'. This covert agent qualifies as an initiator by (17a), which can control Sp by (17f,g). It also prevents the goal argument from being an initiator by (17c),[8] but facilitates it being an undergoer by (17d). As a result, the goal argument of 'tell-PASS' can control Ad but not Sp by (17e,g) (see (13), (15)). In contrast, 'seem' and 'remember' take an experiencer/goal, but no agent, overt or covert. Thus, their goal argument does count as an initiator ((17c)), so it can control Sp and be the antecedent of i-shift, as in (14). However, it cannot count as an undergoer in the absence of an agent ((17d)). Therefore, the dative arguments of these predicates cannot control Ad and thus antecede u-

[8] There seems to be some lower-level crosslinguistic variation on this point, though. A passive agent does not necessarily block a goal argument from counting as an initiator in Lubukusu and Japanese, although it does in Kinande and Magahi. Whether this is related to any other observable properties of the passive constructions in these languages is a topic for future research.

shift, as shown in (18).

(18) Magahi (fieldwork Deepak Alok)

a. #Santee-aa-ke laga h-ai ki Ram tor beijati kar-l-ai.
Santee-FM-DAT seem be-3.NH.S that Ram you.GEN insult do-PFV-3.NH.S
"It seems to Santee$_i$ that Ram insulted you$_{*i,ad*}$."

b. #Ram-ke yaad ha-l-ai ki tu almira-me paisa chhupai-l-eN ha-l.
Ram-DAT memory be-PFV-3.NH.S that you.NH drawer-in money hide-PFV-2.NH.S be-PFV
"Ram remembered that he/you$_{*i,ad*}$ hid the money in the drawer."

This shows clearly the context sensitivity in how "intermediate" thematic roles like goal/experiencer relate to the initiator/undergoer distinction. They have the same basic first-order thematic role across this range of examples, as shown by the fact that they are all marked by the dative postposition *-ke*. I would add that this in turn shows that they are all generated in the same syntactic position, specifier of VP; see Baker (2015, 2024). But these arguments count as initiators if and only if there is no other better initiator (an agent) in the structure, and as possible undergoers only if there is an agent in the structure. This context sensitivity can also be seen by comparing 'remember' with its transitive counterpart 'remind'. The two predicates are morphologically as well as semantically related in Magahi: 'remember' is the result of using the noun *yaad* 'memory' in a light verb construction with the intransitive verb 'be/become', whereas 'remind' is the result of using *yaad* with the (di)transitive light verb 'give'. The transitive version has an agentive subject, the reminder. In this case, the rememberer can control Ad and thus u-shift as in (19), a minimal contrast with (18).

(19) Magahi (fieldwork Deepak Alok)

Santee-aa Ram-ke yaad dia-lk-ai ki tu almira-me paisa chhupai-l-eN ha-l.
Santee-FM Ram-DAT memory give-PFV-3.NH.S that you.NH drawer-in money hide-PFV-2.NH.S be-PFV
"Santee$_i$ reminded Ram$_k$ that he/you$_{k,ad*}$ hid the money in the drawer."

This shows again that an experiencer/goal can be an undergoer if and only if there is a better initiator in the structure, such as an agent.

Next we can consider indexical shift when the matrix verb is 'hear' in the light of these themes. 'Hear' is not a passive verb, and its experiencer argument is not marked dative, but it is thematically similar to 'tell-PASS' in that it has an experiencer/goal as the subject and an agent/source argument is present as an oblique phrase not in Spec TP or not at all. In the African languages, 'hear' constructions were interesting in that the hearer subject could control C-agreement via SoK when by itself, but in a subset of the languages this is blocked when a source phrase is present. This result may have seemed a bit quirky and idiosyncratic. But Magahi turns out to be strikingly parallel. The experiencer/goal of 'hear' can control i-shift if there is no source phrase present as shown in (20). This is different from 'was told', which is similar in semantic content, but the tellee controls u-shift not i-shift, as shown in (13), (15) and (16) above. An NP with an experiencer/goal role can be considered an initiator in the absence of an agent argument ((17c)), and initiators are qualified to control Sp ((17f, g)). Passive verbs have agents, covertly or overtly, which prevent this, but 'hear' (like 'seem') need not have one.

(20) Magahi (fieldwork, Deepak Alok)

a. Jaun-waa sunl-ain ki hamar bahinii await h-ai.
John-FM heard.3.NH.S-HH.AL that my sister come.PROG be-3.NH.S
"John$_{i}$ heard that his/my$_{i,sp*}$ sister came." (said to a teacher)

b. Santee-aa sunk-ai ki ham parichhaa paas ho ge-l-i.
Santee-FM heard-3.NH.S that I exam pass become go-PFV-1.S
"Santee$_{i}$ heard that he/I$_{i,sp*}$ passed the exam."

Indeed, in Magahi the hearer can control i-shift even when a source phrase is present, as in (21). I claim that this is because a source phrase does not necessarily count as an agent, preventing a goal argument from being categorized as the initiator by (17c).[9]

[9] We might also go on to infer from this that the source phrase in Magahi can count as either a PP-adjunct, as in Lubukusu, or as an oblique NP argument of the verb, as in Kinande. The particle *se* would be ambiguous as to whether it is an

(21) Magahi (fieldwork, Deepak Alok)
Santee-aa Bantee-se sunk-ai ki ham parichhaa paas ho gel-i.
Santee-FM Bantee-INS heard-3.NH.S that I exam pass become go.PFV-1.S
"Santee$_i$ heard from Bantee that he/I$_{i,sp*}$ passed the exam."

However, when a source phrase is present along with the verb 'hear', Magahi also allows another possibility. The source phrase is enough like an agent that it can count as the initiator of the event, in which case it controls Sp and antecedes i-shift. When this happens, the hearer can control Ad, thereby anteceding u-shift, as seen in (22).

(22) Magahi (fieldwork, Deepak Alok)
Santee-aa Bantee-aa-se sunl-ai ki (tu) hamraa bajaar-me dekh-l-eN.
Santee-FM Bantee-FM-INS heard-3.NH.S that (you) me.ACC market-in see-PFV-2.NH.S
"Santee$_i$ heard from Bantee$_k$ that he/you$_{i,ad*}$ saw him/me$_{k,sp*}$ in the market."

I take this to be compatible with (17d): initiator-sources are enough like full agents to allow a goal argument to count as an undergoer, which allows it to control Ad. However, the experiencer/goal cannot control Ad if there is no source, as in (23) (contrast with (20a) above).

(23) Magahi (fieldwork Deepak Alok)
Jaun-waa sunl-ai ki tor bahinii awa-it h-au.
John-FM heard-3.NH.S that your sister come-PROG be.3.S-NH.AL
"John$_i$ heard that your$_{*i,(ad*)}$ sister came."

This illustrates again the context sensitivity of how less canonical agent and patient arguments are handled in the Foley-Van Valin-Dowty-type systems of macro/proto-roles that I have adopted. Experiencer-goal arguments are on the borderline of this core distinction of thematic theory, so they can count as either agent-like arguments that control Sp or as patient-like arguments that control Ad,

---

adposition or a case marker, as is common.

depending on what other arguments are expressed. (22) also shows again that a nominal need not trigger agreement on T in order to control Sp and thus be the antecedent of a shifted ‘me’ in Magahi.

One can also compare constructions with ‘hear’ to constructions with the verb ‘ask’ in Magahi. Like ‘tell’ and ‘hear’, ‘ask’ can take two nominal arguments as well as a CP complement. In terms of the case-marking of its arguments, ‘ask’ looks very much like ‘hear’: it has a nominative subject and its internal argument, the askee, is marked with the postposition *se*. I suppose that this is because askee is a (potential) source, in that the asker is hoping to get information **from** the askee. Despite this, the overall thematic structure of ‘ask’ is more like that of ‘tell’ than like that of ‘hear’, since ‘ask’ has an agentive subject and since the question content is directed from that subject **to** the oblique internal argument. Therefore, the internal argument of ‘ask’ is thematically a goal as well as a (potential) source, whereas the internal argument of ‘hear’ is a pure source. This layered thematic analysis matters for how control of Sp and Ad proceeds. In (24), ‘ask’ behaves like ‘tell’ rather than like ‘hear’ in that its subject is a true agent and thus can be the initiator by (17a) and can control i-shift, even in the presence of the oblique internal argument. At the same time, the oblique object of ‘ask’ is a goal as well as a (potential) source. It is also in the context of an agent, so it can be an undergoer by (17d) and thus can control Ad and u-shift in (24).

(24) Magahi (fieldwork Deepak Alok)
Raam profesar saaheb-se puchhk-au ki kaa ham apne-ke dekhl-i-ain he.
Ram professor HH-INS ask-NH.AL that Q I you.HH-ACC saw-1.S-HH.AL be
“Ram$_i$ asked the professor$_k$ whether he/I$_i$ saw him/you$_k$.”

This comparison of ‘ask’ with ‘tell’ and ‘hear’ confirms that a thematic analysis of the matrix verb/event is crucial; one cannot just operate off the clause’s surface case pattern.

As a final case of thematic role matching, consider morphological causatives in Magahi. These provide another situation in which there can be mismatches between thematic roles like agent and grammatical functions like subject. In particular, the causee of a syntactic causative construction can count as an agent of sorts, even though it does not occupy the Spec TP position. In Kinande and Ikalanga, we saw some

evidence that the causee can control SoK, even though it cannot license C-agreement with SoK given the T/Agree Condition. With Magahi indexical shift, the evidence for this is clearer. Consider for example (25), a causative built on the triadic verb 'tell'. This has a reading in which 'I' refers to the causee Chhotu and 'you' refers to Ram the goal of the telling. This is expected if 'tell' counts as a separate verb from causative 'make'. Then the agent of 'tell' controls Sp and its goal controls Ad, in accordance with (17). The causee 'Chhotu' not being agreed with and not being in Spec TP does not prevent it from controlling Sp, and hence being the antecedent for i-shift. A structure for (25) is given in (26).

(25) Magahi (fieldwork, Deepak Alok)
Bittu-aa Chhotu-aa-se Ram-ke kah-wal-k-ai ki (pro) toraa dekhl-i.
Bittu-FM Chhotu-FM-INS Ram-DAT tell-CAUS-PFV-3.NH.S that (I) you.ACC saw-1.S
"Bittu$_n$ made Chhotu$_i$ say to Ram$_k$ that he/I$_{i\ (n,sp^*)}$ saw him/you$_{k\ (k,ad^*)}$."

(26) [Bittu$_n$ T [t$_n$ Voice [make [Chhotu$_i$ Voice [tell Ram$_k$ [Sp$_i$ Ad$_k$ that [I$_i$ saw you$_k$]]]]]]]

Similarly, (27) has a causative of the dyadic verb 'think'. Here too the causee can control Sp and hence i-shift, supporting the claim that the controller of Sp is thematically determined rather than structurally determined, to the extent that the two are different.[10]

[10] It is also possible for the causer to be the antecedent of i-shift in both (25) and (27). On my account, this must be because these verbs can also be analyzed as lexical causatives. Under this analysis, 'make-say' and 'make-think' are triadic verbs, not significantly different from 'tell'. As such, the (morphogically complex) verb that selects the CP complement also has an agent argument that controls Sp inside that complement, compatible with the GOCS. I do not know if Magahi has causatives that cannot behave like simple triadic verbs or not.

(27) Magahi (fieldwork, Deepak Alok)

Bittu-aa Chhotu-aa-se/ke soch-wal-k-ai ki (pro) Ram-ke dekhl-i.
Bittu-FM Chhotu-FM-INS/DAT think-CAUS-PFV-3.NH.S that (I) Ram.ACC saw-1.S
"Bittu$_n$ made Chhotu$_i$ think that he/I$_{i,(n,sp*)}$ saw Ram."

The causee in these examples is also like the *by*-phrase in a passive and the *from*-phrase associated with 'hear' in that it can control Sp without triggering subject agreement on the finite verb.

This completes my discussion of the thematic role matching issues that are involved in the control component of indexical shift. We have seen that initiator (proto-agent) arguments are the controllers of i-shift and undergoer (proto-patient) arguments are the controllers of u-shift. How experiencer-goal arguments fit into this is more complex and varies from example to example. However, the variation is patterned, and involves the kinds of factors that have been previously identified as affecting thematic role classification for other phenomena, including choices about which NPs will be projected as underlying subjects and objects (Foley & Van Valin 1984, Dowty 1991).

As a comparative note, I acknowledge that these complexities about which argument of the matrix verb counts as the "author" or "logophoric center" for verbs like 'hear', 'seem', and 'remind' have been discussed some in the semantic literature on indexical shift (and logophoricity). Usually, these discussions are framed in terms of which argument of the matrix verb counts as being an attitude holder, rather than which argument is an initiator or proto-agent. Some readers may feel that talk about attitude holders is more to the point, and my formulation in terms of initiators and undergoers is no more than a somewhat clumsy approximation to this. But I see two potential advantages to formulating the relevant generalizations in terms of macroroles. The first is that it may capture more clearly the fact that only one argument in any given clause can be (say) the controller of Sp and antecedent of i-shift, even though which argument it is varies across examples. It is a feature of the Foley/Van Valin/Dowty approach that there is only one initiator/proto-agent per clause, much like there is only one subject per clause.[11] In contrast, if an attitude

[11] This is clearer for Foley & Van Valin's version than it is for Dowty's. Dowty (1991: 576, 599) mentions that some arguments may share the same proto-role in his system, distinguishing his view from Foley and Van Valin's in this regard.

holder is roughly “the person who mentally represents the content described by the CP complement”, then it seems to me that in many examples both arguments qualify as possible attitude holders, including both arguments of ‘tell’, ‘hear’ and ‘remind/remember’ even though they behave differently.[12] It is useful to have thematic notions of agent, source, causer, and goal to help sort this out. The other potential advantage of the implementation in terms of macroroles is that is part of a much broader system. By using it, I suggest that choices and tradeoffs as to what counts as the initiator or undergoer for purposes of controlling Sp and Ad are conceptually similar to the well-known choices and tradeoffs as to what counts as the patient or location for ‘spray’/’load’ verbs or what counts as the agent or theme for psych verbs like ‘like’ or ‘please’—the kinds of data that originally motivated Foley & Van Valin’s and Dowty’s proposals. Whether it is right to see variation as to which argument of a verb can be the antecedent for shifted ‘I’ as an instance of the same phenomenon is too big a question to consider more directly here. But it is something important to consider in the big picture, I claim.

3.3. Structural conditions on the controller of Sp and Ad

Next let us turn to the evidence that indexical shift in Magahi is shaped by the GOCS, the most fundamental syntactic condition on obligatory control. This puts conditions both on where an NP must be in order to obligatorily control something and on where a CP must be to have null DPs at its edge undergo obligatory control.

Consider first the restrictions on what can be a controller. The GOCS states that an NP can be an OC controller of a controllable null DP if it is an argument of the verb that heads the phrase that contains the clause that the null DP is at the edge of. One core consequence of this is that a characteristic sort of clause-level locality holds of obligatory

---

However, I do not see that he presents any compelling cases in which two nominals in the same clause are proto-agents to the same degree, even if he allows for this in principle.

[12] Alternatively, one could characterize an author/attitude holder more narrowly, as the one who is the source of the content expressed by the complement CP. Then it is clear why the goal of ‘tell’ cannot be the antecedent for i-shift, but not why the subject of ‘hear’ can be.

control relationships. Consider an abstract structure like (28), where there are two levels of clausal embedding. Here the GOCS implies that Sp2 and Ad2 can be controlled by Z and W, arguments of the next higher clause, but not by X and Y, arguments of the highest clause. Upward C-agreement in the African languages shows this kind of clause-level locality, as do cases of the obligatory control of PRO in languages like English.

(28) X told Y [Sp1 Ad1 that [Z told W [Sp2 Ad2 that [I saw you]]]].

However, we have to be careful about how we look for this effect in Magahi, because indexical shift is in some sense optional in this language. As a result, an example like (29) is possible with 'I' in the lowest clause shifting to refer to the higher subject 'Santee' rather than the intermediate subject 'Bantee'.

(29) Magahi (fieldwork, Deepak Alok)
Santee-aa kahl-ai ki Bantee-aa socha h-ai ki (ham) tej h-i.
Santee-FM said-3.NH.S that Bantee-FM think be-3.NH.S that (I) smart be-1.S
"Santee$_i$ said that Bantee$_k$ thinks that he/I$_{i,k,sp*}$ am smart."

My analysis of this with many precedents is that this is not the result of Sp2 in the lower CP being controlled directly by the matrix subject 'Santee' past subject 'Bantee'. Rather, it is the result of 'I' in the lowest clause being bound at a distance by the controlled Sp1 of the higher CP—perhaps via Sp2 being bound by Sp1. The representation of the relevant reading of (29) is (30b), not (30a).

(30) a. Sp*$_n$ **Santee$_i$** said that [Sp1$_n$ Ad1 that [Bantee$_k$ think [**Sp2$_i$** Ad2 [**I$_i$** am smart]]]

b. Sp*$_n$ **Santee$_i$** said that [**Sp1$_i$** Ad1 that [Bantee$_k$ think [(Sp2$_i$) Ad2 [**I$_i$** am smart]]]

We can confirm that this is true by placing a first person indexical in the middle clause and seeing how that relates to the interpretation of an indexical in the lowest clause. Suppose that (30a) were a possible representation for (29). Then it should still be possible for 'I' in the lowest clause to refer to Santee even if there is an unshifted indexical

in the middle clause, because such an 'I'/'me' in the middle clause would not be bound by the controlled Sp, $Sp2_i$. But this prediction is false. (31) shows that if 'I' or 'my' in the middle clause refers to the speaker Sp*, then 'I' in the lowest clause cannot be coreferential with the highest subject *Santee*. If however, 'I'/'my' in the middle clause does shift to refer to Santee, then 'I' in the lowest clause can too. This is what we expect if (30b) is a possible representation and (30a) is not. (Note that a significant assumption here is that any first person pronoun must be bound by the closest c-commanding Sp element, as required by the Person Licensing Condition, introduced in §3.4.5. For more on this, see §4.4 below.)

(31) Magahi (fieldwork, Deepak Alok)

a. Santeeaa kahl-ai ki (ham) socha h-i ki (pro) toraa bajaar-me dekhl-i.
Santee-FM said-3.NH.S that I think be-1.S that (I) you.ACC market-in saw-1.S
"Santee$_i$ said that I$_{sp*}$ think that he/I$_{*i,sp*}$ saw you$_{ad*}$ in the market." Or "Santee$_i$ said that he/I$_i$ think that he/I$_{i,*sp*}$ saw you$_{ad*}$ in the market."

b. Santee-aa socha h-ai ki Bantee-aa hamar baabaa-ke kahk-ai ki ham igjaam-me phel ho ge-l-i.
Santee-FM think be-3.NH.S that Bantee-FM my.GEN grandfather-DAT told-3.NH.S that I exam-in fail happen go-PFV-1.S
"Santee$_i$ thinks that Bantee told my$_{sp*}$ grandfather that I$_{*i,sp*}$ failed the exam." Or "Santee$_i$ thinks that Bantee told my$_i$ grandfather that I$_{i,*sp*}$ failed the exam."

These examples demonstrate the local control of Sp, the ghostly DP most parallel to SoK as found in the African languages. The same kind of reasoning applies to Ad, as I showed in §3.4.3 using data from allocutive marking. It is also possible to replicate the result using u-shift. The critical example is (32).

(32) Magahi (fieldwork, Deepak Alok)

Santee-aa Bantee-aa-ke kahk-ai ki tu Ram-ke khal-eN ki (tu) parichaa paas ho gel-eN.

Santee-FM Bantee-FM-DAT told-3.NH.S that you Ram-DAT told-2.NH.S that you exam pass become go-2.NH.S

"Santee told Bantee$_k$ that you$_k$ told Ram that you$_{k,*ad*}$ passed the exam." Or "Santee told Bantee$_k$ that you$_{ad*}$ told Ram that you$_{ad*,*k}$ passed the exam."

If 'Bantee' could control Ad2 directly, at a distance, then it should be possible for 'you' in the lowest clause (the exam-passer) to refer to Bantee even when 'you' in the middle clause (the one who spoke to Ram) refers to Ad*. In contrast, if control of Ad shows clause-level locality in accordance with the GOCS, then 'you' in the lowest clause should only be able to refer to Bantee if 'you' in the middle clause also refers to Bantee. The second prediction is the correct one.[13]

This line of reasoning does not depend on the details of the structure of the middle clause. All that matters is that the clause that immediately contains the controlled Sp and Ad is not in the VP headed by the verb that the putative long-distance controller is an argument of. In particular, it should not matter whether the intermediate clause is a full-fledged CP or not, as an account in terms of the Phase Impenetrability Condition might. Nor should it matter whether the intermediate clause hosts Sp and Ad coordinates of its own, as an account in terms of Relativized Minimality might. In this light, consider (33), where the highest verb 'expect' takes an infinitival/nominal complement, rather than a full finite CP, and the subject of the intermediate verb 'say' is an oblique nominal, not a nominative subject. There is no full CP structure associated with the nonfinite clause built around 'say'; rather it is like an English gerund construction. Nevertheless, the thematic subject 'Bantee' of this nonfinite clause can control the Sp associated with the finite clause built around 'pass the exam', resulting in shift of 'I' to refer to Bantee. In contrast, the highest subject 'Santee' cannot control the Sp of the most embedded clause, because the CP 'that I passed the exam' is not the complement of the verb 'expect'. Therefore, 'I' in the most

---

[13] These predictions are true but not different from those of the shifty operator view,. This effect is sometimes called "no intervening binder" (Anand 2006) or "local determination" (Deal 2020). See Deal (2020: 39-42) for discussion.

embedded clause cannot refer to Santee. This shows the locality of the control of Sp more directly. (Note that the finite CP complement of 'say' extraposes rightward here, as usual, but the nonfinite complement of 'expect' does not. This nonfinite complement is evidently not a bounding category for the Right Roof Constraint.)

(33) Magahi (fieldwork, Deepak Alok)
Santee-aa [Bantee-aa-se t_{CP} kah-e-ke] ummid kar h-ai] [ki ham parichha paas ho ge-l-i].
Santee-FM Bantee-FM-INS say-INF-GEN expect do be-3.NH.S that I exam pass happen go-PFV-1.S
"Santee$_i$ expects Bantee$_k$ to say that I$_{k,*i,sp*}$ passed the exam."

The other major consequence that the GOCS has for the controllers of Sp and Ad is that they must be arguments of the immediately superordinate verb, not some other constituent of the clause.[14] For example, the subject of the matrix verb can control Sp and thus antecede i-shift in an example like (34a), but the possessor of the subject cannot ((34b)), nor can the subject of a relative clause that modifies the subject ((34c)). This seems to be a syntactic restriction on indexical shift, and one that is reminiscent of obligatory control.[15]

(34) Magahi (fieldwork, Deepak Alok, Alok 2020: 176 (83))
a. Santee-aa kaha h-ai ki ham jaldiye mil-e aibo.
Santee-FM say be-3.NH.S that I soon meet-INF come.FUT.1.S
"Santee$_i$ said that he/I$_{i\ (sp*)}$ will come soon."

---

[14] Some of this discussion comes from Alok and Baker (2022).

[15] The examples in (34b,c) are a bit awkward in Magahi even with 'I' referring to Sp*, in that Magahi does not like to have inanimate subjects. Nevertheless, the examples are much worse with 'I' referring to Santee. Alok (2020: 176 fn. 16) reports that in a sentence like 'Santee's face told Bantee that I passed the exam', 'I' can refer to Santee, the inalienable possessor of the subject. I assume that this is a case of metonymy, where Santee's face is very closely associated with Santee and can be used as a way of referring to him. If so, then 'Santee's face' and 'I' are actually coreferential. See Chapter 5 for some discussion of metonymy in the context of logophoric pronouns.

b. #Santee-aa-ke likhkal chiThii-aa kaha h-ai ki ham jaldiye mil-e aibo
Santee-FM-GEN written letter-FM say be-3.NH.S that I soon meet-INF come.FUT.1.S
"Santee$_i$'s letter said that I$_{*i\ (sp*)}$ will come soon."

c. #ChiThii-aa je Santee-aa likh-k-ai kaha h-ai ki ham jaldiye mil-e aibo.
letter-FM REL Santee-FM write-PFV-3.NH.S say be-3.NH.S that I soon meet-INF come.FUT.1.S
"The letter that Santee$_i$ wrote said that I$_{*i\ (sp*)}$ will come soon."

Similarly, (35) gives tentative evidence that the goal argument can control Ad and hence antecede u-shift, but the possessor of the goal argument cannot. Having 'his' in the expression 'his phone' be the antecedent of u-shift ((35b)) is ungrammatical compared to having 'him' as the goal argument be the antecedent of u-shift ((35a)).[16]

(35) Magahi (fieldwork, Deepak Alok)
a. Santee-aa okraa text bhej di-au ki (tu) parichaa paas ho gel-eN.
Santee-FM 3SG.DAT text send give-NH.AL that you exam pass become went-2.NH.S
"Santee texted to him$_i$ that he/you$_{i\ (ad*)}$ passed the exam."

b.#Santee-aa okar phon-maa-ke text bhej di-au ki (tu) parichaa paas ho gel-eN.
Santee-FM him.GEN phone-FM-DAT text send give-NH.AL that you exam pass become went-2.NH.S
"Santee send a text to his$_i$ phone that you$_{*i\ (ad*)}$ passed the exam."

This also follows from the GOCS.

This point deserves some emphasis, because it is a reasonably clear difference between my control-based theory and the shifty operator account. In purely semantic terms, it seems reasonable to say that Santee counts as the author in the context associated with the matrix event of saying in all three examples in (34); he certainly is the source

[16] These examples use a pronoun for the goal/possessor rather than a name like *Bantee* because the pronoun's dative and genitive case forms are distinct.

of the content expressible as “I will come to visit soon” in all three. Therefore, the shifty operator account might well predict that ‘I’ referring to Santee will be possible in all three examples. Indeed, some constructions that are perspectival or “logophoric” in a broad sense do show this kind of grammatical laxity. For example NOC PRO in English is possible in (36b,c), where the antecedent of PRO is not an argument of the matrix clause, as well as in (36a), where the antecedent of PRO is an argument of the matrix clause.

(36) English (Landau 2001: 110, personal knowledge)
  a. It damaged Lisa$_i$ [PRO$_i$ to perjure herself].
  b. It damaged Lisa$_i$’s reputation [PRO$_i$ to perjure herself].
  c. It damaged the reputation Lisa$_i$ built up over the years [PRO$_i$ to perjure herself].

This is different from indexical shift in Magahi, which does not have this kind of latitude, being possible in (34a) but not in (34b) or (34c). There is thus an additional constraint on indexical shift: not only does the understood antecedent of ‘I’ in the embedded clause need to count as an author of some content semantically, but it needs to be a grammatical argument of the matrix verb. By attributing this syntactic restriction to obligatory control, I am saying that the indexical shift paradigm in (34) and (35) in Magahi is akin to the OC paradigm in (37) and different from the NOC paradigm in (36); (37b,c) can only mean that the letter is promising that the letter itself will come soon.

(37) English (personal knowledge)
  a. Mary$_i$ promised [PRO$_i$ to come soon].
  b. #[Mary$_i$’s letter]$_k$ promised [PRO$_{k,*i}$ to come soon].
  c. #[The letter [Mary$_i$ sent]]$_k$ promised [PRO$_{k,*i}$ to come soon].

3.4. Structural conditions on the clause that contains Sp and Ad

The GOCS also has implications for where a clause containing shifted indexical pronouns can be. It says that obligatory control is something that happens specifically with null DPs contained in clauses that are generated inside the projection of the lexical head which CP is an argument or modifier of—usually a verb, but it can also be a noun or an adjective. In other words, obligatory control is something that

happens in complements and low adjuncts. So far, we have only considered CP complements. Now let us consider the possibility of indexical shift in other types of CPs.

Consider first relative clauses, which are canonically adjoined somewhere inside an NP/DP projection, and thus are not immediate constituents of the verb phrase. These are indeed an environment in which upward C-agreement with the matrix subject has not been attested in the literature and is not possible in Ibibio. In Magahi, relative clauses can have unshifted allocutive agreement, reflecting the social status of the addressee of the sentence as a whole, as in (38).

(38) Magahi (Alok 2020: 11 (16))
*[Laikwaa [je uhan khaRaa h-**au**]] hamar bhaai h-ai.*
boy REL there stand be.3.NH.S-NH.AL my brother be-3.NH.S
"The boy who is standing there is my brother." (to a peer)

This shows that the CP that constitutes the relative clause can contain Ad; it is a full FinP (and more), not some kind of truncated clause that does not have room for such an element. As such, the relative clause presumably has room for Sp as well, given that Sp and Ad are both arguments of the Fin head(s) in Magahi. However, Sp and Ad cannot be controlled by arguments of the matrix verb in this environment. For Sp, this is shown by the comparison in (39). 'Me' can shift to refer to the subject of a verb like 'imagine' when it is in a CP complement, as in (39a), but not when it is in a CP relative clause that modifies a DP complement, as shown in (39b).

(39) Magahi (fieldwork, Deepak Alok)
a. Santee kalpanaa kark-ai ki ego sudar laiki hamraa-se biaah kart-ai.
Santee imagine did-3.NH.S that one.CL beautiful girl me-INS marriage do.FUT-3.NH.S
"Santee$_i$ imagines that a beautiful girl will marry him/me$_{i\ (sp*)}$."

b. Santee, ego sudar laiki je hamraa-se biaah kart-ai, okra baare-me kalpanaa kark-ai.
Santee one.CL beautiful girl REL me-INS marriage do.FUT-3.NH.S her about-LOC imagine did-3.NH.S
"Santee$_i$ imagined (about) a beautiful girl who'll marry me$_{*i,sp*}$."

(40) gives two other examples illustrating the impossibility of i-shift

in a relative clause in Magahi.[17] (Here the relative clause has been extraposed to postverbal position, as is common in Magahi.[18])

(40) Magahi (fieldwork, Deepak Alok)

a. Saantee-aa kitabi-aa bhulaa del-ai je (ham) kharid-l-i ha-l.
Santee-FM book-FM lost gave-3.NH.S REL I buy-PFV-1.S be-PFV
"Santee$_i$ lost the book that I$_{*i,(sp*)}$ bought."

b. Saantee-aa ego bartan ban-l-ai je ham Bantee-aa-ke de-b-ai.
Santee-FM one pot make-PFV-3.NH.S REL I Bantee-FM-DAT give-FUT-3.NH.S
"Santee$_i$ made a pot that I$_{*i,(sp*)}$ will give to Bantee."

It is worth clarifying that there is no absolute ban on 'I'/'me' inside a relative clause receiving a shifted reading. Such a reading is possible in the more complex sentence in (41). Here 'me' in the relative clause can be coreferential with the highest subject *Santee* or with Sp*, although it cannot be coreferential with the closer subject *Bantee*.

(41) Magahi (fieldwork, Deepak Alok)

Santee-aa kahk-ai ki Bantee-aa ego sudar laiki-ke baare-me sochk-ai je hamraa-se biaah kart-ai.
Santee-FM said-3.NH.S that Bantee-FM one.CL beautiful girl-GEN about-LOC thought-3.NH.S REL me-INS marriage do-3.NH.S
"Santee$_i$ said that Bantee$_k$ imagined (about) a beautiful girl who will marry him/me$_{i,*k,(sp*)}$."

This sentence has two embedded Sps, one in the relative clause 'who

---

[17] Notice that it does not make a difference in Magahi whether the main verb of the sentence is an intensional verb like 'imagine' or 'look for' or a nonintensional verb like 'lose'. I-shift in Magahi is empirically simpler than logophor-licensing in Ibibio in this respect; see §5.2.3 on Ibibio.

[18] It is possible that these are corelative constructions (cf. Srivastav 1991) rather than simple extraposed relative clauses. If so, then I assume they are basically a kind of high adjunct clause, with a special interpretation. Such adjuncts are also not a domain of OC, hence not a context for indexical shift, as discussed below.

would marry me' and one in the complement of 'say'. The one in the relative clause cannot be obligatorily controlled by *Bantee* (or anything else), in line with the GOCS. However, Sp in the complement of 'say' can be controlled by *Santee*. That Sp is then the closest [+1] binder for Sp in the relative clause and an ultimate binder for 'me' in the relative clause. Therefore 'me' can refer to Santee, but not to Bantee. The structure is roughly as in (42).

(42) Sp^{*}_{i} $Santee_{k}$ said [$Sp1_{k,(i)}$ that $Bantee_{n}$ imagined
[a beautiful girl [$Sp_{k,*n,(i)}$ who would marry $me_{k,*n,(i)}$]]]

Similarly, allocutive marking shows that Ad can be inside a relative clause but it cannot be controlled by a goal argument of the matrix verb in a sentence like "Santee told Grandfather the news that Bantee told him"; see Chapter 3, example (56).[19]

Consider next adjunct clauses. The expectation that comes from the GOCS is that there should be two kinds of behavior: high adjuncts which are merged into the clause outside the (greater) verb phrase should not show indexical shift, whereas low adjuncts which are merged inside the verb phrase could allow it. Indeed, Magahi has two kinds of CP adjuncts along these lines. One class includes temporal adjuncts, causal adjuncts, and conditional clauses. Like relative clauses, their verbs can bear unshifted allocutive marking, showing that they contain an Ad close enough for the Fin head in the adjunct clause to agree with it. However, they do not allow i-shift, showing that any Sp they contain cannot be controlled by the matrix subject.

(43) Magahi (fieldwork, Deepak Alok)
a. Santee-aa Bantee-aa-ke beijjattii karl-ai kaaheki Bantee-aa pahile hamar beijjattii karl-ai hal.
Santee-FM Bantee-FM-DAT insult did-3.NH.S because Bantee-FM first my.GEN insult did-3.NH.S be-PFV.
"$Santee_{i}$ insulted Bantee because Bantee insulted $me_{*i\ (sp*)}$ first."

---

[19] It should be possible to test this with u-shift as well, but I have not done so. A test example would be something like 'Santee gave Grandfather the note that your friend left for you.' The prediction would be that 'your' and 'you' cannot refer to the grandfather in such a sentence.

b. Jab Santee-aa hamar beijjattii karl-ai ta Bantee-aa okra baRaalii kar-ti ha-l-ai.
when Santee-FM my.GEN insult did-3.NH.S PRT Bantee-FM his.GEN praise do-PROG be-PFV-3.NH.S
"When Santee insulted me$_{*i\ (sp*)}$, Bantee$_i$ was praising him."

In contrast, i-shift is possible in 'so that' adjuncts, introduced by the C-like head *taaki*, as shown in (44).[20]

(44) Magahi (fieldwork, Deepak Alok)

a. Bantee lukaa ge-l-ai taaki hamraa koi na dekh sake.
Bantee hide go-PFV-3.NH.S so.that me.ACC someone not see can
"Bantee$_i$ hid so that no one will see him/me$_i$."

b. Bantee-aa ghare rukl-ai taaki ham bimmar na ho jaa-i.
Bantee-FM home stay-3.NH.S so.that I sick not become go-1.S
"Bantee$_i$ stayed home so that he/I$_{i\ (sp*)}$ would not become sick."

This is not surprising. Syntactically, we can take 'so that' clauses to be generated lower than other adjunct clauses, inside VP. Indeed, given the connection between purposes/goals and agency, it is plausibly right for rationale clauses to be added at the VP level, in the scope of the Voice head which adds the agent. (Again, I do not have independent evidence for the precise attachment sites of these different kinds of adjunct clauses) Semantically, these rationale clauses can have an attitude-like semantics in which they express a goal that is in the mind of the person who performs the action denoted by the matrix clause; they are the adjunct clauses that are most like CP complements in this respect.[21] Moreover, the matrix subject can trigger C-agreement on a rationale clause in Lubukusu and Ibibio, so there is a parallel here between C-agreement and indexical shift in Magahi, as expected if both involve obligatory control of ghostly DP operators.

---

[20] The ordinary C *ki* of complement clauses can also be used to introduce this type of adjunct clause, with no obvious difference in meaning or structure.

[21] Thus *Mary went out into the yard in order to catch a unicorn* doesn't imply that unicorns exist, the way that *Mary caught a unicorn* does. Similarly, *Lois Lane waited for an hour in order to interview Superman* does not imply *Lois Lane waited an hour to interview Clark Kent*.

Although ‘so that’ clauses allow i-shift, it turns out that they do not allow u-shift or shifted allocutive marking—a potentially surprising asymmetry. I discuss this in §4.5.4, in the context of Shift Together phenomena. There I will claim that ‘so-that’ clauses are basically argument-like dependents of an agentive Voice head, slightly revising the view outlined here. See also §8.5.2.

A third syntactic environment to consider is CPs in subject positions. These are also expected not to allow indexical shift where the CP is a true external argument, although they might in cases where the CP is initially merged in a complement position as an internal argument. When it comes to CPs occupying the true syntactic subject position (Spec TP), the issue is moot in Magahi: *ki* clauses cannot end up in this position, perhaps because they are more verbal than nominal in their categorical properties. Thus, (45) is bad with CP in the preverbal subject position unless CP is embedded in a NP/DP with a carrier noun like ‘rumor’ or a demonstrative like *ii* ‘this’.

(45) Magahi (fieldwork, Deepak Alok)
[(Aphawaah) ki Santee-aa inaam jitl-au] sahii ha-l-ai.*
rumor that Santee-FM prize won-NH.AL true be-PFV-3.NH.S
(Not: “That Santee won the prize was true.” OK: “The rumor that Santee won is true.”)

In this respect, Magahi is like many other languages, including Kinande and Ibibio. It is, however, possible to have a CP that is associated with an external argument thematic role appear after the verb, extraposed to the right periphery. For example, in (46) the *ki*-clause that appears sentence finally is the understood subject of the predicate ‘help’. The sentence is grammatical, but unlike clauses that have been extraposed from object position, ‘I’ inside the clause cannot refer to Santee, an argument of the matrix verb.

(46) Magahi (fieldwork, Deepak Alok)
Santee-aa-ke (ii) parasid hobe me madad kark-ai ki ham puruskaar jit-l-i.
Santee-FM-ACC this famous become LOC help did-3.NH.S that I prize win-PFV-1.S
“It helped Santee$_i$ in becoming famous that I$_{*i,sp*}$ won the prize.”

This subject clause is not inside VP, either before or after extraposition. Rather, it starts in Spec VoiceP, above the VP, and

lands right-adjoined to TP. Neither of these structural positions is a context of obligatory control according to the GOCS, so Sp in the periphery of the embedded CP cannot be controlled by *Santee*, the other argument of 'help'. As a result, i-shift cannot happen here.[22] Compare Landau (2001), who shows that complement clauses are domains of OC whereas subjects and clauses extraposed from subject position are contexts of NOC in English and a range of other European languages.

Another construction of interest is CPs that function as the complement of a noun like 'rumor' or 'news'. In the African languages, the head C can agree out of the NP when the complex NP appears in object position. Apparently, then, this structure does allow OC. Indeed, Magahi also allows i-shift in this context. Often this kind of CP is extraposed rightward, such that it does not form a constituent with the noun on the surface, as in (47a). However, it is also possible for the CP and the noun to appear as a unit before the verb (along with *ii* 'this'). In both versions, 'I' in the CP can have shifted reference referring to Santee, the subject of the main verb.

(47) Magahi (fieldwork, Deepak Alok)

a. Santee-aa [aphawaah] suruu kark-ai [ki ham viraasat-me baRimanii paisaa pai-l-i].
Santee-FM rumor start did-3.NH.S that I inheritance-LOC much money get-PFV-1.S
"Santee$_i$ started a rumor that he/I$_{i,(sp*)}$ inherited a lot of money."

b. Santee-aa [ii aphawaah [ki ham puruskaar jiti ge-l-i]] sagaro phailak-ai.
Santee-FM this rumor that I prize win go-PFV-1.S everywhere spread-3.NH.S
"Santee spread everywhere the rumor that he/I$_{i,(sp*)}$ won the prize."

These can be analyzed in a way that is parallel to what I said for the African languages, such that the CP is an argument (or NP-internal

---

[22] Another factor here is that as the object of 'help' *Santee* in (46) might not have the right agent-like thematic role to control Sp in the extraposed CP. However, some examples of LD anaphora in Japanese suggest that the object of 'help', arguably a benefactee, is enough like an experiencer to count as an initiator.

adjunct) of the noun ‘rumor’, Sp inside the CP is OCed by a covert argument of ‘rumor’, and the covert argument is in turn bound by the matrix subject *Santee*. This analysis is consistent with the GOCS.

This N+CP structure is possible in subject positions too; it is another way Magahi has of getting the equivalent of a CP subject (see (45)). Interestingly, this version does not allow i-shift in the complex subject anteceded by an object of the root clause (even though the object here is an experiencer who has the content “I failed the exam” in mind).

(48) Magahi (fieldwork, Deepak Alok)
[Ii batiyaa [ki ham parichhaa-me phel ho ge-l-ai]] Santee-aa-ke gossaa di-laa de-l-ai.
this news that I exam-LOC fail become go-PFV-3.NH.S Santee-FM-DAT anger give-PFV give-PFV-3.NH.S
“The news that $I_{*i\ (sp*)}$ failed the exam made $Santee_i$ angry.”

Here the CP built around ‘fail’ is clearly not generated inside the VP headed by ‘give anger’, so direct OC does not happen. Apparently, it is also not the case that there is a null argument of ‘news’ that can be the OC controller of Sp and can itself be controlled/anteceded by the experiencer ‘Santee’. Perhaps that form of NOC is blocked by the presence of the demonstrative along with the NP subject in (48).[23] We see then that carrier nouns do not have much effect on control and indexical shift in Magahi: indexical shift is possible inside a CP associated with an internal argument position, with or without a carrier noun, and indexical shift is impossible inside a CP associated with a thematic subject position, with or without a carrier noun. This is in line with the fundamental inside-VP/outside-VP distinction built into the GOCS, although there is more to nail down about the possible role of null arguments of nouns in these constructions.

A final place where CPs can occur that deserves some more discussion is as unembedded root clauses. These obviously are not contexts of obligatory control according to the GOCS. They are not merged with the projection of a verb or other lexical predicate; indeed,

---

[23] The conjecture that the demonstrative in (48) may play a role in preventing the psych object *Santee* from anteceding a null argument of ‘news’ is inspired by the fact that logophoric pronouns and LD anaphors are possible in contexts like these in Ibibio and Japanese. See §5.2.3 for discussion.

they are not merged with anything. Sp and Ad can appear in the periphery of a root clause. For Ad, this is shown by the possibility of allocutive marking in root clauses, as discussed in Chapter 3. A more theory-internal reason to say this for both Sp and Ad is that first and second person pronouns are possible in root clauses, and they must be bound by Sp or Ad (Baker 2008; see §4.4 for discussion). The question now is whether Sp and Ad in this context can be controlled by another NP—in this case, another NP in the discourse context. For PROs that are subject to NOC in English, discourse antecedents are possible, as in sequences like *Sam$_i$ was in trouble. [PRO$_i$ Perjuring himself$_i$ before the judge] had been a serious mistake.* But the evidence shows that Sp and Ad in a matrix clause cannot receive a discourse antecedent in this way in Magahi. For example, (49) is not a well-formed discourse in Magahi in which 'me' in the second sentence is interpreted as referring to *Santee*, the subject of the first sentence. If the Sp in the periphery of the second sentence could take *Santee* as its antecedent in discourse, NOC style, this would be possible. Indexical shift is different from exempt anaphors in some languages in this respect; see Baker & Ikawa (2024) and references cited there.[24]

(49) Magahi (fieldwork, Deepak Alok)
#Santee-aa-ke gossaa aayel ho. Bantee-aa hamra beijjatii kar-o he.
Santee-FM-DAT anger come.PFV be-H.AL Bantee-FM my.GEN insult do-H.AL be
"Santee$_i$ was angry. Bantee had insulted me$_{*i,,(sp*)}$."

---

[24] One limited situation in which it looks like a shifted indexical can appear in a root clause is in an example like (i). This is possible if and only if the content of the second sentence is also something that Santee said. Baker & Ikawa (2024) claim that parallel examples using logophoric pronouns in Ibibio are derived by ellipsis. I assume this is true for Magahi too: the second sentence in (i) is syntactically embedded under "Santee said that…" but then it undergoes focus movement and the rest of the sentence is elided. See also §5.2.3.

(i) Magahi (fieldwork, Deepak Alok)
Santee-aa kahl-o ki hamraa gossaa aa-yel h-o. Bantee-aa hamra beijjatii kar-o he.
Santee-FM said-H.AL that me.DAT anger come-PFV be-H.AL Bantee-FM my.GEN insult do-H.AL be
"Santee$_i$ said he/I$_{i, (sp*)}$ was angry. Bantee had insulted him/me$_{i, (sp*)}$."

Nor can Sp in the root clause pick up its reference from a perspectival adverb, as shown in (50). Again, indexical shift is different from exempt anaphors in some languages in this respect (OK is *According to Eric*$_i$, *his children only depend on himself*$_i$ (Charnavel 2020: 685)).

(50) Magahi (fieldwork, Deepak Alok)
Santee-aa-ke anusaar, Sita hamraa pasand kara h-ai.
Santee-FM-DAT according Sita me.ACC like do be-3.NH.S
"According to Santee$_i$, Sita likes me$_{*i,(sp*)}$."

Thus, it is not enough to say that shifted 'I' must refer semantically to some kind of perspectival center, and the subject of an attitude verb is merely a special case of that. Rather, the syntactic structure matters: shifted 'I' can only be in an environment of obligatory control, such as the complement of a verb or a low adjunct clause. Sp and Ad are possible in root clauses, but they cannot undergo some version of discourse-sensitive nonobligatory control in such clauses. Rather, they are fixed as referring to the speaker of the sentence (Sp*) and the addressee of the sentence (Ad*) (see §4.4 for more discussion).

This section has investigated conditions on the control of Sp and Ad by arguments of the superordinate verb in Magahi as revealed by data from indexical shift. I have shown that there are many substantive parallels between the control of Sp in Magahi and the control of SoK which results in upward C-agreement in the African languages. Both language groups are subject to the GOCS and a thematic role matching condition. The similarities appear in complements of 'tell' type verbs, 'think' type verbs, clausal locality effects, causatives, passives, the special properties of 'hear', and purposive adjuncts. Distinctive Magahi constructions that follow the same general principles include dative subject constructions and triadic verbs with oblique objects like 'ask'. Also covered is the fact that neither C-agreement nor indexical shift is possible in relative clauses, high adjunct clauses, sentential subjects, or root clauses. The African languages are not identical with each other in every respect, but the range of variation is small, and the behavior of Magahi falls very well within that range of variation. Parameterized matters are relatively minor: they include whether a source phrase or a passive agent counts as an argument or not, whether a morphological causative is lexical or syntactic or both, and perhaps whether carrier nouns like 'news' have covert arguments or not. (Furthermore, the T/Agree Condition governs the realization of C-agreement with SoK but does not apply to the

control of Sp and Ad in Magahi, for reasons discussed in §3.3.2.)

Now generativists agree that allocutive marking is done by agreement with a null DP in the CP periphery. So is C-agreement in Bantu in the original Diercks/Baker proposal, together with the null DP being anteceded/controlled by a matrix argument. If that is the right theory for allocutive marking and Bantu C-agreement, then the strong similarities between the distribution of these agreement phenomena and indexical shift in Magahi constitute evidence that a control-based theory is right for indexical shift too.

4. Pronoun Binding and the Person Licensing Condition

4.1. Introduction

The basic structure for an indexical shift construction under my analysis is repeated in (51) (=(3)). It consists of three main ingredients: the licensing of Sp and Ad in the periphery of finite clauses, the obligatory control of Sp and Ad by suitable arguments of the matrix verb, and Sp and Ad's binding of pronouns in their c-command domain that match them in features.

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

	HH	+1	+2HH	+1	+2HH
				(*3rd)	(*3rd)
Control		Control	(Agree)	Binding (2x)	

Chapter 3 already discussed to some extent the licensing of Sp and Ad in Spec saP of root clauses in all languages and the licensing of Sp and Ad in Spec FinP of all finite clauses in some languages, including Magahi. §4.3 discussed in detail the control of Sp and Ad by the arguments of a CP-selecting verb like 'tell'; this part of the analysis is closely parallel to the control of SoK and OoK by arguments of the matrix verb that results in upward C-agreement in various African languages, as I emphasized throughout. This section now turns the focus onto the binding relationship that holds between Sp and Ad and pronouns in the TP that is selected by the C-type head that licenses Sp and Ad. This is the newest element in the analysis, the one that is not especially relevant to C-agreement and allocutive marking constructions. That Sp and Ad can bind pronouns inside their domain

could go without saying, since any DP can in principle be a binder. However, there are some special properties and patterns at work in this case, and the time has come to put them under the spotlight.

The special properties come from my claim that participant pronouns have to be bound by the ghostly operators Sp and Ad, that being the source and nature of their first person-ness or second person-ness. Indeed, they must be locally bound, in the sense of being bound by the closest c-commanding Sp or Ad. This requirement is formulated as the Person Licensing Condition (PLC) which I first postulated in Baker (2008: 126) and introduced again at the end of Chapter 3. I restate this slightly as in (52), with the added clauses italicized.

(52) a. A [+1] feature on a pronoun *that does not otherwise have a grammatically assigned semantic value* must licensed by the pronoun being locally bound by the closest c-commanding element that is [+1] (a Sp or another first person pronoun).
b. A [+2] feature on a pronoun *that does not otherwise have a grammatically assigned semantic value* must be licensed by the pronoun being bound by the closest c-commanding element that is [+2] (an Ad or another second person pronoun).

These conditions can be seen is a specific type of relativized minimality (Rizzi 1990), in that a binding relationship between an operator X and a bindee Y cannot be established over another operator Z of the same type as X. As such, the PLC is abstractly like the wh-island condition and similar phenomena. In Chapter 3, the PLC played a supporting role. It helped to answer one relatively narrow question: why does an Ad that is not controlled by an argument of the immediately superordinate verb need to be bound by the next highest Ad—not just any higher Ad. In this chapter, the PLC comes to the fore as an important constraint on indexical shift, and indeed on the use of first and second person pronouns more generally. It will also be crucial in Chapter 6, on so-called indexiphors.

The clauses added to the PLC in (52) clarify that it applies to [+1] and [+2] elements that do not already have a fixed grammatical interpretation of some kind. In Baker (2008), the PLC was applied to ordinary first and second person pronouns, regulating how they are bound by Sp and Ad. In §3.4.5, I argued that it also applies to instances of (Sp and) Ad that do not undergo obligatory control, explaining the behavior of uncontrolled Ad in (say) adjunct clauses. It is not surprising that the same principle should apply to both elements,

since both are [+2] and pronominal (nominals with no encyclopedic meaning). However, Ad* and Sp* in root clauses are not bound by anything, and yet they are grammatical. Similarly, Ad and Sp in complement clauses where they undergo OC in an indexical shift construction are bound by their controllers, which can very well be third person nominals, yet they are also grammatical. The point of the italicized additions in the version of the PLC in (52) is to sort these cases out in a coherent way. I now discuss the subcases one by one.

4.2. The PLC and Sp/Ad in the root clause

Consider first the special status of Sp* and Ad*. My leading idea is that being first or second person is a recursively defined notion. A very small set of linguistic items are intrinsically first or second person in both a formal and a semantic sense. This set might indeed contain only Sp* and Ad*, the arguments of sa1 and sa2, unembeddable functional heads that are found only in root clauses, presumably for semantic reasons (because they denote a speech act; see Portner et al. (2019) for discussion). This is stated in (53).

(53) a. The DP specifier of sa1 denotes the speaker of the speech act expressed by saP and is [+1]. (This DP is called Sp*.)
b. The DP specifier of sa2 denotes the addressee of the speech act expressed by saP and is [+2]. (This DP is called Ad*.)

This provides the basis step for the recursive characterization of first person and second person, whereas the PLC provides the recursive step. Because Sp* and Ad* are assigned an interpretation explicitly by the functional heads that select them, they are not subject to the PLC in (52). This allows them to appear at the top of a root clause, where nothing else c-commands them. I think of Sp* as a syntactically represented version of the so-called "author" coordinate in a Kaplanian context (Kaplan 1989), and Ad* as a syntactically represented version of the "addressee" coordinate (see Spadine (2020) for a similar idea).[25] The advantage of having these coordinates represented as DPs in the syntax, not just as members of a context

[25] However, I realize that there might be semantic subtleties to this claim that I am not fully aware of, so I leave this at an intuitive level. Ideally one would not need both a Kaplanian context tuple and syntactically represented Sp* and Ad*, since that seems like a duplication of effort. But there are worse things.

tuple that exists only in the semantic machinery, is that then they can be targets for Agree in the syntax, as happens with addressee agreement in Magahi and with speaker agreement in Dargwa.

4.3. The PLC and ordinary pronouns

Elements with participant features that do not have an intrinsic interpretation and hence are subject to the PLC include ordinary first and second person pronouns in Magahi and presumably in all languages.[26] Pronouns like 'I', 'me' and 'you' do not have what they refer to fixed by the meaning of the heads they are arguments of, the way that Sp* and Ad* do. Rather, verbs like 'see' and Voice heads like active Voice allow their DP arguments to refer to any individual in the domain of discourse. Nor do verbs, Voice heads, or other standard theta-markers fix the phi-features of their argument(s); those can in general be first person, second person, or third person. Even if the argument of a V or Voice happens to be first or second person, that does not fix its interpretation directly in the current view. My proposal is that [+1] and [+2] are not semantically interpreted features per se, but formal features that signal to the language user which of several operators a given pronoun happens to be bound by. [+1] pronouns are ones that are ultimately bound by Sp* (perhaps by way of other [+1] items, which are themselves bound by Sp*). In such cases, [+1] pronouns end up denoting the speaker of the speech act because that is what their binder denotes. Similarly, [+2] pronouns are ones that are ultimately bound by Ad*, and which therefore end up denoting the addressee of the speech act.[27] Thus a simple sentence

---

[26] The other kind of elements that are [+1] or [+2] are Sp and Ad in Spec FinP as opposed to Spec saP. I discuss these below, claiming that if they are controlled then they do have a grammatically assigned semantic value and are not subject to the PLC, but if they are not controlled then they do not have an assigned semantic value and they are subject to the PLC.

[27] Looking ahead, I am treating [+1] and [+2] in a way that is analogous to how the feature [+log] works in languages with dedicated logophoric pronouns. [+log] is a formal feature, with distinctive exponents at PF in some languages, which has no intrinsic meaning but signals that a particular pronoun is bound by a particular operator. See von Stechow (2003), Anand (2006), Pearson (2013), Park (2018); the approach is sometimes attributed to unpublished work by Irene Heim in 2002.

with participant pronouns like (54a) has the representation in (54b).

(54) Magahi (fieldwork, Deepak Alok)
a. Ham toraa dekha-l-i-(au)
I you.ACC see-PFV-1.S-(NH.AL)
"I saw you."

b. Sp^*_i Ad^*_k C [I_i saw you_k]

This can be thought of as a syntactically expressed version of the Kaplan's idea that first person pronouns have their reference fixed by the author coordinate of the context (here identified with Sp*) and second person pronouns have their reference fixed by the addressee coordinate of the context (equated which Ad*).[28]

Magahi offers some empirical support for the claim that second person pronouns in a root clause must be bound by Ad*, given that it makes Ad* visible by its allocutive agreement. Fin in the matrix CP copies the formal features of Ad* in Magahi by Agree. In addition, the formal features of Ad* and a pronoun bound by Ad* must be

---

[28] The other side of the Kaplanian view is that indexicals are not sensitive to quantification over worlds, hence the strangeness of (ia). This contrasts with (ib), where the referent of definite description 'the speaker' can vary with different worlds. (See Deal (2020: 14-16) for discussion.) (ii) confirms that *ham* 'I' in Magahi is an indexical in this sense too.

(i) a. #Whenever Obama talks, I am tall.
b. Whenever Obama talks, the speaker is tall.

(ii) Magahi (fieldwork, Deepak Alok)
#Jab kabhii Santee-aa bola h-ai, ham baRi lambaa h-i.
When sometimes Santee-FM speak be-3.NH.S I very tall be-1.s
"Whenever $Santee_i$ speaks, $I_{*i\ (sp*)}$ am very tall."

In my terms, this is because Sp* and Ad* are at the very top of the root clause, outside the scope of the quantificational time adjuncts. Therefore, they cannot be bound by the quantifier, and 'I' and 'you' must depend on them directly by the PLC. The immunity of indexicals to this kind of quantification is not evidence that the speaker and addressee coordinates are not present in the syntactic representation on this view, but rather evidence about where they are in the syntactic representation (at the top of it). This is exactly parallel to Deal's (2020: 35-36) assumption that context-shifting operators are very high in the C-domain, higher than the highest possible attachment site of adverbial quantifiers.

compatible, as in other cases of bound variable anaphora. It follows that allocutive marking and second person pronouns in argument positions must match in features in simple Magahi sentences. In particular, they must have the same honorificity features, which are the features other than [+2] itself that are active for Agree in Magahi.[29] This is correct, as shown in (55). The high honorific second person pronoun *apne* cannot be used with nonhonorific allocutive marker *-au* ((55a)), nor can the not-high-honorific second person pronoun *toraa* be used with the high honorific allocutive marker *-ain* ((55b)).[30] As far as Alok and I could tell, this is true even if one imagines complex situations in which there might be different kinds of addressees present simultaneously. For example, a person might have their honored professor and her young child over for tea. One can imagine offering coffee to the young child (in an NH relationship to the speaker) but intending the primary audience of the sentence to be the professor (the person who the speaker wants to impress, and who will probably decide whether the child gets coffee or not). Even in a socially layered situation like this, (55b) is not possible with *toraa*.

(55) Magahi (fieldwork, Deepak Alok)

*a. Toraa/*apne kauphii chah-**au**?*
you.NH.DAT/* you.HH-DAT coffee want.3.NH.S-NH.AL
"Do you want coffee?"

*b. Apne-ke/*toraa kauphii chah-**ain**?*
you.HH-DAT /*you.NH.DAT coffee want.3.NH.S-HH.AL
"Do you want coffee?"

Similarly, two second person pronouns in the same one-clause sentence must match in features and reference because both are bound

---

[29] Plural features figure to be more complex, since they are semantically interpreted and a plural pronoun can be partially bound by a singular DP. I expect, then, that examples like 'you-PL prize won-SG.AL' would be possible where one is talking directly to a single person and saying that a group of people including that person won a prize. However, agreement in Magahi does not copy number features, so we cannot check this prediction in this language.

[30] Here it is important that the second person pronoun is not a nominative subject, since such subjects trigger [+2] subject agreement on T and this is incompatible with allocutive marking (see Alok & Baker (2018) for an analysis).

by the structure's only Ad, Ad*. Thus, one cannot shift addressees internal to a single CP, for syntactic reasons. This is shown in (56).[31]

(56) Magahi (fieldwork, Deepak Alok)
#Ram ***apne****-ke* ***tor*** *kitaab lauTal-ai.*
Ram you.HH-DAT you.NH,GEN book return.PFV-3.NH.S
("Ram returned your book to you." bad even with pointing.)

It is also the case that two first person pronouns in the same clause must match in reference, since they are both bound by the unique Sp*; if they were not bound by Sp*, they could not be first person by the PLC. This however is less striking empirically, since Magahi does not have overt speaker agreement as a window on the features of Sp*, nor are first person pronouns differentiated for honorificity in Magahi. It is also hard to imagine scenarios in which who is the speaker shifts as a single clause is being uttered.[32]

Technically the PLC is satisfied in a slightly different way when the root clause has two instances of 'I/me' or 'you', one of which c-commands the other, as in (57).[33]

(57) Magahi (fieldwork, Deepak Alok)
a. Ram hamraa ***hamar*** *kitaab lauTa-l-ai.*
Ram me.DAT my.GEN book return-PFV-3.NH.S
"Ram returned my book to me."

---

[31] In English, it possible to have two second person pronouns in the same sentence refer to different people if they are accompanied by pointing devices of some kind (e.g. *You on the right side of the room must not interrupt you on the left side of the room*). This does not seem to work in Magahi. I do not know why the languages seem to be different in this way.

[32] A possible case to consider is dream scenarios where there can be two first persons with different semantic values in sentences like Lakoff's famous *I dreamed that I was Brigette Bardot and I kissed me*. Even here, though, 'I' and 'me' in some sense refer to the same person in different guises. I have not tried this in Magahi and do not speculate about how it might fit into my framework.

[33] Sentences of the form 'I/you found my/your book' do not raise this issue, because Magahi has a subject-oriented possessive anaphor (*apan*) that does not vary for phi-features and it must be used in such sentences.

Sp*$_i$ Ad*$_k$ [Ram returned me$_i$ my$_i$ book].

b. Ram ***toraa*** ***tor*** *kitaab lauTal-ai.*
Ram you.NH-DAT you.NH.GEN book return.PFV-3.NH.S
"Ram returned your book to you."
Sp*$_i$ Ad*$_k$ [Ram returned you$_k$ your$_k$ book].

Here the closest c-command condition in the PLC implies that 'my' must be a variable bound by 'me', rather than by Sp*, and that 'your' must be a variable bound by 'you', rather than by Ad*. But since 'me' is itself is bound by Sp*, and 'you' is bound by Ad*, this makes little difference in practice, at least in ordinary situations (without quantifiers or focus markers). Either way, 'my' ends up referring to the speaker and 'your' to the addressee. The PLC does nothing remarkable in such sentences, but it does nothing embarrassing either.

Where the closest c-command condition in the PLC does do crucial work is with first and second pronouns inside an embedded clause. Then there is another kind of [+1] or [+2] element that they could be bound by, other than Sp* and Ad* or another first or second person pronoun. In embedded clauses, they can be bound by Sp and Ad in the Spec FinPs of the embedded clause. Moreover, in Magahi Sp and Ad do not have to depend ultimately on Sp* and Ad*; rather, they can be controlled by matrix clause DPs, which are not necessarily [+1] or [+2]. Here the PLC has work to do, regulating which [+1] or [+2] item a given pronoun will take as its antecedent.

I start with the second person case, where the possibility of allocutive agreement gives us a relatively direct look at what is happening with Ad. Alok & Baker (2018) and Alok (2020) observe that empirically there is a tight relationship between the form of allocutive marking in an embedded clause and whether a second person pronoun in that clause is interpreted as shifted or not. Consider the examples in (58). Here the allocutive marking on the embedded verb 'see' differs in honorification features from the allocutive marking on the main verb 'tell'. This shows that Ad in the embedded clause, the goal of the Agree initiated by Fin, is controlled by the object of 'tell'. As a result, the embedded allocutive marking is nonhonorific *-au* in (58a), matching the social status of Bantee, the goal of the telling event. Similarly, in (58b) the embedded allocutive is HH *-ain*, matching the status of the professor, the goal of the telling event in that example.

(58) Magahi (fieldwork, Deepak Alok)

*a. Santee-aa **Bantee-aa-ke**_ kahk-ain ki Ram **toraa**/*apne-ke dekhl-**au** ha-l.*
Santee-FM Bantee- FM-DAT told-HH.AL that Ram you.NH.ACC/*you.HH-ACC saw.3.NH.S-NH.AL be-PFV
"Santee told Bantee$_k$ that Ram saw him/you$_{k,*Ad*}$." (to a teacher)

*b. Santee-aa <u>profesar saaheb-ke</u> kah-au ki Ram apne-ke/*toraa dekhl-<u>ain</u> ha-l.*
Santee-FM professor HH-DAT told-NH.AL that Ram you.HH-ACC/*you.NH.ACC saw.3.NH.S-HH.AL be-PFV
"Santee told the prof$_k$ that Ram saw him/you$_{k,*Ad*}$." (to a peer)

The representation of (58a) is (59).

(59) Sp* Ad*$_i$ Santee told Bantee$_k$ [Sp Ad$_k$ Fin [Ram saw you$_{k,*i}$]]
Agree

So far this is a Chapter 3 topic. The new Chapter 4 twist is what happens when the embedded clause contains a second person pronoun, as the examples in (58) do. It turns out that the second person pronoun in the embedded clause must get the shifted reading in which it refers to the goal of the matrix verb. For example, in (58a) the second person pronoun in the embedded clause must be the nonhonorific form *toraa*, matching the *-au* suffix on the embedded verb 'see'; it cannot be the high honorific form *apne-ke*, matching the *-ain* suffix on the matrix verb. Along with this, the pronoun must refer to Bantee, and cannot refer to the teacher who is being addressed. In other words, indexical shift is both possible and necessary given this kind of allocutive marking. Similarly, in (58b), the second person pronoun in the embedded clause must be the HH form *apne-ke*, not *toraa*, and it must have the shifted reading in which it refers to the professor, not to the speaker's friend who is being addressed. As expected, the [+2] pronoun 'you' in (59) must be bound by one of the intrinsically [+2] Ads that c-commands it. But more specifically, it must be bound by the closest such Ad—by Ad rather than Ad* in (59). This is precisely what the PLC in (52b) requires.

As a foil to these examples, consider the slightly different example in (60). The overall sentence structure of (60) is similar to that of (58a); here too the goal argument of the matrix verb 'tell' is the nonhonorific Bantee while the addressee of the sentence is a highly honored person, such as a teacher. The key difference is the allocutive marking on the

embedded verb ‘should’: in (60) it is HH, matching the root sentence addressee rather than the matrix goal.

(60) Magahi (fieldwork, Deepak Alok)
*Santee-aa Bantee-aa-ke kahk-**ain** ki Ram-ke **apne-se**/*toraa-se baat karke chah-**ain**.*
Santee-FM Bantee- FM-DAT told-HH.AL that Ram-DAT you.HH-INS/*you.NH-INS talk do.INF should-HH.AL
“Santee told Bantee$_k$ that Ram should talk to you$_{ad*,*k}$.” (to a teacher)

This is the case in which Ad in the embedded clause is not controlled by an argument of the matrix verb; rather it is bound by the Ad (Ad*) in the matrix clause—the other option that is generally available to Ad in a complement clause (see §4.5). So (60) has the same overall structure as (58a), but a different indexing, as in (61).

(61) Sp* Ad*$_i$ Santee told Bantee$_k$ [Sp Ad$_i$ Fin [Ram talk to you$_{i,*k}$]]
Agree

Crucially the behavior of the second person pronoun ‘you’ in the embedded clause is different too. With this kind of allocutive marking, indexical shift of ‘you’ in the embedded clause is impossible: ‘you’ must refer to the addressee of the sentence as a whole in (60), and must accordingly be the HH form *apne-se*, not *toraa-se*. This is also in accordance with the PLC. ‘You’ cannot be bound directly by the matrix goal *Bantee*, because ‘you’ is [+2] and *Bantee* is not. The only c-commanding [+2] elements in (24) are Ad and Ad* which binds Ad. Therefore, ‘you’ has to be bound by Ad, with the effect that there is no indexical shift in this case. The examples in (58)-(61) show that indexical shift and allocutive marking are indeed closely related topics in Magahi, as emphasized by Alok & Baker (2018) and Alok (2020). The PLC in (52) provides the connection, on my account. The control of Ad necessarily results in indexical shift in (58), and that indexical shift is impossible apart from control of Ad, as seen in (60).

Now let us consider first person pronouns, to confirm that they behave similarly. Magahi does not have speaker agreement parallel to allocutive/addressee agreement, which would provide a window on whether Sp is controlled in a particular example. But one can get similar evidence by having two first person pronouns in the same embedded clause. In any such structure, if one of the first person

pronouns has a shifted reading in which it refers to the subject of the matrix verb rather than Sp*, then the second first person pronoun must also refer to the subject of the matrix verb. If there is no c-command relationship between the two pronouns, the result is grammatical but constrained as to what it means, as in (62a). If one of the pronouns c-commands the other inside the embedded clause, then the result is ruled out by Condition B of the Binding theory, as in (62b).

(62) Magahi (fieldwork, Deepak Alok)

a. Santee soch-l-ai ki hamar mammii hamraa dekh-l-ai.
Santee think-PFV-3.NH.S that my.GEN mother me.ACC see-PFV-3.NH.S
"Santee$_i$ thinks that his/my$_i$ mother saw him/me$_{i,*sp*}$." or
'Santee$_i$ thinks that his/my$_{sp*}$ mother saw him/me$_{sp*,*i}$.'

*b. *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$_i$ thinks that he/I$_{i,sp*}$ saw him/me$_{i,sp*}$.")

The structure of (62a) is (63). Suppose the 'my' here refers to Santee. This means that Sp must be controlled by 'Santee' and 'my' takes Sp as its immediate antecedent; otherwise, 'my' would violate the PLC. Then consider the [+1] pronoun 'me' in object position. According to the PLC, it must be bound by the closest c-commanding [+1] element, which is Sp, rather than Sp*. Hence it too must refer to Santee.

(63) Sp*$_k$ Santee$_i$ thinks [Sp$_n$ that [my$_n$ mother saw me$_n$]].
n=i
or n=k

The other possibility is that Sp is bound by Sp* rather than controlled by 'Santee'. In that case, both 'my' and 'me' must refer to Sp*--which is indeed another possible interpretation of (62). What is ruled out is having one of the first person pronouns refer to Santee and the other to Sp*. This is an aspect of the famous Shift Together effect, originally pointed out (for Zazaki) by Anand & Nevins (2004) and Anand (2006), and taken to be strong evidence in favor of the shifty operator theory. In my account, it follows from the PLC.[34] (62b) is similar,

[34] Amy Rose Deal (p.c.) has pressed me with the claim that Shift Together follows

except that here the readings in which 'I' and 'me' are bound by the same antecedent (Sp) happen to be ruled out by Condition B of the binding theory: the pronominal object is coreferential with a c-commanding antecedent in the same domain. Therefore, the structure is unacceptable, forced to violate either the PLC or Condition B. This shows that the PLC is a strong grammatical restriction, not a pragmatic preference which can be overridden by other factors.

Second person pronouns in the same embedded clause must also shift together, as expected given that they are subject to the same condition. This is shown in (64). Again, the two pronouns in (64a) must have the same reference: either both refer to Bantee or both refer to Ad*, with no mixing and matching. And again (64b) is bad, where coreference between the two pronouns runs afoul of Condition B.[35]

---

organically from the shifty operator view, whereas I stipulate it by including a locality condition in (52) that does not hold for other bound pronoun constructions. I push back on this in two ways. First, I'm not convinced that this isn't stipulated in the shifty operator view as well, although it is done so in a more axiomatic fashion. The shifty operator view invokes context overwriting, which says that an indexical can only be interpreted relative to the local context. This is a close analog of the PLC saying that an indexical must be bound by the closest Sp or Ad. Second, I believe that (52) is independently motivated in that it also has relevance to other kinds of items that are [+1] or [+2], including pronouns with complex phi-feature bundles (indexiphors; see Chapter 6) and agreeing heads (Baker 2008). There are also apparent counterexamples to Shift Together in languages like Amharic. The shifty operator view treats these as not being true indexicals but rather as "indexiphors"—logophors that look in some respects like indexicals on the surface. I also adopt a version of this idea in Chapter 6, but it will help that the PLC can be parameterized in ways that the context-shifting view cannot naturally be (as far as I can see). I come back to some further comparison of the two theories below.

[35] There is another kind of Shift Together effect in indexical shift languages, which says that a first person pronoun in an embedded clause shifts if and only if any second person pronouns in the same clause shift. Anand & Nevins (2004) treat this as another manifestation of the same phenomenon that we see in (62) and (64), whereas Deal (2020) distinguishes the two kinds of Shift Together, partly on typological grounds. I discuss the second kind of Shift Together in §4.5.

(64) Magahi (fieldwork, Deepak Alok)

a. Santee-aa Bantee-aa-ke kah-l-ai ki tor mammii toraa dekh-l-ai.
Santee-FM Bantee-FM-DAT tell-PFV-3.NH.S that your.NH mother you.NH.ACC see-PFV-3.NH.S
"Santee told Bantee$_i$ that his/your$_i$ mother saw him/you$_{i,*ad*}$." or
"Santee told Bantee$_i$ that your$_{ad*}$ mother saw you$_{ad*,*i}$."

*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.")

(65) Sp* Ad*$_k$ S told Bantee$_i$ [Sp Ad$_n$ that [your$_n$ mom saw you$_n$]].
n=i
or n=k

From here, we can go on to consider how third person pronouns work in contexts that allow indexical shift of participant pronouns. A question that arises given the PLC (or the shifty operator theory) is how one realizes a structure like (66) in Magahi or other languages with indexical shift. The question is what pronoun is used to refer to Sp*, the speaker of the sentence as a whole, in a clause where [+1] pronouns have shifted to refer to the subject of the matrix clause.

(66) Sp*$_i$ sa [Santee$_k$ think [Sp$_k$ Fin that [I$_k$ saw pronoun$_i$]]]

Here 'pronoun' cannot be a first person form 'me', as we saw above. That sort of pronoun would have to be bound by Sp$_k$ (and indeed by I$_k$) by the PLC—a Shift Together effect. Nor could 'pronoun' be second person, since it is not bound by Ad or any other [+2] element on the intended interpretation. The interesting question is whether it could be a third person pronoun. The answer is potentially yes, if third person pronouns are the elsewhere case, used wherever the more specific first and second person pronouns are unavailable. That seems to be the right answer for Slave, where Rice (1989) reports many examples like (67). Here the first person pronoun in the complement of 'want' refers to the nurse, the subject of 'want' and the controller of Sp, whereas the third person pronoun in the complement of 'want' can be interpreted as referring to the speaker of the sentence as a whole.

Indeed, this is the only way to refer to the speaker from this position.[36]

(67) Slave (Rice 1989: 1274 (7))

a. Judóné ri nurse [Teddy ghọ beghárayuhdá] sudeli?
when Q nurse Teddy about 1SG.S.OPT.see.**3SG.O**]
3SG.S.want.1SG.O
"When does the nurse$_k$ want of me$_{sp*}$ that she$_k$ [lit. I] see me$_{sp*}$ [lit. her] about Teddy?"

b. When Q [Sp*$_i$ C [nurse$_k$ want me$_i$ [Sp$_k$ C [pro[+1]$_k$ see pro$_i$**[*+1]** about Teddy.]]]

Leslau (1995: sec 142.8, 142.11) also has some examples of this kind from Amharic. However, Alok and I have not been able to get this judgement in Magahi. The difference seems to be that indexical shift is obligatory in the complements of certain verbs in Slave, whereas it is always optional in Magahi. If one wants to convey in Magahi what is intended in (66), the natural way to do so is not to use indexical shift, as attempted in (68a), but to say (68b) with no indexical shift. This has the structure in (69) rather than the one in (66).

(68) Magahi (fieldwork, Deepak Alok)

a. #Santee-aa soch-l-ai ki (pro) okraa dekh-l-i.
Santee-FM think-PFV-3.NH.S that (I) him.ACC see-PFV-1.S
"Santee$_i$ thinks that he/I$_i$ saw him$_{k,*Sp*}$."

b. Santee-aa soch-l-ai ki (pro) hamraa dekh-l-ai.
Santee-FM think-PFV-3.NH.S that (he) me.ACC see-PFV-3.NH.S
"Santee$_i$ thinks that he$_i$ saw me$_{sp*}$."

(69) Sp*$_i$ sa [Santee$_k$ think [Sp$_i$ Fin that [he$_k$ saw me$_i$]]]

However, the structure in (69) is not an option in Slave, given the obligatoriness of first person indexical shift with the verb 'want' (see

---

[36] In this particular case, the object of 'see' must refer to Sp*, since 'want' has a proleptic object 'me' and this must be coreferential with some pronoun within the complement of 'want'. In other cases, a third person pronoun within the embedded clause can be ambiguous between referring to Sp* or to some third person identifiable in the discourse.

§4.5). This forces (66) to be used despite the unusual ambiguity in which a third person pronoun can be used to refer to the speaker or some non-speech act participant known from the discourse.[37] I tentatively assume that Magahi's preference for (69) over (66) is a pragmatic one, given that indexical shift is somewhat marked anyway and (69) avoids a certain kind of ambiguity.

The other question about third person pronouns in constructions with indexical shift is whether they can refer to the same matrix nominals that shifted first and second person pronouns can. In Magahi, the answer is no. Apart from indexical shift, third person pronouns can refer to NPs in the matrix clause in the usual way, just as in English ((68b) is also an example of this).

(70) Magahi (fieldwork, Deepak Alok)

a. Santee soch h-ai ki u tej h-ai.
Santee think be-3.NH.S that he intelligent be-3.NH.S
"Santee$_i$ thinks that he$_{i,k}$ is intelligent."

b. Santee-aa Bantee-aa-ke kah-l-ai ki u hamraa/Ram-ke dekh-l-ai.
Santee-FM Bantee-FM-DAT tell-PFV-3.NH.S that he me.ACC/Ram-ACC see-PFV-3.NH.S
"Santee$_i$ told Bantee$_k$ that he$_{i,k}$ saw Ram/me$_{sp*}$."

c. Santee-aa Bantee-aa-ke kah-l-ai ki u okraa dekh-l-ai.
Santee-FM Bantee-FM-DAT tell-PFV-3.NH.S that he him.ACC see-PFV-3.NH.S
"Santee$_i$ told Bantee$_k$ that he$_i$ saw him$_k$."

However, a third person pronoun often loses the ability to refer to a matrix argument when there is a shifted indexical along with it in the

---

[37] If my conjecture is right that the structure in (66) is only clearly manifested in languages with obligatory indexical shift, then Matses would be another language to look for it in. However, Munro et al. (2012) do not discuss this type of example. Uyghur is another possible case. Another very intriguing way that languages with indexical shift can realize the structure in (66) is by using a first person pronoun in the embedded clause that triggers third-person (default) agreement rather than first person agreement on the embedded verb. See Spadine (2020) for clear discussion of such a case in Tigrinya. This sort of "disagreement" construction is discussed in §6.3.

embedded clause. This can be seen in the examples in (71).

(71) Magahi (fieldwork, Deepak Alok)

a. Santee-aa soch-l-ai ki okar maiyaa hamraa kaul kar-k-ai.
Santee-FM think-PFV-3.NH.S that his mother me.ACC call do-PFV-3.NH.S
"Santee$_i$ thinks that his$_{k,*i}$ mother called him/me$_i$." (if 'me'=Santee, then *'his'=Santee)

b. Santee-aa Bantee-aa-ke kah-l-ai ki okar maiyaa toraa kaul kar-k-ai.
Santee-FM Bantee-FM-DAT tell-PFV-3.NH.S that his mother your.NH.GEN call do-PFV-3.NH.S
"Santee$_i$ told Bantee$_k$ that his$_{*i,*k,n}$ mother called him/you$_k$." (if 'you'=Bantee, then *'his'=Bantee or Santee)

c. Santee-aa Bantee-aa-ke kah-l-ai ki u toraa dekh-l-ai.
Santee-FM Bantee-FM-DAT tell-PFV-3.NH.S that he you.NH.ACC see-PFV-3.NH.S
"Santee$_i$ told Bantee$_k$ that he$_{*i,n}$ saw you$_k$." (if 'you'=Bantee, then *'he'=Santee.)

d. Santee-aa Bantee-aa-ke kah-l-ai ki u hamraa dekh-l-ai.
Santee-FM Bantee-FM-DAT tell-PFV-3.NH.S that he me.ACC see-PFV-3.NH.S
"Santee$_i$ told Bantee$_k$ that he$_{*k,n}$ saw me$_i$." (If 'me'=Santee, then *'he'=Bantee.)

There is no blanket prohibition against a third person pronoun referring to an NP in the matrix clause from within a complement clause that contains a shifted indexical. This is possible when the third person pronoun refers to an NP in the matrix clause which is not one that 'I' or 'you' could refer to—such as the possessor of the subject. 'Him' can refer to 'Santee' in (72) while 'I' refers to Santee's mother.

(72) Magahi (fieldwork, Deepak Alok)

Santee-aa-ke maiyaa soch-l-ai ki (ham) okraa bajaar-me dekh-l-i.
Santee-FM-GEN mother think-PFV-3.NH.S that (I) him.ACC market-in saw-PFV-1.S
"Santee$_i$'s mother$_k$ thinks that she/I$_k$ saw him$_{i,n}$ in the market."

This effect is not part of the core data that Deal (2020) tries to explain within her shifty operator analysis. Nor is it obvious why it should hold in those terms. Third person pronouns are not interpreted relative to a context in the way that indexicals are. Therefore, there is no obvious reason why the presence of a context-shifting operator should affect them. Patterns like this have been observed in the previous literature. Anand (2006: §2.6.5) has a brief discussion under the heading of the "obviation effect". He relates it to Schlenker's (2003) discussion of presupposition maximization, which leads speakers to avoid the use of negative feature values like third person.[38] See also Podobryaev (2014: 101) for an "Elsewhere 3rd person principle" and Spadine (2020: 169) for a "Realize Person Features" principle. These are blocking-type preference principles of unclear theoretical status. Along these lines, 'he' in (71) referring to Santee or Bantee is blocked by the preferred possibility of shifted 'me' referring to Santee or shifted 'you' referring to Bantee. This is descriptively accurate, and far be it from me to say that this analysis cannot be made to work. However, a pragmatic account needs to be stated with considerable care, given that a third person pronoun referring to a matrix argument is not blocked in (70), even though 'I' would be possible referring to Santee and 'you' would be possible referring to Bantee in these examples too. One thus would need to be very careful about what is compared to what within a pragmatic blocking account.

In contrast, my analysis can attribute this robust effect in Magahi to the familiar Rule H (Fox 2000, Safir 2004, Büring 2005). This says that in a structure like [… NP > pronoun1 > pronoun2…], where each element c-commands the following element and pronoun1 depends on the NP for its reference, pronoun2 can only depend on NP by depending directly on pronoun1. This condition plays a role in explaining why (73c) is bad in English, even though (73a) and (73b) are both possible. (73b) shows that it is possible in general for *her* in the embedded clause to take the matrix subject *Mary* as its antecedent. However, this is not possible in (73c), where *she* also takes *Mary* as its antecedent and *she* c-commands *her*. Rule H implies that in this configuration a pronoun in the object position of the embedded clause can only refer to Mary by taking *she* as its direct antecedent. That then

---

[38] Anand (2006: 114 (342)) makes the intriguing observation that this effect does not hold in Zazaki, although it does in Amharic, Navajo and Slave. I do not know why this difference should exist.

requires the pronoun to have the reflexive form *herself*, as in (73d), since its antecedent c-commands it within the same clause.

(73) English (personal knowledge)
  a. Mary$_i$ thinks that she$_i$ embarrassed John at the party.
  b. Mary$_i$ thinks that John embarrassed her$_i$ at the party.
  c. *Mary$_i$ thinks that she$_i$ embarrassed her$_i$ at the party.
  d. Mary$_i$ thinks that she$_i$ embarrassed herself$_i$ at the party.

In the current context, Rule H can also be used to explain (71). Consider for example (71c), with the structure given in (74).

(74) Sp* Ad* sa [Santee$_i$ told Bantee$_k$ [Sp$_i$ Ad$_k$ Fin that [he$_i$ saw you$_k$]]].

Here Ad must be controlled by *Bantee* in order for 'you' to refer to Bantee, as intended. Therefore, Sp must be controlled by *Santee* (this is a form of Shift Together; see §4.5). Now 'he' refers to Santee on the intended interpretation. But so does Sp, and Sp c-commands 'he' and is c-commanded by *Santee*. Therefore, 'he' must depend directly on Sp, not *Santee*, by Rule H, just as *her(self)* must depend on *she* rather than *Mary* in (73c,d).[39] But Sp is [+1]. Therefore, the pronoun that it binds must be [+1] as well; it must be 'I' not 'he'. The other examples in (71) can be explained in the same way. Note that this account is possible because Sp and Ad are syntactically represented on my account, not just part of the interpretative apparatus. Since they are syntactically represented, they have well-defined c-command domains and are visible to binding theoretic principles like Rule H.

This has some further significance for comparing theories. Part of Anand's (2006) oft-cited argument that some *de se* items (like shifted 'I') do not involve pronoun binding while others (like logophors) do is based on the so-called *de re* blocking effect. Logophors are supposed to be susceptible to this effect, whereas shifted indexicals (by implication) are not. *De re* blocking gives the contrast in (75) in Yoruba (Adesola 2005). (75b) shows that it is possible in principle for a logophor and a plain pronoun to refer to the same matrix clause

[39] Rule H might allow the pronoun to depend on *Santee* rather than Sp if this would give a different interpretation. I do not explore whether there might be special situations (e.g. with focus or ellipsis) where 'he' can depend directly on *Santee*.

antecedent in Yoruba. But (75a) shows that this is not possible when the plain pronoun c-commands the logophoric pronoun. Anand assumes that logophors refer to their antecedents *de se* and plain pronouns refer to their antecedents *de re*. The generalization, then, is that a *de se* element cannot be c-commanded by a *de-re* element—drawing a connection between this Yoruba contrast and the behavior of pronouns in dream complements in English.

(75) Yoruba (Anand 2006: 57)

a. Olu so pé o rí bàbá òun.
Olu say that he see father LOG
"Olu$_i$ said that he$_{k,*i}$ saw his$_i$ father."

b. Olu so pé bàbá rè ti rí iyá òun.
Olu said that father his ASP see mother LOG
"Olu$_i$ said that his$_{i,n}$ father saw his$_i$ mother."

I return to this contrast in a range of languages with logophoric pronouns in Chapter 5. Indexical shift analogs of this contrast would be (76). In (76a), an ordinary third person pronoun c-commands a *de se* first person pronoun capable of referring to Santee. In (76b), the third person pronoun does not c-command the first person pronoun.

(76) Magahi (fieldwork, Deepak Alok)

a. Santee soch-l-ai ki Bantee-aa okraa hamar kitab lauTaa det-ai.
Santee think-PFV-3.NH.S that Bantee-FM him.DAT my book return give-3.NH.S
"Santee$_i$ thinks that Bantee will return to him$_{k,*i}$ his/my$_i$ book." (if 'my'=Santee, then *'him'=Santee)

b. Santee-aa soch-l-ai ki okar maiyaa hamraa kaul kar-k-ai.
Santee-FM think-PFV-3.NH.S that his mother me.ACC call do-PFV-3.NH.S
"Santee$_i$ thinks that his$_{*i,k}$ mother called him/me$_i$." (if 'me'=Santee, then *'his'=Santee)

Indeed, (76a) is bad in Magahi with my=him=Santee, as (75a) is in Yoruba. And we have a reason why: like (74), (76a) violates Rule H plus the feature matching condition on bound pronouns. We could also describe this as *de re* blocking, assuming that the shifted

indexical ‘my’ refers to its antecedent *de se* and the third person pronoun refers *de re*. There is no clear difference between a logophor and a shifted indexical here. However, (76a) does not stand out as an instance of *de re* blocking because (76b), where there is no c-command relationship between the two pronouns, is also ruled out by the same principles, whereas (75b) is possible in some languages. I argue in Chapter 5 that this is because logophoric pronouns are nondistinct in formal features from ordinary pronouns in some languages. However, first person pronouns (including Sp) are distinct from third person pronouns in all languages. [+1] and [+2] are universal features that conflict with third person in all languages. In contrast, [+log] is a language particular feature and a subtype of third person; whether it conflicts in features with an ordinary third person pronoun or not depends on the details of the feature system of a particular language. Once this difference is abstracted away from, it is not clear that there is an important difference between shifted indexicals and logophoric pronouns in terms of *de re* blocking. When all the pieces are lined up, I will claim that this undercuts Anand’s argument that logophoric pronouns are bound variables whereas shifted indexicals are not.

4.4. The PLC and embedded Sp and Ad

So far we have seen that the Person Licensing Condition applies to restrict the binding of first and second person pronouns in the domain of Sp and Ad operators, thereby capturing certain Shift Together effects which are attributed to context overwriting in the shifty operator theory. However, the PLC does not restrict Sp* and Ad* in a root clause, because those elements, although they also bear the features [+1] and [+2], have a “grammatically assigned semantic value” imposed on them as the specifiers of sa, the speech act head(s). Sp* is set as referring to the speaker of the speech act, and Ad* to its addressee. My system allows for one other kind of [+1] or [+2] element: namely the ghostly DPs Sp and Ad in Spec FinP, including FinPs in embedded clauses. I complete the discussion of pronoun binding in these ghostly operator constructions by considering how the PLC applies to this case.

Sp and Ad in Spec FinP are parallel to Sp* and Ad* in Spec saP in many respects, but a crucial difference is that Fin does not have the specialized speech-act denoting meaning that sa does. One consequence of this is that FinPs (or larger CPs that contain them) can be selected by verbs as complements, and can appear in other

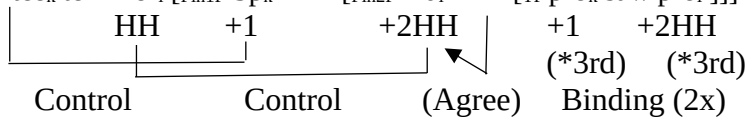
embedded positions, like adjunct clauses and relative clauses. Another consequence of this is that Fin1 and Fin2 do not fix a specific interpretation for their arguments the way that sa1 and sa2 do. Nevertheless, I assume that Fin1 does impose the formal feature [+1] on its argument and Fin2 does impose the formal feature [+2] on its argument, as stated in (77). Fin1 and Fin2 can be thought of as grammaticalized versions of sa1 and sa2, which retain the formal features of the originals but not the semantic substance.

(77) a. The DP specifier of Fin1 is [+1]. (This DP is called Sp.)
b. The DP specifier of Fin2 is [+2]. (This DP is called Ad.)

Sp and Ad then have the fixed formal features of first and second person pronouns, but not the special semantics that is usually attributed to them (which in the current proposal is found inherently only on Sp* and Ad*, but is inherited by their bindees).

These feature values lead to two possible outcomes, I claim. One is that Sp and Ad can undergo obligatory control, when the clause containing them is in the right structural position, merged inside VP (see §4.3.4). Since Sp and Ad do not have intrinsic interpretations, this does not lead to semantic incoherence; rather, it results in indexical shift given the structure in (3), repeated again in (78). OC is a way of Sp and Ad receiving a "grammatically assigned semantic value", on a par with having an interpretation imposed on them by sa1 and sa2. Therefore, controlled instances of Sp and Ad are also not subject to the PLC given the italicized restrictions in (52); they are not bound by [+1] and [+2] elements, and do not need to be. First and second person pronouns in the core of the embedded clause satisfy the PLC by being bound by Sp and Ad, respectively. However, Sp and Ad are bound only by 'Santee' and 'the professor', both third person nominals.

(78) Santee_k tell Prof_i [$_{\text{Fin1P}}$ Sp_k Fin1 [$_{\text{Fin2P}}$ Ad_i Fin2 [$_{\text{TP}}$ pro_k saw pro_i]]]



The second possibility is that Sp and Ad are not controlled by arguments of the verb that selects the CP that contains them. OC always fails when the clause that immediately contains Sp and Ad is not generated inside VP—in relative clauses and high adjunct clauses,

for example. When OC does not happen, Sp and Ad do not get a syntactically defined semantic value, so they are subject to the PLC. As a result, they need to be bound by elements that are themselves [+1] and [+2]—often by higher instances of Sp and Ad (including Sp* and Ad*). Indeed, they have to be bound by the closest higher instances of Sp and Ad, given the relativized minimality character of the PLC. This can be seen in (79), under the analysis sketched in (80).

(79) Magahi (fieldwork, Deepak Alok)
Santee-aa ***masTar saaheb-ke*** *kahl-ai ki Bantee-aa ghare ruk gel-**ain** taaki bimaar na paD-**ain/*au**.*
Santee-FM teacher HH-DAT told-3.NH.S that Bantee-FM home stay went-HH.AL so.that sick NEG got-HH.AL/*NH.AL
"Santee told the teacher that Bantee stayed home so that he would not get sick." (said to a friend)

(80) [Ad*$_n$ sa [S told teacher$_i$ [Ad2$_i$ C B [stay...] [Ad3$_{i,*n}$ C he not.sick]]]]
NH HH HH HH
*NH

Here Ad* in the root clause is nonhonorific, by hypothesis (this can be made explicit by using the verb *kahlau* 'told.NH.AL'). Ad2 in the complement of 'told' is high honorific, since it is controlled by 'teacher', the goal argument of 'told'. (This control is optional, but we can see that it happened in (79) because the allocutive marking on 'stay' is *-ain*, not *-au*). The instructive thing is what happens with Ad3 in the CP adjoined within the complement clause. This cannot be controlled by an argument of 'stay', because the CP is not inside the VP headed by 'stay'. Therefore, it has to be bound by the closest c-commanding [+2] element. This is Ad2—not Ad*. As a result, Ad3 must be HH, and the allocutive marking that surfaces on 'get sick' can be *-ain* but not *-au*. This conclusion recaps reasoning from Chapter 3, now presented in the context of a fuller exposition of the PLC.

We can replicate this observation by using i-shift examples to show that the Sp coordinate in an adjunct clause is also subject to the PLC. Suppose a complement clause has a 'because'-clause adjoined to it, and there are first person indexicals in both the 'because'-clause and the embedded clause proper, as in (81).

(81) Magahi (fieldwork, Deepak Alok)

Sita soch-l-ai ki hamar bhaiwaa Bantee-aa-ke maar-l-ai kaaheki Bantee-aa hamra beijatii kai-l-ai.

Sita think-PFV-3.NH.S that my.GEN brother Bantee-FM-ACC hit-PFV-3.NH.S because Bantee-FM my.GEN insult do-PFV-3.NH.S

"Sita$_i$ thinks that her/my$_i$ brother hit Bantee because Bantee insulted her/me$_{i,*sp*}$." (If 'my'=Sita, then 'me'=Sita not Sp*; also if 'my'=Sp*, then 'me'=Sp* not Sita.)

Suppose that 'my' in 'my brother' has the shifted meaning, where Sita is pondering the violent behavior of her own brother and its causes. Then 'my' in the adjunct clause expressing the insultee must refer to Sita as well, and not Sp*, the speaker of the whole sentence. In other words, Sita's brother is defending her honor, not the speaker's. This shows that Sp in the adjunct clause must be bound by Sp in the CP complement of 'think', not by Sp* in the root clause. This demonstrates that Sp like Ad obeys the locality condition embedded in the PLC when it is not in a position to undergo OC. I conclude that the PLC holds for uncontrolled Sp and Ad as well as for ordinary first and second person pronouns in argument positions. This results in a range of Shift Together effects for elements in different clauses as well as for pronouns in the same clause.

5. The optionality of indexical shift and shift together

5.1. Introduction

One important aspect of indexical shift in Magahi that I have not faced directly in this chapter yet is the fact that it is in some sense optional: 'I' and 'you' in complement clauses or 'so that' clauses can shift such that they refer to the arguments of the immediately superordinate verb, but they do not need to do so. Rather, they can simply refer to Sp* and Ad*, as we have seen throughout. The simple example in (82), for instance, can have the English-like meaning "Santee told Bantee that I saw you" as well as the indexical-shifted meaning "Santee told Bantee that he saw him."

(82) Magahi (fieldwork, Deepak Alok)

Santee-aa Bantee-aa-ke kahk-ai ki ham toraa dekh-l-i ha-l.
Santee-FM Bantee-FM-DAT told-3.NH.S that I you.ACC see-PFV-1.S be-PFV
"Santee$_i$ told Bantee$_k$ that I$_{sp*}$ saw you$_{*k,ad*}$." or: "Santee$_i$ told Bantee$_k$ that he/I$_i$ saw him/you$_{k,*ad*}$."

Magahi is different in this respect from some other indexical shift languages, such as Uyghur, where pronouns inside full finite (nonnominalized) CP complements must undergo indexical shift (Sudo 2012, Shklovsky and Sudo 2014, Major 2022). In Uyghur, a first person pronoun in a structure like (83a) must refer to the matrix subject, not Sp*. Similarly, the second person pronoun in (83b) must refer to the matrix goal, not Ad*, and in (83c) both 'I' and 'you' must shift together to refer to Ahmet and Aygül, respectively.

(83) Uyghur (Shklovsky & Sudo 2014: 383 (4b), (5b); 395 (38))

a. Ahmet [men ket-tim] di-di.
Ahmet 1SG leave-PST.1.SG.S say-PST.3
"Ahmet$_i$ said that he/I$_{i,*sp*}$ left."

b. Tursun Muhemmet-ke [(pro) xet jaz-ding] di-di.
Tursun Muhemmet-DAT you letter write-PST.2.SG say-PST.3
"Tursun told Muhemmet$_i$ that he/you$_{i,*ad*}$ wrote a letter."

c. Ahmet Aygül-ge [men seni yaxshi kör-ymen] di-di.
Ahmet Aygül-DAT 1SG 2SG.ACC well see-IPFV.1SG say-PST.3
"Ahmet$_i$ told Aygul$_k$ that he/I$_{i,*sp*}$ likes her/you$_{k,*ad*}$.'

From the point of view of the theoretical tools I have adopted, the Uyghur pattern is the more expected one. A key ingredient of the theory is that Sp and Ad undergo OC by suitable arguments of the matrix verb. Now a natural understanding of "obligatory control" is that it is obligatory. That seems right for Uyghur, but wrong for Magahi, and one would like to know why.

Another important layer to this question of optionality is that sentences like (82) in Magahi seem to have one dose of optionality rather than two. If u-shift and i-shift are both optional in Magahi complement clauses, and the two options are independent of each other, then (82) should have four possible readings rather than two: It should also be able to mean "Santee told Bantee that I (Sp*) saw him

(Bantee)" and "Santee told Bantee that he (Santee) saw you (Ad*)." However, it cannot mean these things; rather i-shift takes place in the complement clause if and only if u-shift takes place. This is another aspect of the Shift Together phenomenon, first discussed by Anand & Nevins (2004) for Zazaki, and an important part of theoretical discussions of indexical shift ever since. It has also been observed in Uyghur (Sudo 2012, Shklovsky and Sudo 2014), Nez Perce (Deal 2020), as well as Magahi (Alok and Baker 2018, Alok 2020) and other languages. It shows that although both kinds of indexical shift can be in some sense optional, they are not *independently* optional. This needs to be understood.

This section unfolds as follows. In §3.4.5, I briefly presented a solution to question of what the source of optionality is for allocutive shift in Magahi, proposing that it is due to the CP complement extraposing to the periphery of the clause that contains it, causing what is thematically a complement clause to behave like an adjunct clause for purposes of the GOCS. A quick comparison of Magahi with Uyghur supports this idea. Both are generally head-final SOV languages, but with CP complements there is a difference: Magahi systematically shows S-V-CP order, as seen in (82),whereas Uyghur maintains what could be the underlying S-CP-V order in (83). The language that does not show overt evidence of CP extraposition does not show optionality in indexical shift either, suggesting that the two may be related. I explore this hypothesis further in §4.5.2. In §4.5.3, I discuss another factor that can make indexical shift optional or obligatory in particular languages—the possibility that a clause can be nominalized—using data from Uyghur and Slave. Then in §4.5.4, I discuss Deal's (2020) proposal that violations of Shift Together can arise because one of the two operators that result in indexical shift can fail to by selected by particular verbs in particular languages. Although my theory could get the same results in an analogous way, I argue that there is not enough evidence across languages to justify this instance of parametrization. Rather, the few examples of apparent Shift Together violations that have been documented (especially in Slave) can be explained by saying that the null subject of the clause is an indexiphor rather than a true indexical pronoun.

5.2. Optionality through extraposition

First, let us fill in the theory of why extraposition is potentially very relevant to the fact that indexical shift is optional in some languages but not others. The GOCS says that OC applies to clauses that are

inside VP. Is that where finite CP complements are in Magahi? The answer appears to be "sometimes". CP complements are (by hypothesis) generated inside VP, but they move outside of VP by PF. My hypothesis, then, is that the optionality we see in Magahi indexical shift is a connectivity effect: the moved CP can be interpreted for purposes of control either in its base position inside VP or in its derived position outside VP. If it is interpreted inside VP, then the GOCS applies, such that Sp and Ad are both controlled by the corresponding arguments of the matrix verb. If, however, CP is interpreted outside of VP, the structure is not significantly different from one with a high CP adjunct adjoined to TP, like a 'because'-clause or a 'when'-clause. This is not a context of obligatory control, according to the GOCS. As a result, indexicals do not shift in these CP-adjuncts (see (43)), and I suggest that they can not shift in extraposed CP complements for the same reason. On this account, the fact that 'I' and 'you' inside a moved clause have two readings each in (82) is parallel to the fact that *herself* inside a moved DP has two readings in (84), depending on whether the DP containing it is interpreted for Binding theory in its base position or its moved position (Chomsky 1993) among many others). In contrast, when *herself* is inside a DP that does not move by *wh*-movement, *herself* has only one reading, as in (84b).

(84) English (personal knowledge)
a. Kyle$_i$ wonders [[which pictures of himself$_{i,k}$]$_n$ [Max$_k$ likes ~~DP~~$_n$ best]].
b. Kyle$_i$ thinks that [Max$_k$ likes [the pictures of himself$_{k,*i}$] best]].

The LFs of the two readings of (82) in Magahi are sketched in (85). In contrast, Uyghur only has the structure in (85a), so it has only one reading, the indexical-shifted one.

(85) a. Santee [$_{VP}$ Bantee [$_{CPi}$ Sp Ad I you see] tell] ~~CP~~$_i$.
√OC

b. Santee [$_{VP}$ Bantee ~~CP~~$_i$ tell] [$_{CPi}$ Sp Ad I you see].
*OC

This account of the optionality of indexical shift in CP complements in Magahi contributes to an explanation of the Shift Together

phenomenon seen in (82), given reasonable auxiliary assumptions. First, we must hold that it is impossible for a constituent including Ad to extrapose leaving a constituent containing Sp behind (and vice versa). In terms of the structure proposed by Alok (2020), we must rule out Fin2P moving out of Fin1P, leaving Sp behind inside the VP. If that were possible, then Ad could be interpreted outside VP, as bound by the higher Ad*, while Sp is interpreted inside VP as controlled by the matrix subject. But this is easy to rule out, given that it is always whole clauses that extrapose, not partial clauses stranding some C-like heads in situ (for example, *ki* cannot be left before *kahk-ai* 'told' in (82)).

Another assumption about CP extraposition to shore up is that it moves CP all the way out of VP in Magahi. If it merely right-adjoined CP to VP, the GOCS would still force obligatory control, and CP complements would behave not like high adjuncts but rather like low adjuncts. The necessary assumption is supported by the fact that in Magahi (as in other languages) the rightward-moved CP comes after not only the thematic verb but also the finite auxiliary whenever there is one. This can be seen in (86).

(86) Magahi (fieldwork, Deepak Alok)
Santee-aa ***chaaha h-ai*** *[ki ham parichha paas ho jaa-i]*.
Santee-FM want be-3.NH.S that I exam pass become go-1.S
"Santee$_i$ wants that he/I$_{i,sp*}$ pass the test."
*Not: *Santeeaa* ***chaaha*** *[ki ham parichha paas ho jaai]* ***hai****.*

The last theoretical point to shore up is how reconstruction works. Another way to avoid Shift Together could be to have the CP as a whole extrapose, but then allow some kind of "scattered deletion" at LF, such that Sp is interpreted inside VP and Ad is interpreted outside VP (or vice versa). However, it is standardly assumed that connectivity effects must be coherent: a moved constituent is interpreted for a given purpose in one position or another but not both (Fox 2000). I assume that something like that holds in this domain too. This leads to a new research question: how does the optionality of indexical shift in a language like Magahi interact with other types of reconstruction effects? For example, does one get obligatory reconstruction for Condition C in clauses with shifted indexicals but not in clauses with unshifted indexicals? What about quantifier scope and bound variable anaphora? I have not investigated such questions, and they go beyond the bounds of this inquiry, but they could be

worthy topics for future investigation.

A prediction of this account is that when Sp and Ad in a CP complement are not controlled by arguments of the matrix verb, they should behave in the same way as Sp and Ad do in high adjunct clauses. In particular, they should obey the PLC, such that they are bound by the closest c-commanding Sp and Ad. In singly embedded clauses like (82) and (86), this is Sp* and Ad*, which results in unshifted readings of the first and second person pronouns. But in a doubly embedded clause, the prediction is that they will have the same potentially shifted reading that pronouns in the intermediate clause have. This is borne out. For Sp, we can test this using i-shift. Consider (87), looking at readings in which 'my' in the intermediate clause shifts to refer to 'Santee', the subject of the highest clause, but where 'I' in the lowest clause does not shift to refer to 'Bantee', the subject of the intermediate clause. These would have a structure like (88), where the CP complement of 'think' is interpreted in its VP-internal position but the CP complement of 'tell' is interpreted in its extraposed position.

(87) Magahi (fieldwork, Deepak Alok)
Santee-aa socha h-ai ki Bantee-aa hamar baabaa-ke kahk-ai ki ham igjaam me phel ho ge-l-i.
Santee-FM think be-3.NH.S that Bantee-FM my.GEN grandfather-DAT told-3.NH.S that I exam in fail happen go-PFV-1.S
"Santee$_k$ thinks that Bantee told his/my$_k$ grandfather that he/I$_{k,*sp*}$ failed the exam." (also possible: I=my=Sp*)

(88) Sp*$_i$ C Santee$_k$ [$_{CP2}$ Sp2$_k$ that Bantee$_n$ [my$_k$ grandfather ~~CP3~~ told] [$_{CP3}$ Sp3$_{k,*i}$ that I $_{k,*i}$ failed the exam]] think ~~CP2~~

Here Sp* refers to the speaker, as usual. Sp2 is by hypothesis controlled by 'Santee' in this version of the sentence, with the CP2 that contains it interpreted inside the VP headed by 'think'. Our primary interest is what happens with Sp3 in CP3, the lowest clause, the complement of 'tell', when this is interpreted in its extraposed position, adjoined to the intermediate TP. The prediction based on the PLC is that it must be bound by Sp2, not by Sp*, since it is still in the c-command domain of Sp2. (This assumes that clausal extraposition is subject to Ross's (1967) right roof constraint, such that it can only adjoin higher up in the same clause it was generated in.) So 'I' here is predicted to refer to Santee (if not Bantee), and not to Sp*. This is correct. I conclude that complement clauses that resist OC are subject

to the strictures of the PLC, which constrain their interpretation. Indeed, the sentence in (87) is interpreted like (81), in which the lowest clause is an adjunct rather than a complement, as expected.[40]

We can also check this prediction for Ad, using shifted allocutive agreement, comparing complement clauses in (89) with how adjunct clauses behave in (79)/(80) ((89) was also discussed in §3.4.5). In (89), the NH marker *-au* on the root clause verb 'tell' shows that the sentence as a whole is addressed to a friend of the speaker. The H marker *-o* on the intermediate verb 'think' is different from *-au* but matches the status of 'grandfather', the goal of the root clause. This implies that the intermediate CP is interpreted in its low position, where it is controlled by the matrix goal. (Having *-au* marking on this verb is also possible.) Now the key question is what kind of allocutive marking can one have on the verbal complex 'pass go' in the lowest clause. Here there is no possibility of a locally controlled reading, because 'think' does not have a goal/theme/undergoer argument that could control the Ad of its complement. So Ad must be a bound [+2] pronoun in this case, with CP3 interpreted in its extraposed position. The crucial observation is that allocutive marking on 'pass go' can be *-o* matching the intermediate verb but cannot be *-au* matching the highest verb. This shows that the lowest Ad is constrained to take as its antecedent the closest c-commanding [+2] item, the intermediate Ad rather than the highest one (Ad*). This obeys the PLC and matches the behavior of base-generated adjunct clauses like in (79).

(89) Magahi (fieldwork, Deepak Alok)
*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

[40] See Anand & Nevins (2004) for a parallel analysis of a similar sentence in Zazaki illustrating "no intervening binder" (see also Deal (2020: 42), who calls the relevant constraint "local determination"). Although closely related, there is a difference in that the shifty operator analysis has no operator in the lowest clause, whereas I have DP operators there but they are bound rather than controlled. On my analysis, this is forced at least for Ad by the fact that there must be something in the vicinity of the lowest Fin for it to agree with in (89).

test." (to a peer)

I conclude that when complement clauses do not undergo OC it is because they are behaving like adjunct clauses, as an account based on CP extraposition expects.

Finally, we can investigate this hypothesis typologically, by seeing if it makes the right crosslinguistic predictions. Based on the comparison between Magahi and Uyghur, the expectation is that in languages in which word order shows that CP complements are extraposed indexical shift should be optional (like Magahi), whereas in languages where word order shows that CP complements stay in situ indexical shift should be obligatory (like Uyghur). We did not know of enough languages with shifted allocutive marking in embedded clauses to study the question from this angle in Chapter 3, but more crosslinguistic data is available on indexical shift. I consider the results encouraging. Languages in which indexical shift is said to be obligatory (in some constructions) include Slave, Navajo, Matses, and Laz as well as Uyghur.[41] All are verb final languages, and indeed the verb comes after the CP complement in all these languages. For example, indexical shift is obligatory with certain verbs in Slave, including 'say'. In (90a), the first person pronoun must refer to the matrix subject, not the speaker, and in (90b) the second person pronoun in the embedded clause must refer to the matrix goal, not to the addressee of the sentence as a whole. In both cases, the embedded clause comes immediately before the matrix verb (plus a postposition in (90b)). Indeed, Rice (1989: 1239) is explicit that CP complements cannot extrapose rightward in Slave (although CP subjects can).

(90) Slave (Rice 1989: 1279 (37), 1277 (24))

a. John [hı̨dowedzíné k'e deshı̨ta duhła] *hadi.*
John tomorrow on bush 1.SG.S.OPT.go 1.S.say
"John$_i$ said that he/I$_{i,*sp*}$ is going to the bush tomorrow."

---

[41] Different from CP extraposition rightward is CP topicalization leftward. Rice (1989) shows that this is possible in Slave, and it appears to be independent of indexical shift. CPs of verbs that require indexical shift still show indexical shift if the CP is topicalized, and CPs of verbs that forbid indexical shift still forbid it if the CP is topicalized. If this is right, then I have to say that reconstruction is obligatory with this form of CP movement in Slave.

b. John ts'ódanike [yerigha yejai táʔerase] gho góhdi.
John child.PL why window 2.PL.S.broke about 3.S.ask.4.PL.O
"John asked the children$_i$ why they/you$_{i,*ad*}$ broke the window."

The facts for Matses, a Panoan language, are similar. Indexical shift is obligatory in the complements of the verbs 'say', 'tell' and 'suppose mistakenly'. Thus Munro et al. (2012: 48) write that in this language "The overarching constraint that can be seen is perspective persistence: indexicals must remain from the point of view of the original speaker" (i.e. the referent of the subject of a verb like 'say'). Thus, the null first person pronoun in the complement clause of (91a) refers to Dashe, not to Sp*, and the second person pronoun in the complement clause of (91b) refers to Sp*, the person that the referents of 'they' were talking to, not Ad*, the person that Sp* is talking to. Correlated with this, the embedded CP comes immediately before the matrix verb 'say' in both examples.

(91) Matses (Munro et al. 2012: 47 (14), 51 (22))
a. Dashe [kachina pe-o-mbi] ke-o-şh.
Dashe chicken eat-PST-1 say-PST-3
"Dashe$_i$ said that he/I$_{i,*sp*}$ ate chicken."

b. [Mibi bëda-mbo ik-e-k] ke-o-şh, ubi chui-ek.
you good-AUG be-NPST-IND say-PST-3 I speak about-SS
"They said about me$_{sp*}$ that I/you$_{sp*,*ad*}$ are a good person."

A fourth language of this type is Laz, which has obligatory indexical shift in certain kinds of CP complements and shows S-CP-V order.

(92) Laz (Demirok and Öztürk 2015): 45 (2a))
Arte-k [ma noseri vore ya] iduşun-am-a.
Arte-ERG I smart be.1.SG that think-IPFV-3SG
"Arte$_i$ thinks that he/I$_{i,*sp*}$ am smart."

So the often-cited languages that have obligatory indexical shift also do not have CP-extraposition, at least not obligatorily.[42]

[42] Two other languages that Deal (2020) mentions as having obligatory indexical shift are Navajo, an Athapaskan language related to Slave (Schauber 1979,

Now consider languages that are known to be like Magahi in having optional indexical shift. The Iranian language Zazaki is a classic case (Anand and Nevins 2004). It has mixed word order, with nominal objects appearing before the verb. However, finite CP complements come after the verb, extraposed rightward as in Magahi.[43]

(93) Zazaki (Anand & Nevins 2004 (13))
Vızeri Rojda Bill-ra va [kɛ ɛz to-ra miradiša].
yesterday Rojda Bill-to said that I you-to angry.be.PRES
"Yesterday Rojda$_i$ said to Bill$_k$ that she/I$_{i,sp*}$ is/am angry at him/you$_{k,ad*}$."

The North American language Nez Perce also has optional indexical shift. The language has very free word order when it comes to NPs and verbs, but Deal (2020) consistently gives finite CP complements in sentence-final position. This suggests they also extrapose rightward.

(94) Nez Perce (Deal 2020: 56)
a. Isii-ne A. hi-i-caa-qa [cewcewin'es-ki (pro) 'e-muu-ce-Ø]?
who-ACC A. 3.s-say-IPFV-REC.PST phone-with I 3.O-call-IPFV-PRES
"Who$_i$ did A$_k$ say she/I$_{k,sp*}$ was calling?"

Speas 2000) and Dobon (Davies 1981). Dobon generally frames a finite embedded clause between two speech predicates, resulting in forms like 'Yesterday he to.me **say-(SS.)3SG** [I home go-1SG] **say-3SG**' (QUOTE do-3SG), meaning: "Yesterday he told me that he was going home" (Davies 1981: 1). Although the first verb generally comes with more descriptive content (overt arguments and modifiers, more lexical content), there is reason to say that the clause is syntactically the complement of the second verb: (i) this fits the language's SOV word order, (ii) the second verb is strictly obligatory whereas the first can be omitted in context (Davies 1981: 2-3), and (iii) the first verb can be marked with subordinate same-subject morphology whereas the second verb always has full matrix clause finite inflection. Given that CP is the complement of the final verb, Dobon does not allow CP extraposition and does have obligatory indexical shift, fitting well within my generalization.

[43] However, because of Zazaki's mixed word order, a Kayne (1994)-style analysis in which complements follow the head but NP objects move leftward might be more plausible for Zazaki than for some more rigidly head-final languages.

b. Manaa we'nikt 'u-us haama-nm, ke ko-nya T-nm pee-Ø-n-e R-ne ['ee 'o-opayata-yo'qa]?
how name.NOM 3.GEN-be.PRES man-GEN C REL-ACC T-ERG 3.S/3.O-say-PFV-REM.PST R-ACC 2SG 3.O-help-MOD
"What is the name of the man$_k$ that T told R$_i$ that he/you$_{i,ad*}$ should help?"

So we see a good correlation between the presence of CP extraposition and the optionality of indexical shift, with four reasonably well-studied languages on one side of the ledger and three on the other side. I take this to be very promising for the view that the former causes the latter. More generally, I do not know of any language that has clear CP extraposition in which indexical shift is obligatory. The other side of the correlation may be more complex, in that there are a few languages without obvious CP extraposition which nevertheless have optional indexical shift—Amharic, for example (see Schlenker 2003, Anand 2006).[44] This could be either because CP movement is harder to see in these languages (perhaps it can be string-vacuous; see fn. 34 of Chapter 3 on Tamil), or other structural factors also play a role. This is a topic for future research.

5.3. Optionality through nominalization

In fact, I already know of one other structural factor that can block otherwise obligatory control of Sp and Ad, hence indexical shift. This is nominalization. I mention this for completeness, as a starting point for further typological investigation. We saw above that in Uyghur full finite CP complements necessarily undergo indexical shift, as seen in (83); (95a) repeats one of these examples. However, like other Turkic languages, Uyghur has a second form of complementation which is essentially synonymous but involves gerund-type nominalization; (95b) for example is very similar to (95a). Nominalization in (95b) is easily recognized by the convergence of several factors: the subject of the embedded "clause" is genitive; the nominalizing affix *lik* is found on the verb, agreement with the subject comes from the possessive paradigm, and the embedded clause as a

---

[44] Turkic languages other than Uyghur also pose questions for this correlation, including Turkish (Major 2022: 804-805 (51)) and Sakha (Nadya Vinokurova, p.c.). For speakers of Japanese and Korean who allow indexical shift (Park 2016, Shiori Ikawa p.c.), it is also optional without obvious CP extraposition.

whole bears accusative case. All these properties are different from the form of complementation seen in (95a). Correlated with this structural difference is the fact that indexical shift is blocked in (95b), where 'I' must refer to Sp*, not Ahmet. This nominalized form of complementation also blocks u-shift.

(95) Uyghur (Shklovsky & Sudo 2014: 383 (4b), (4a); Major 2022)

a. Ahmet [men ket-tim] di-di.
Ahmet 1SG leave-PST.1SG say-PST.3
"Ahmet$_i$ said that he/I$_{i,*sp*}$ left."

b. Ahmet [mening ket-ken-lik-im-ni] di-di.
Ahmet 1SG.GEN leave-REL-NMLZ-1SG.POSS-ACC say-PST.3
"Ahmet$_i$ said that I$_{*i,sp*}$ left."

So indexical shift in Uyghur is optional in the sense that there are two synonymous (as far as we know) forms of complementation to choose from, one which requires it and one which forbids it. However, it is either obligatory or forbidden given the form of complementation.

This effect of nominalization is readily capturable in terms of the GOCS. We can say that any clause-like constituent containing Sp and Ad is not the direct complement of the verb in (95b), because the nominal head of the nominalized clause intervenes. For example, (95b) could have the structure in (96), where FinP is technically the complement of NMLZ, not 'say'.[45]

(96) Ahmet [$_{VP}$ say [$_{NP}$ NMLZ [$_{FinP}$ Sp Ad Fin [I left]]]]

The GOCS does not enforce *Ahmet* controlling Sp in this structure, because the CP containing Sp is not a direct argument of 'say'. This is comparable to Landau's (2013: 43-46) observation that in English OC is not required in certain gerundival complements with the nonfinite

[45] This is the nominalized structure that is most likely to allow indexical shift, considered for the sake of argument. It is very possible that nominalization happens lower in (95b), with a nominal head replacing T and taking a VoiceP complement. Such a structure would not have a FinP projection, so Sp and Ad would not be licensed in the nominalized clause and indexical shift would be impossible for that reason. (The structure in (96) is more likely for (97)-(98) in Slave, where the internal syntax of the clause is entirely verbal.)

verb bearing the affix *-ing*, which can be nominal rather than verbal, whereas complement clauses using the *to*-infinitive always show OC. Similarly, in Japanese embedded clauses headed by *koto* require OC when *koto* is categorically a C, but allow NOC when *koto* is categorically a N (Fujii 2006). Then if OC is blocked in (96), any Sp and Ad inside the nominalized constitutent must be bound by a higher Sp and Ad, in this case Sp* and Ad*, as happens with Ad and Sp in high adjunct clauses or relative clauses.[46]

Crucially the same structural factor—the presence of NMLZ—that blocks OC for Sp also blocks OC for Ad. This follows under the assumption that nominalization cannot take place between the head that licenses Sp and the head that licenses Ad. In Magahi terms, this would happen if a nominal head took a Fin2P complement and was itself the complement of Fin1. It seems plausible to rule this out in as much as the two Fin heads can be thought of as two projections ("shells") of the same head, or closely related heads. (Compare Speas & Tenny's (2004) idea that Sp and Hr are introduced in the analog of a Larsonian shell structure.) This then gives another form of Not-Shift Together. Averaging over the two complementation structures available in Uyghur, both Sp and Ad must be controlled if there is no nominalizing layer to hide them, and neither can be controlled if there is a nominalizing layer. This is a kind of Shift Together.

Another language in which nominalization plays a role in conditioning indexical shift is Slave. Rice (1989) claims that some clausal complements in Slave are dominated by an NP node, whereas others are not. A relatively clear sign of nominalization is the clause triggering object agreement on the selecting verb, such that it bears the so-called areal gender *go-*. The verb 'know' is an example, as in (97). Rice (1989: 1230) writes "The complements are dominated by NP: the areal pronoun *go-/ke-* is present and other nouns can occur in the place

---

[46] The structure in (96) looks very much like the one in (i) for the [N+CP] construction found in examples like (47). The only evident difference is that the NP is headed by an affixal noun NMLZ in (96) but by the ordinary noun 'rumor' in (i). Yet (i) allows indexical shift, whereas (96) does not. The difference is that carrier nouns like 'rumor' and 'news' can also have covert arguments, which count as the possessor/agent argument of the noun. This often-covert argument is the true OC controller of Sp in (i). When it is itself controlled by the agent-subject *Santee*, the result is indexical shift. In contrast, NMLZ in (96) has no additional arguments that can mediate control in this way.

(i) Santee$_i$ [$_{VP}$ spread [$_{NP}$ *(ec$_i$)* rumor [$_{FinP}$ Sp$_i$ Ad Fin [I$_i$ win prize]]]]]

of the complement sentence."

(97) Slave (Rice 1989: 1229 (43))
[John ʔode nágwe] ***ko****dįhshǫ.*
John where 3.S.lives 3.S.know.areal.O
"He knows where John lives."

All matrix verbs that agree with CP in this way count as what Rice calls "indirect discourse" verbs; they do not allow indexical shift. (98) shows two examples in which embedded first person can only refer to Sp*, one with the overt C *gu* and one with a null complementizer.

(98) Slave (Rice 1989: 1272 (2), 1275 (10))
a. John [ʔerákeʔée wihsi gú] ***ko****dįhshǫ.*
John parka 1SG.S.made C 3.S.know.areal.O
"John$_i$ knows that I$_{sp*,*i}$ made a parka."

b. [Judeni ráhgwe] ***ko****dįhshǫ yįle.*
where 1.SG.S.live 3.S-know-areal.O NEG
"She$_i$ doesn't know where I$_{sp*,*i}$ live."

In contrast, verbs like 'say' that require indexical shift have complements that do not trigger object agreement on the verb (there no *go-/ke-* prefix on the matrix verbs in (90) above). Rice (1989: 1274) also says that all verbs in Slave that select an overt complementizer (*gu* or *ni*) are verbs that do not allow indexical shift, whereas verbs that require indexical shift never occur with an overt complementizer. Both of these overt complementizers seem to be at least somewhat nominal in that they are possible with CP subjects as well as with CP complements (Rice 1989, §43.1).

Overall, the signs that a clause is nominalized are subtler in Slave than in Uyghur.[47] Nevertheless, there is evidence that having a nominal

[47] Slave might have an even more subtle kind of nominalization as well. Rice (1989: 1161, 1274) shows that there are verbs whose complements do not trigger agreement on the verb but are weak islands, in the sense that adjunct question words cannot move out of them. The complements of these verbs also do not allow indexical shift. In contrast, this sort of extraction is allowed with verbs like 'say' and 'want', which do allow indexical shift. Perhaps weak islandhood is also a sign of nominalization, which blocks the OC of Sp and Ad.

layer above the clause prevents the control of Sp and Ad in Slave too, blocking indexical shift in a complement clause that might otherwise have it. This is another factor to consider when seeking to understand the distribution of indexical shift across different verbs, constructions, and languages. Some instances in which indexical shift does not take then place in a language that otherwise allows it could be due to nominalization rather than to extraposition of the clause.[48]

5.4. Can C select Sp without Ad?

This section has considered the optionality that is found with indexical shift in some languages but not others—an optionality that seems to be at odds with the role of obligatory control in these constructions. I have identified two factors that can lead to apparent optionality: the possibility of CP extraposition plus optional reconstruction, and the possibility of selecting for a nominalized clause rather than an ordinary CP. Both extraposition-reconstruction and nominalization affect whole clauses, not separating one part of the CP periphery from the rest. As a result, they affect Sp and Ad equally, either allowing both to be controlled or preventing both from being controlled. As such, they contribute to an explanation of the Shift Together generalization of Anand & Nevins (2004) and much subsequent work.

However, there is another, arguably even more basic factor to consider as well. This is the question of whether it is optional for Sp and Ad to be present in the embedded clause in the first place. If one or both of these is absent from the periphery of some clause, then the indexicals that it would bind will not have a shifted reading.

When Sp and Ad are considered together as a set, this may well be a factor in determining whether indexical shift is possible in particular

---

[48] Laz (Demirok & Öztük 2015) and Tigrinya (Spadine 2020) would be interesting languages to look into more to investigate these issues. Both languages have two complementation structures, one of which requires indexical shift (marked with C head *ya* in Laz and *ʔil*+Agr in Tigrinya) and the other of which forbids it (marked with *na* in Laz and *kim-zi-* in Tigrinya). There are some structural differences between the two types of complements, having to do with the order of the C-type head and the verb. However, it is not clear if this relates to *na* and *kim zi-* being nominal in a way that *ya-* and *ʔil*+Agr are not. Another possibility is that *na* and *kim zi-* simply do not license Sp and Ad in these languages, as a lexical property. See Deal (2020) on the role selectional properties play in the distribution of indexical shift.

languages or constructions. For example, English and many other European languages do not allow indexical shift at all. This could be because Fin heads in English do not license Sp and Ad, but only unembeddable sa heads do. At a language-internal level, CPs headed by *ya* in Laz require indexical shift whereas CPs headed by *na* forbid it (see fn. 48). This could be because *ya* has selectional properties that license Sp and Ad whereas *na* does not. All theories have room for brute-force selectional differences like these, and rightly so.

A more distinctive question is whether (in my terms) it is possible for C-type heads to license Sp but not Ad, or Ad but not Sp. If so, this would be a way of generating Shift Together violations in which first person indexicals shift but second person indexicals do not, or vice versa. This could happen in a structure like (99). Here all null DPs in a controllable position are in fact controlled, but there happens to be only one of them. And there should be no objection to 'you' being bound by Ad* at a distance, since bound variable anaphora is not restricted by phases or other measures of absolute syntactic locality. This would result in 'me' undergoing indexical shift but not 'you'.

(99) a. Sp^*_i Ad^*_k C [$Simon_n$ say/think [Sp_n that [you_k hit me_n]]]
b. Sp^*_i Ad^*_k C [$John_n$ tell $Mary_m$ [Sp_n that [I_n help you_k]]]

Deal (2020) claims that there is indeed parametric variation of this kind. Based on her extensive review of the indexical shift literature up to that time, she claims that this sort of Shift Together violation is found in some languages but not others. In languages where Shift Together is strictly respected, the operator that accomplishes second person indexical shift (for Deal, this is Op_{ADDR}) is bundled together with the operator that accomplishes first person indexical shift (Op_{AUTH}) into a single lexical item, which is either present or not. However, in other languages these constitute two distinct functional heads, with Op_{AUTH} lower than Op_{ADDR} in the functional sequence. Languages like Zazaki, Uyghur, Nez Perce, and Magahi have bundled Ops (=Op_{PERS}, see also Anand 2006), so they obey Shift Together. But other languages can have Op_{AUTH} present without Op_{ADDR}. (Not vice versa: for Deal the fixed functional sequence does not allow Op_{ADDR} to be present without the lower head Op_{AUTH} also being present.) Her principal example of a language in which a structure like (99a) is possible yielding i-shift without u-shift is Slave. In the examples in (100) from Rice (1989), first person elements in the embedded clause

refer to the matrix subject, whereas second person elements refer to Ad*. ((99a) is the structure for (100a).)

(100) Slave (Rice 1989: 1279, 1283)

a. Simon náseneineht'u hadi.
Simon 2SG.S.hit.1SG.O 3.say
"Simon$_i$ said that you$_{ad*}$ hit him/me$_i$."

b. Negháyuhdá nudeli.
1SG.S.OPT.see.2SG.O 3SG.S.want.2SG.O
"She$_i$ wants (of you$_{ad*}$) the she/I$_i$ see you$_{ad*}$."

Although there is no strong theoretical reason why my theory could not also adopt representations like (99), I think there are empirical reasons to doubt that UG allows this possibility. The first is simply that Slave is the only language that we know of that allows sentences like (100) (depending on how we characterize the cases; see below). The second is that even in Slave the effect is verb-specific in a particular way. In fact, verbs that select a goal argument in Slave, such as 'tell' and 'ask', do obey Shift Together. Both first and second person indexicals in their complements must shift, as seen in the examples in (101). There is no option of 'I' referring to the matrix subject and 'you' referring to Ad* in this situation, based on Rice's thorough and precise discussion.

(101) Slave (Rice 1989: 1273 (5), 1277 (23), (26))

a. Rosie [ʔerákeeʔée wihsį] sedeyį́dí.
Rosie parka 1SG.S.made 3.S.told.1SG.O
"Rosie$_i$ told me$_{sp*}$ that she/I$_{i\ (*sp*)}$ made a parka."

b. John [ʔaraɲíła] yéhdi.
John 2SG.S.go.home 3.S-told-4SG.O
"John$_i$ told her$_k$ for her/you$_k$ to go home."

c. Se-gha náu̲hdí sédi̲di yi̲lé.
1SG-for 2SG.S.buy 2SG.S.tell.1SG.O PST
"You told me to buy it for you." (lit: You$_{ad*}$ told me$_{sp*}$ you$_{sp*}$ buy it for me$_{ad*}$.')

In other words, structure (99a) may be attested, but the variant in (99b) is not, even in Slave. One could take this to be a selectional matter, saying that 'say' and 'want' select for Sp/Op$_{AUTH}$ but not for

Ad/Op$_{ADDR}$ in Slave, whereas ‘tell’ and ‘ask’ select for both. One could even say that this selection is not arbitrary but is rooted in the lexical semantics of these verbs. But if Shift Together is really parameterized across languages, one should be able to see (99b) in some language, such that a verb with both agent and goal arguments selects only Op$_{AUTH}$ (Deal’s version) or controls only Sp (mine). This does not seem to be attested. I consider this a very suspicious gap in the data.[49]

A curious detail of Magahi also bears on this matter. When the subject of the CP complement of a dyadic verb like ‘think’ is a null pronoun (pro) licensed by rich agreement on the verb, then Magahi behaves like Slave in (100): the first person null pronoun can refer to the matrix subject while ‘you’ refers to Ad*. This is seen in (102a). But like Slave, this apparent violation of Shift Together is only possible under dyadic matrix verbs like ‘think’ and ‘say’, never under a triadic verb with a goal argument like ‘tell’, ‘ask’ or ‘remind’. Even more curiously, the possible Shift Together violation goes away when the subject of the embedded clause is the overt first person pronoun *ham*. Thus, (102b) is different from (102a) in that *ham* ‘I’ can only refer to Sp* in (102b) as ‘you’ continues to refer to Ad*.

(102) Magahi (Alok 2020: 253 (5))

a. Santee-aa soch-l-ai ki (pro) toraa dekh-l-i.
Santee-FM think-PFV-3.NH.S that I you.NH.ACC see-PFV-1.S
“Santee$_i$ thought that he/I$_{i,sp*}$ saw you$_{ad*}$.”

b. Santee-aa soch-l-ai ki ham toraa dekh-l-i.
Santee-FM think-PFV-3.NH.S that I you.NH.ACC see-PFV-1.S
“Santee$_i$ thought that I$_{sp*,*i}$ saw you$_{ad*}$.”

So Shift Together violations of this kind when they arise are narrow and fragile. Even languages that have some such examples have

---

[49] The transitive verb ‘want’ in Slave does trigger i-shift in its complement but not u-shift (see (100b)). But I attribute this to the fact that the object of ‘want’ is not a goal argument, hence not thematically eligible to control Ad. Rather it seems to be a sort of proleptic object, which has to be coreferential with some DP in the complement clause (e.g. She$_i$ wants of you$_{ad*}$ that she/I$_i$ sees you$_{ad*}$’ Rice 1989 1283 (66)). With this proleptic object set aside, (quasi)-transitive ‘want’ behaves like ‘say’ and intransitive ‘want’ in Slave.

closely related constructions in which Shift Together is obeyed.

What we see in (102b) is an instance of what Deal (2020: 84-85) presents as a *defective addressee* effect. The question arises as to how Shift Together works in a sentence in which the matrix clause has no goal for second person elements to shift to. The answer is that second person elements get a null value—and hence are ruled out. This effect is particularly striking in Uyghur, where indexical shift is obligatory in certain kinds of complements. Obligatory shift, plus Shift Together, plus there being no matrix-clause goal for 'you' to shift to conspire to give the result that 'you' cannot be used at all in the CP complement of the relevant verbs (cognitive verbs) in Uygur(!). (103a) shows that i-shift is obligatory when 'believe' takes a finite CP complement headed by *dep* in Uyghur. (103b) shows that a second person pronoun has no possible interpretation in this grammatical context, making the sentence as a whole ungrammatical.

(103) Uyghur (Sudo 2012: 231)

a. Ahmet [(pro) kim-ni jaxshi kör-iman dep] bil-du?
Ahmet I who-ACC well see-IPFV.1SG.S that believe-IPFV.3.S
"Who does Ahmet$_i$ believe that he/I$_{i,*sp*}$ like(s)?"

*b. *Ahmet [(pro) kim-ni jaxshi kör-isen dep] bil-du?*
Ahmet you who-ACC well see-IPFV.2SG.S that believe-IPFV.3.S
("Who does Ahmet$_i$ believe that you$_{ad*}$ like?")

Nez Perce and Magahi are like Uyghur in this respect, except that indexical shift is optional. If it applies, 'I' can shift, but 'you' cannot be in the embedded clause, as in Uyghur. If it does not apply, then 'I' refers to Sp* and 'you' refers to Ad*. In other words, i-shift is ruled out in the context of 'you' in these languages/constructions (for Nez Perce, see Deal (2020: 94 (171)); for Magahi, see (102b)).

The important thing to grasp is that this defective addressee effect is fundamentally a kind of Shift Together, given a particular understanding of the goal role. According to Shift Together, the agent of a verb like 'think' is the same as Sp in the complement clause if and only if the goal of 'think' is the same as Ad. Since 'think' has no goal, Ad then has no reference, and 'you' inside its scope cannot be used to refer to anyone. As a result, it cannot be used at all. (This is more or less a transposition of Deal's way of handling defective addressees into my terms; see also Deal's (2020: 73) definition of the *addr* function.) I would not have foreseen that the absence of a goal is

taken to be the equivalent of there being a goal with no reference in this way. But apparently it is. What is clear, though, is that this is a funny kind of Shift Together effect, *not* a counter example to Shift Together. In its own way, this points to the robustness of Shift Together, since if a language is ever going to tolerate Shift Together violations, this seems like the perfect opportunity to do so.

The behavior of 'so that' adjunct clauses in Magahi also bears on this topic. I mentioned above that these adjunct clauses allow i-shift but not u-shift. That is true even when the matrix verb is one like 'speak', which does license a goal argument. So (104a) is possible, but (104b) is not with a shifted reading in which 'you' refers to 'Bantee'.

(104) Magahi (fieldwork, Deepak Alok)

a. Bantee-aa ghare rukl-ai taaki ham bimmar na ho jaa-i.
Bantee-FM home stay-3.NH.S so.that I sick not become go-1.S
"Bantee$_i$ stayed home so that he/I$_{i\ ,sp*}$ would not become sick."

b. Baabaa Bantee-aa-se batiail-thi taaki tu dukhii na ho.
grandfather Bantee-FM-INS speak-3.H.S so.that you sad NEG be
"Grandfather talked to Bantee$_i$ so that you$_{*i\ ,ad*}$ wouldn't be sad."

The example in (104b) contrasts with (105), a superficially very similar sentence with the C head *ki*, the complementizer normally found with CP complements. In (105), u-shift is possible.

(105) Magahi (fieldwork, Deepak Alok)

Baabaa Bantee-aa-se batiail-thi ki tu dukhii na ho.
grandfather Bantee-FM-INS speak-3.H.S that you sad NEG be
"Grandfather talked to Bantee$_i$ (saying) that he/you$_{*i\ (ad*)}$ should not be sad."

Given this contrast, it is tempting to say that *taaki* is a C head that selects Sp but not Ad, whereas *ki* licenses both Sp and Ad. Indeed, Alok and I thought this for a long time; see Alok (2020: 271-272). That would account for why 'you' cannot shift to the goal of the matrix verb 'speak' in (104b), assigning it a structure like (99b). However, this turns out to be the wrong analysis. One clear fact that it misses is that allocutive marking is possible in a *taaki* clause, as shown in (106).

(106) Magahi (fieldwork, Deepak Alok)

*Bantee-aa ghare ruk-l-o taaki (pro) bimmar na ho ja-i-**o.***

Bantee-FM home stay-PFV-H.AL so.that (I) sick NEG become go-1.S-H.AL

"Bantee$_i$ stayed home so that I$_{sp*,*i}$ do not get sick." (said to grandfather)

This shows that there must be an Ad in the 'so that' clause after all; otherwise, there would be nothing nearby for Fin in the adjunct clause to agree with in (106). So it is not that there is no Ad in the structure, but rather that Ad cannot be controlled by the goal of the matrix verb in this case; it has to be bound by Ad* in the root clause instead. The other reason for saying that *taaki* clauses include Ad as well as Sp is that they show a defective addressee effect, similar to the *ki* clause that appears with 'think' in (102b). To see this, suppose that the *taaki* clause contains both a first person pronoun and a second person pronoun, as in (107). Since Ad in a *taaki* clause can only refer to Ad*, Sp in the taaki clause can only refer to Sp*, not to the matrix clause subject, by Shift Together. The presence of a second person element in the embedded clause inhibits i-shift in that clause which would otherwise be possible, just as in (102b).[50]

(107) Magahi (fieldwork, Deepak Alok)

Baaba Bantee-aa-se bola-thi taaki (ham) tor samaachar jaan saki.

grandfather Bantee-FM-INS speak-3.H.S so.that (I) your news know can

"Grandfather$_i$ talked with Bantee$_k$ so that I$_{sp*,*i}$ can know your$_{ad*,*k}$ news."

The analysis that works for this wider range of facts is to say that the 'so that' clause has both Sp and Ad, but it is a quasi-argument (dependent) of the active Voice head in the matrix clause, not the complement of the verb per se. As such, it is presumably merged as an adjunct to VoiceP rather than inside VP. According to the standard view (Hale and Keyser 1993, Chomsky 1995, Kratzer 1996), Voice takes an agent argument but no goal argument; any goal role is

---

[50] Indeed, this is also true in (106), with allocutive marking playing the role of a second person element in the embedded clause.

assigned lower down, in the VP or ApplP complement of Voice. As such, active Voice can be considered as being essentially like 'think' and dyadic 'say', rather than being (in and of itself) a triadic verb like 'tell' or 'ask'. There is an Ad inside the 'so that' clause, visible for allocutive agreement, and it must shift together with Sp. But because of the special position of this type of clause, Shift Together only shows up in the form of a defective addressee effect, since active Voice never has a goal argument. This pattern thus testifies again to the robustness of Shift Together. It also provides an interesting new argument for decomposing verbs into distinct theta-marking heads, separating Voice from V and Appl. I provide more details for this analysis in §8.5.2.

Why then is *pro* different from the overt pronoun in Magahi in (102), in a way that seems to violate Shift Together? I return to this in Chapter 6, arguing that it is because *pro* in Magahi is really an indexiphor in the sense of Deal (2020). Anticipating the discussion there, the idea is that since the subject is null in (102a), we cannot observe directly what its features are. I claim that it is an indexiphor—a logophoric pronoun that triggers first person agreement on the verb—rather than a true first person indexical pronoun. Since logophors are bound by a different kind of ghostly DP operator (1lOp) than true indexicals are, there is no expectation that the control of 1lOp and that of Ad will be linked, the way that Sp and Ad are. On this view, (102a) is not a counterexample to the claim that indexical pronouns shift together because the subject here is not a true indexical; it only looks like one because it is null and triggers a particular kind of agreement on the verb. I then extend this analysis to the apparent violations of Shift Together in (100) from Slave as well. Like Magahi, Slave is also a pro-drop language of sorts, so the elements triggering first person agreement could be logophors rather than indexicals. The door is open, then to, say that Shift Together is really universal for true indexicals across languages and constructions, and that languages do not make use of the possibility of having C heads that license Sp but not Ad (or Ad but not Sp). (See Chapter 6 for more on the analysis of indexiphors across a range of languages.)

Why should this be? In fact, it does not seem to be connected to any of the deeper principles of my theory, and I present only a tentative conjecture. Recall that Sp and Ad can be arguments of two different kinds of C-type heads: sa and Fin. For sa, the fact that Sp and Ad always go together might be intrinsic to the meaning of sa, ultimately to the fundamentals of what speech acts are. They always have a

speaker of some sort, by definition. However, one might also say that they always have an addressee of some sort, also by definition. For many canonical speech acts, this is obvious. There are also some special kinds of speech acts, including ones where the speaker addresses themself (only), or where the speaker addresses nobody. But the claim would be that even such cases have a syntactically present Ad, even if it is somehow bound or negatively quantified over (as in the defective addressee cases discussed above). This same idea probably does not carry over to the embeddable instances of Sp and Ad in the specifiers of a head like Fin; Fin does not have the special substantive semantics of sa heads that would support this reasoning. However, if we continue to think of Sp and Ad in FinP as grammaticalized "echoes" of Sp* and Ad* in saP, perhaps they inherit a semantically bleached version of this property from their semantically more substantive kin. The result could be that natural languages rarely if ever have constructions in which Sp is licensed without Ad, or Ad is licensed without Sp. However, we should keep looking for possible cases of the (99) structure to confirm this.

6. Other languages, other theories

6.1. Introduction

In the early sections of this chapter, I shamelessly let Magahi data and my own theory guide the exposition. Inasmuch as the patterns of indexical shift in Magahi fall squarely in the center of how indexical shift has been found to behave in other relatively well-studied languages, this should do no serious damage to the topic, and it serves to foreground the new data that I have to offer. The focus began to broaden in the last section, as I considered issues about the optionality of indexical shift and the Shift Together phenomenon. Indeed, what degrees and patterns of optionality there are in indexical shift is perhaps the most important parameter of crosslinguistic variation that we know about beyond the basic question of why indexical shift is possible in some languages but not others. However, this is not all there is to say about the topic of crosslinguistic variation. I now round out the discussion by considering briefly other types of crosslinguistic variation in indexical shift, using Deal's (2020: Chapter 3) discussion as my outline. At the same time, I further compare my theory of indexical shift to Deal's (2020) version of the shifty operator theory. I

focus on hers because it is recent, comprehensive, influential, and explicitly strives to account for crosslinguistic variation in indexical shift (including her new Nez Perce data). As such, her theoretical goals are very similar to mine. My theory and hers also share the core idea that indexical shift is fundamentally caused by certain kinds of operators (functional heads or null DPs licensed by them) which are found only in the peripheries of certain kinds of clauses (Deal 2020: 45-48). For discussion of earlier theories of indexical shift and some criticism of them, I refer the reader to Deal's work, especially her Chapter 2, which for the most part I agree with, and whose empirical basis Alok and I replicated in Magahi.[51]

6.2. Variation in the verbs involved in indexical shift

The first type of variation that Deal discusses is variation in which matrix verbs allow indexical shift in their CP complements. Some languages allow this only with 'say'-class verbs, including Zazaki, Farsi, Kurmanji, Dhaasanac, and Somali. Others extend this to 'think'-class verbs (nonfactive cognition verbs) but not to 'know' class verbs (factive verbs). This set includes Navajo, Slave, Laz, Korean, and Japanese. Still others allow indexical shift even with 'know' class verbs—indeed with essentially all verbs that take finite CP complements. Nez Perce is such a language. So too is Magahi; see Alok (2020: §3.5.1). Examples with 'say', 'tell', and 'think' have been amply illustrated above. (108) adds an example with 'know'.

(108) Magahi (Alok 2020: 145 (26a))
Santee-aa jaana ge-l-o ki hamraa dillii jaa-yelaa he.
Santee-FM know go-PFV-H.AL that me.DAT Delhi go-INF be
"Santee$_i$ knew that he/I$_{i,sp*}$ has/have to go to Delhi." (said to grandfather)

---

[51] The exception is that my theory includes aspects of what Deal calls the Binding theory, which she attributes to von Stechow (2003). Her primary argument against such a view, following Anand (2006), is that it does not capture Shift Together effects. In my version, the crucial effects follow once the Person Licensing Constraint is included in the system. See §4.4 and Chapter 6 for discussion. (See also discussion of the *de re* blocking effect in those sections.)

Deal's account of this variation is essentially a selectional one: some verbs select larger clauses than others. Those that select larger clauses have room for the functional heads that constitute context-shifting operators in Deal's framework, whereas those that select smaller clauses may not. My framework allows for essentially the same approach. The only difference is that for Deal a verb may or may not select a complement that includes the functional heads which shift contexts, whereas in my account a verb may or may not select a complement that includes the functional heads that license the ghostly DPs Sp and Ad, which are the vehicles of indexical shift.

Considering this a bit further, what might set 'say' class verbs apart from others in some languages is that they can select the largest complements, saPs, which express speech acts. Then languages that have such verbs and that license Sp only in Spec sa1P and Ad only in Spec sa2P will only allow indexical shift under 'say'-class verbs. (See Miyagawa (2012) for this sort of reasoning applied to embedded allocutive marking in Japanese.) In contrast, languages like Magahi which license Sp and Ad in a lower projection (FinP) allow indexical shift in a much wider range of complement clauses.

In contrast, the resistance of complements of verbs like 'know' to indexical shift in some languages might well be related to the tendency of such verbs to have nominalized complements, where nominal projections can disrupt the obligatory control relation that indexical shift depends on, as discussed in §4.5.3 (see, for example, the long tradition of saying that factive verbs have complements with NP-over-CP structures). For example, Slave's verb 'know' selects a CP that is nominal in the sense of bearing (areal) gender and triggering object agreement on the verb, and this rules out indexical shift in the CP (see (97), (98) and discussion). Similarly, Deal (2020: 69) mentions Korean as a language in which the complement of 'know' is a nominalized clause, and hence does not allow for indexical shift. In contrast, the complement of 'know' in Magahi is not different in structure or morphology from the complement of 'think', and both allow indexical shift. These might be the primary "joints" in selectional phenomena across languages. There is also the possibility of more idiosyncratic selectional properties being stipulated for individual verbs as well, as in Slave, where Rice (1989: 1276) says that which verbs allow indexical shift and which do not is not semantically predictable.

6.3. Shift of locative and temporal indexicals

The second area of crosslinguistic variation that Deal discusses involves which indexicals shift in a given language. She considers four types of indexicals—first person, second person, locative, and temporal—whereas I discuss only the first two types in this work. Deal argues for an implicational hierarchy, such that locative indexicals shift only if first and second person indexicals shift, second person indexicals shift only if first person indexicals shift, and first person indexicals shift only if temporal indexicals shift. The only part of this hierarchy that falls within my purview is the relationship between first person indexical shift and second person indexical shift, and here I interpret the crosslinguistic evidence a bit differently from Deal, as discussed in §4.5.4. Her primary example of a language-construction in which i-shift happens but u-shift does not is Slave with the matrix verbs 'say' and 'want'. I proposed that this is really an instance of indexiphoricity, not true indexical shift—like Magahi sentences with a pro-dropped DP triggering first person agreement as opposed to sentences with an overt unambiguous first person pronoun. Therefore, I do not have an analog of Deal's parameter that the author-shifting functional head can be bundled together with the addressee-shifting head into a single unit, or it can appear by itself, depending on the language (although I could add this, if new empirical discoveries call for it). Rather, I always bundle the two together.

How do I justify offering an analysis of first and second person indexical shift without considering how the theory might apply to locative and temporal indexicals? This may seem like an artificially narrow study of the phenomenon to some. However, I have chosen to study i-shift and u-shift in comparison with phenomena like allocutivity and upward C-agreement on the one hand and logophoric pronouns on the other hand. These phenomena do not have any clear analogs for 'here'-shift or 'now'-shift that would give me leverage on those topics using my chosen method. Moreover, 'I', 'you', 'here' and 'now' may seem like a natural class from the perspective of indexicality, but from other perspectives they may not be. For example, the notion of grammatical person, a core feature for Agree and inflectional morphology, is relevant to 'I' and 'you' but not to 'here' and 'now'. Of course, it would be nice to understand everything all at once, but that is too big a task to do in one step. As we work toward that goal, it is practical, perhaps inevitable, to group subtopics in different ways and make different comparisons and see what sticks on our road to the grand unified theory of everything.

There are also empirical/typological reasons for taking a narrower approach to indexicals for now. We just do not have that much good data available on locative and temporal indexical shift yet, in my opinion. For example, Deal's conclusion that locative indexical shift implies first and second person indexical shift depends heavily on Nez Perce. That is the only language she discusses in which locative indexical shift is optional, first and second person indexical shift is optional, and whether the former happens depends on whether the latter happens.[52] Magahi facts do not necessarily fit smoothly with Deal's picture. Alok (2020) includes a very preliminary discussion of locative and temporal indexicals in Magahi. He shows that locative indexicals can shift along with first person in (109).

(109) Magahi (Alok 2020: 274)
Content: Santee is in his village talking to his friend about their friend Ram:
Ham JNU-me gel-i-au ha-l pichhlaa saal. UhaaN, ham Ram-se mil-i-au. U khak-it ha-l-ai ki ham aglaa saal yahan-se nikal jaayem.
I JNU-in went-1.S-NH.AL be-PFV last year there I Ram-INS met-1.S-NH.AL he tell-PROG be-PFV-3.NH.S that I next year here-INS pass go.FUT.1.S
"Last year, I went to JNU (a university in Delhi). I met Ram there. He_i was saying that he/I_i will pass out from here (=JNU) next year."

However, it seems also to be possible for 'here' to shift in Magahi without indexicals shifting. Alok (2020: 275 (39)) gives the example in (110) in the context of his discussion of shifted locative indexicals not necessarily being in direct quotations. 'Here' in the embedded clause refers to the location of Atul's speaking event (Delhi), but the subject 'he' of that clause refers to Atul. In this way, the shifted locative indexical behaves differently from a shifted second person pronoun, which prevents using a third person pronoun rather than a first person pronoun to refer to the matrix subject (see §4.4.3, as

---

[52] However, Deal does cite a reasonable number of languages in which locative indexicals shift together with first and second person ones, as well as a reasonable number of languages in which first and second person indexicals shift without locatives shifting. The crosslinguistic evidence for her asymmetry may thus be stronger than her evidence internal to particular languages.

shown in (71c) above, repeated here as (111).

(110) Magahi (Alok 2020: 274)
Jab hammni dillii-me ha-l-eN ta Atul kahk-ai hal ki [u ihaiN paidaa hol-ai hal].
when we Dehli-in be-PFV-2.S PRT Atul said-3.NH.S was that he here born happen-3.NH.S was
"When we were in Delhi, Atul$_i$ said that he$_i$ was born here (=in Dehli)."

(111) Magahi (fieldwork, Deepak Alok)
Santee-aa Bantee-aa-ke kah-l-ai ki u toraa dekh-l-ai.
Santee-FM Bantee-FM-DAT tell-PFV-3.NH.S that he you.NH.ACC see-PFV-3.NH.S
"Santee$_i$ told Bantee$_k$ that he$_{*i,n}$ saw you$_k$." (If 'you'=Bantee, then not 'he'=Santee.)

(110) implies that a locative indexical can shift without the clause being a domain of i-shift—in my terms, without Sp being controlled by *Atul*, which would force a pronoun referring to Atul to be first person, by Rule H plus the fact that Sp is first person. This is a seeming problem for Deal's hierarchy.[53] Similarly, Alok (2020: 274-276) shows that 'today' can get shifted readings in complement clauses in Magahi, but it cannot shift without 'I' shifting, whereas 'I' can shift without 'today' shifting. This is the opposite of Deal's hierarchy for temporal indexicals, based primarily on Korean facts from Park (2016). (Nez Perce does not have temporal indexicals to investigate.) Deal (2020: 78-79) also observes that there is more variation across languages in whether temporal adverbials count as genuine indexicals or not, whereas first and second person pronouns seem always to be true indexicals. As a result, the theoretical status of the relevant elements may be unclear in particular languages. Nor does it make much conceptual sense to me that temporal shifting operators should be the lowest in the clause structure whereas locative shifting operators are the highest. Deal does not try to motivate the functional hierarchy that she posits, but simply stipulates it. Finally, I

[53] It is also possible for 'here' to shift with or without 'I' shifting in Korean (Park 2016). For Deal, this apparent exception to her hierarchy is due to the locative shifting operator being bundled with the temporal-shifting operator in Korean.

find it odd that in her account a locative-shifting operator can bundle together with a temporal-shifting operator, as in her analysis of Korean, even though these heads are not adjacent in her underlying functional hierarchy. Overall, I am left with many questions about this domain, both empirical and theoretical, and have little certainty about what the robust patterns are. Much more work is needed in this area going forward.

As we learn more about this topic, there are (at least) three ways things could turn out. One is that it could turn out that DP pronoun indexicals and locative/temporal adverbial indexicals are not really the same kind of thing after all. Then there shouldn't be a (fully) unified analysis of them. My analysis would apply to DP pronoun indexicals, while some quite different theory applies to adverbial indexicals.

A second possible outcome is that we could learn that adverbial indexicals are the same kind of thing as DP-pronominal indexicals, such that they should have a unified analysis, and that my account can be generalized to locatives and/or temporal indexicals. It takes some effort to imagine this, but it does not seem impossible. It would involve contemplating a representation like (112) for a sentence like (110). The idea would be that there is a null XP "Loc" at the edge of a clause that denotes a location. In a root clause, Loc denotes the location of the speech act (Loc*). This is in line with my notion that the coordinates in a Kaplanian context can be syntactically represented in the periphery of the clause, extending this from Sp and Ad to Loc (and perhaps Time). Then an indexical like 'here' has to be bound by the closest Loc, just as 'I' must be bound by Sp and 'you' by Ad. So far, so good. Then the crucial step to get 'here'-shift in some languages is to say that Loc in an embedded complement clause can be controlled by an adverb like 'in Delhi' in the matrix clause, rather than simply being bound by Loc*. This is sketched in (112).

(112)[Loc*_k C [in Dehli_i Atul said [Loc_i C [he was born here_i.]]]].

The control of a locative element in this way is one step more remote from familiar instances of the control of PRO than the control of Ad and Sp are, but that may not be a bad thing. We know that the infinitival complements of some control verbs (so-called exhaustive control) have to take place at the same time as the matrix verb. 'Manage' differs from 'decide' in this way, as in (113) (Landau 2001, Wurmbrand 2003), etc.

(113) English (personal knowledge)
   a. Pat decided on Tuesday [PRO to go into NYC on Thursday].
   b. Pat managed on Tuesday [PRO to go into NYC (*on Thursday)].

Perhaps this means that a temporal element in the infinitival clause is obligatorily controlled by the time adverb in the matrix clause in (113b) but not in (113a). This distinction extends to location as well: the managing event needs to take place at the same location as the show-seeing event in (114b), whereas the deciding event need not take place at the same location as the show-seeing event in (114a).

(114) English (personal knowledge)
   a. Pat decided in Newark [PRO to see a show in NYC].
   b. Pat managed in Newark [PRO to see a shown (*in NYC)].

This generalization of my account to locative and temporal adverbs may or may not prove to be right, but I do not see it as a non-starter.

The third possible outcome of further inquiry into locative and temporal indexical shift is that it is the same phenomenon as first and second person indexical shift, but my account cannot be generalized to cover the locative and temporal cases. That will presumably mean that I have been wrong about the first and second person cases, having followed the wrong intuitions about what should be given a unified analysis with what. That would be a shame, but it happens. As my old bridge partner used to say: “You pay your money, and you take your chances.” (And it would not imply that the rest of this work is wrong.)

6.4. Other types of crosslinguistic variation

The third type of crosslinguistic variation in indexical shift according to Deal (2020) concerns whether it is optional or obligatory, and if so, with which verbs and which indexicals. For example, i-shift and u-shift are obligatory with certain matrix verbs in Matses, Laz, Navajo, Uyghur, Kobon, and Slave, whereas they are optional with certain matrix verbs in Zazaki, Nez Perce, Amharic, Korean, Japanese, and Magahi. I have already discussed this type of variation at length in §4.5. My idea is that optional nominalization and CP-extraposition plus optional reconstruction are two sources of apparent optionality in indexical shift. In languages where CPs do not extrapose and there is only one form of complementation, a non-nominalized kind, indexical

shift may be obligatory. Various minimal comparisons confirm that these are relevant factors in whether indexical shift appears to be obligatory or not. (The account may not be complete, however, and there is room to discover other factors that may be relevant to this.)

The last parameter of variation that Deal emphasizes is whether shifted indexicals must be interpreted as referring to their antecedents *de se* or not. Empirically speaking, Alok (2020: 168-171) shows that shifted first person indexicals in Magahi do need to be interpreted *de se*, with 'I' in the complement clause referring to the matrix subject only if the referent of the matrix subject is aware that the state of affairs expressed by the embedded clause holds of them. However, shifted second person indexicals in Magahi do not need to be interpreted *de se* (*de te*). In this respect, Magahi falls within the range of variation charted by Deal; it replicates the pattern documented for Uyghur by Sudo (2012). This is, however, a more purely semantic topic, and a tricky one. I have nothing to contribute to the theory of *de se* interpretation here. I assume that this is semi-independent of the syntactic issues. (See also Baker & Ikawa 2024 and Chapter 5 for some discussion of empirical controversies about *de se* interpretation in the case of logophoric pronouns.)

7. Conclusion

This chapter has argued that indexical shift is the result of the ghostly DP operators that trigger speaker and addressee agreement in some languages binding pronouns that match them in phi-features within the clause that they c-command. This accounts for the close relationship between allocutive marking and indexical shift in Magahi, and it easily generalizes to indexical shift in languages that do not show speaker/addressee agreement. I showed that the same principles of thematic-role-based obligatory control that are at work in upward C-agreement and allocutive marking are at work in this domain too. The parallels between upward C-agreement and first person indexical shift are particularly close, once the effects of the T/Agree Condition on C-agreement are factored out. The binding relationship between the ghostly DPs and first and second person pronouns is regulated by my Person Licensing Constraint. This requires that ordinary first and second person pronouns be bound by the closest Sp and Ad, and that uncontrolled Sp and Ad be bound by the next highest Sp and Ad, thereby capturing most of the Shift Together and No Intervening Binder/Local Determination effects discussed in previous literature. I went on to argue that Shift Together is more universal than thought,

and that it follows from the obligatoriness of obligatory control applying to both Sp and Ad. This led me into a discussion of what I take to be the main source of crosslinguistic variation in indexical shift: the question of whether it is obligatory, optional, or forbidden in a particular language and construction type. I attributed this to the possibility of nominalization and/or CP-extraposition bleeding obligatory control, depending on the language. Throughout the chapter, I compared my theory to the shifty operator theory of Anand (2006) and Deal (2020), claiming that the theories are based on a similar intuition and derive many of the same results in parallel ways, but there are a handful of specific respects in which my analysis has advantages. This is in addition to the fact that my approach embeds an analysis of indexical shift within a broad picture of how complementizers relate to the NPs around them to form a range of rare constructions scattered around the world.