

Chapter 1: Introduction: Rare “Wings” are Universal “Forelimbs”

Mark C. Baker
Rutgers University
Fall 2022

1.1 The challenge of crosslinguistically rare syntactic constructions

It has become clear as typology and generative linguistics have advanced that there exist crosslinguistically rare constructions. There are grammatical features that are salient aspects of one language but have no obvious analog in most other languages of the world, however exactly one may choose to individuate languages. For those who (like me) think that universal grammar (UG) is something real and important, this poses an interesting challenge. The key question is does UG specify the possibility and core features of the rare construction or not? On the one hand, if one thinks that it does not, then a significant feature of the language under study is apparently unconstrained by UG, and that is not a very satisfying view within the paradigm. In the limit, one could start to wonder what UG is for if it does not shape and guide the acquisition of important aspects of particular languages. On the other hand, if one thinks that UG does specify the possibility of the rare construction, then one needs to face the possibility that there is a chapter of UG that most languages do not read at all, it seems. Moreover, if it is not rare for a language to have a rare construction—because there are many rare constructions in the world’s languages taken as a whole—then the book of UG will contain many chapters that most languages read only one or two of. Although I admit that there are books on my shelves that are like that, some of which I am glad to have anyway, this is also not very satisfying as a view about UG.

I illustrate this conceptual issue with a series of examples, which introduce some the specific topics that this book is concerned with. Consider first switch-reference (SR). SR can be defined as a construction in which an affix or particle at the periphery of an embedded clause signals whether the subject of the embedded clause refers to the same entity as the subject of the main clause (same subject, SS) or not (different subject, DS). (1) is a reasonably typical pair of examples from Shipibo, a Panoan language spoken in the Amazonian area of Peru. Here the suffix *-ax* on the embedded verb signals SS, whereas *-tian* signals DS (see Valenzuela 2003, Baker and Camargo Souza 2020).

- (1) a. José=ra [_{<pro>} Rosa oin-**ax**] xobo-n ka-ke.
José=EV he Rosa see-SS.ABS house-LOC go-PFV
‘He_i seeing Rosa_j, José_i went home.’
- b. [José-kan Rosa oin-ke-**tian**]=ra, (ja) xobo-n ka-ke.
José-ERG Rosa see-PFV-DS=EV 3SG home-LOC go-PFV
‘When José_i saw Rosa_j, he/she_k (someone else) went home.’

The term “switch-reference” was coined by Jacobsen (1967), it evoked a flurry of interest in the 1970s and 1980s, and the first major generative study was Finer (1984, 1985). It is not a super rare phenomenon, as these things go. It is found in quite a few languages of Western North

America (70-some in McKenzie’s (2015) survey). It is also characteristic of substantial region of Australia (Austin 1981), of New Guinea, and it is found in a fair number of South American languages as well. But even if there are hundreds of languages that have SR, there are probably not thousands, and there seem to be more languages that do not have it than that do have it. Indeed, the phenomenon is absent from large areas of the world, including Europe, Africa, mainland Asia, and Eastern North America. This case is particularly interesting in that SR is generally a big deal in languages that have it. An average page of text in Shipibo may well have a dozen or more examples of SS/DS marking, and it is used in some complex, subtle, and expressive ways, seen only in complex sentences consisting of more than one clause. A theory of UG that does not say anything about SR would seem to be missing a key part of the genius of Shipibo. Furthermore, the literature on SR strongly suggests that its essentials are quite stable in interesting ways in the languages that have it. SR in Shipibo and Quechua is not that different from SR in Washo and Mojave in North America, or from SR in Diyari and Jiwari in Australia. If we looked only at those languages that have SR, it would seem to provide a great case for UG. Nevertheless, the majority of languages do not have an SR system, so that it does not feel at all “universal” in the ordinary intuitive sense.¹

Another example of a crosslinguistically rare construction is the use of dedicated logophoric pronouns. These are special pronouns that are used in an embedded clause to refer to an argument—usually the subject—of the matrix clause. (2a) is a reasonably typical example from Ibibio, a Niger-Congo language spoken in Nigeria. Here the pronoun *ímò* cannot be used in a matrix clause, and must refer to the matrix subject *Okon*. In contrast, the ordinary pronoun *anye* in (2b) can be used freely in matrix clauses, and in embedded clauses it can refer to the object of the matrix verb ‘tell’ or to someone else salient in the context, as well as to the matrix subject (just like pronouns in English).

- (2) a. Okon á-ké-dòkkó Edem ké Emem í-maá-ghá **ímò**.
 Okon 3.SG-PST-tell Edem that Emem 3.SG-like-NEG LOG
 ‘Okon_i told Edem_k that Emem does not like him_{i,*k,*n}.’
 ,i
- b. Okon á-ké-dòkkó Edem ké Emem í-maá-ghá **anye**.
 Okon 3.SG-PST-tell Edem that Emem 3.SG-like-NEG 3SG
 ‘Okon told Edem that Emem does not like him.’ 3sg=Edem, Okon, or other.

This term “logophoric pronoun” was coined by Hagege (1974), and the first landmark generative study was Clements (1975); Cully (1994) is a well-known typological overview. Such pronouns are a reasonably salient feature in the languages that have them. There will not be a dozen examples on most pages of a text, but there will be examples in an average short story. Crosslinguistically, this phenomenon has an even narrower distribution than SR. As far as is known, depending on the definitions, it may be found only in one region of the world, namely West Africa (Cully 1994). As a result, one might not be so tempted to use UG for a case like this. However, it is robust in the sense that it found in many languages in this region rich in languages, and it is attested in languages from different families (e.g., languages from the Chadic branch of Afroasiatic as well as Niger Congo languages) and with different typological

¹ Although there is of course no contradiction with what Chomsky has always meant by Universal grammar, which is the innate knowledge that a child brings to bear on learning their native language—not a surface feature claimed to be present in all languages.

characters (Yoruba is isolating, whereas Ibibio is quite agglutinative). So it is not very satisfying to just say that logophoric pronouns are just some idiosyncratic ornament to core language that happened to develop only once or twice because of special historical circumstances. It has also long been thought that logophoric pronouns are similar to the more widespread phenomenon of long distance anaphors in (for example) the languages of Europe and East Asia (Clements 1975, Sells 1987, etc.), a hypothesis that we will have reason to consider in some detail.

A third rare construction is upward complementizer agreement. This can be defined as a construction in which the complementizer of the clausal complement of a verb agrees in person/number/gender features with another argument of that verb—almost always in fact the subject. (3) exemplifies this with a typical pair from the Bantu language Kinande spoken in the Eastern Congo.

- (3) a. Kám bale mw-a-kabw-ir-a a-bá-kalí **a-ti** Maryá
 CL1.Kambale AFF-CL1.TNS-told-APPL-FV CL2-CL2-women CL1-that CL1.Mary
 mw-á-gúl-ir-é e-hí-lole. (Kinande)
 AFF-CL1.TNS-buy-ASP-FV CL19-CL19-bananas
 ‘Kambale told the women that Mary bought bananas.’
- b. A-ba-kali mo-ba-kabw-ir-a Kambale **ba-ti** Maryá
 CL2-CL2-women AFF-CL2.TNS-told-APPL-FV CL1.Kambale CL2-that CL1.Mary
 mw-á-gúl-ir-é e-hí-lole. (Kinande)
 AFF-CL1.TNS-buy-ASP-FV CL19-CL19-bananas
 ‘The women told Kambale that Mary bought bananas.’

This is mentioned in Baker (2008), among other places, and the first full-fledged generative study is Diercks (2010, 2013). This construction is not as significant an overall design feature for languages that have it as SR is for Shipibo, or even as logophoricity is for Ibibio. However, it is arguably a special case of what is a general design feature for the Bantu languages: the fact that they are particularly rich in agreement, extending it to a wide variety of functional heads. Like logophoricity, upward C agreement may only be found in one area of the world—a band of languages from Nigeria and Angola in the west to Kenya in the east. There are languages that are very rich in agreement in other parts of the world, but they do not have this kind of agreement.² Another interesting feature of this case is that several of the languages seem to have developed C agreement independently: the agreeing C comes from ‘say’ in Kinande but from ‘be’ in Lubukusu; Kipsigis is a Nilo-Saharan language, not a Bantu language, and Angolan languages have developed their agreeing Cs from pronouns rather than verbs (Kawasha 2007). But despite these historical differences, this range of languages seem to have developed essentially the same construction. That is the sort of thing that could be explained by having UG shape the space of what an agreeing C-like particle could be like despite superficial differences in the input. But should UG cover something that only languages in one area of the world do?

Other constructions that I consider in this work have other kinds of distributions across languages. A rather well-discussed construction that I include is indexical shift. This can be

² Mohawk is a good example of a language from another part of the world in which agreement is ubiquitous on all types of lexical categories, but Cs do not bear agreement. In the 100+ language survey I undertook in Baker (2008), the only language from another region that might have upward C-agreement was the Arapesh language of New Guinea, and descriptions of that are not very complete (e.g. as to whether the agreeing head really is a complementizer rather than say a serial verb).

characterized as a construction in which first person pronouns inside the CP complements of certain verbs can be interpreted as referring to an argument of the selecting verb—usually its subject. (4) is an example from Magahi, an Indo-Aryan language spoken in Eastern India (Alok & Baker 2018, Alok 2020).

- (4) a. Santee-aa soch h-ai ki (**ham**) tej h-i.
 Santee-FM think be-3.NH.S that (I) smart be-1.S
 ‘Santee_i thinks that I_{i,sp*} am smart.’
- b. Santee-aa Bantee-aa-ke kahl-ai ki (**ham**) tej h-i.
 Santee-FM Bantee-FM-DAT tell:PFV-3.NH.S that (I) smart be-1.S
 ‘Santee_i told Bantee_k that I_{i,*j,sp*} am smart.’ (I=Santee or I=Sp*)

Indexical shift was brought to the attention of generative linguists as something distinct from direct quotation by Schlenker (1999, 2003); see Deal (2020) for a thorough recent overview. Its crosslinguistic distribution is not yet clear, as the tests to distinguish it from direct quotation need to be more widely applied. It is turning up in more languages now that we know to look for it, including a wide range of Turkic languages, Zazaki and Kurmanji, Amharic, Nez Perce, and Matsigenka, as well as Magahi. Still there are probably fewer languages that have this phenomenon than that do not; it is not known in the languages of Europe, and it is not possible in Niger-Congo languages like Ibibio, Kinande, and Lubukusu. Moreover, this seems not to be as strongly an areal phenomenon as logophoric pronouns and upward C agreement are. Rather indexical shift languages seem to be sprinkled throughout the world—although this picture could change some given the challenges of distinguishing indexical shift from direct quotation.

The last of the five core constructions that I focus on in this work is allocutive agreement. This can be characterized as a verb bearing some kind of agreement, not (only) with its subject or object, but with features of the person that the sentence is addressed to. The paradigm case is Basque, with Oyharcabal (1993) as an early generative landmark. (5) is an example, where the auxiliary verb is different depending on whether the sentence is addressed to someone with whom one has a formal relationship (5a) or a male person with whom one has a close relationship (5b), or a female person with whom one has a close relationship (5c).

- (5) a. Pette-k lan egin di-**zi**.
 Peter-ERG work do AUX.3.ERG-2SG.H.AL
 ‘Peter worked.’ (to a person with a distant, formal relationship))
- b. Pette-k lan egin di-**k**.
 Peter-ERG work do AUX.3.ERG-2SG.M.AL
 ‘Peter worked.’ (to a close male)
- c. Pette-k lan egin di-**n**.
 Peter-ERG work do AUX.3.ERG-2SG.F.AL
 ‘Peter worked.’ (to a close female)

This is a relatively isolated phenomenon in Europe (although see Alok & Haddican to appear on possible allocutivity in Gallician). It has recently been turning up in a range of South Asian

languages as well, both Dravidian and Indo-Aryan. Miyagawa (2012, 2017) argues that it is present in Japanese too, although it is less clear that the relevant morphology is agreement in this language. Antonov's (2015) typological study also lists Pumé (Venezuelan), Nambikwara (Brazilian), Mandan (Siouan), and Baja (Cushitic). This gives the impression of a phenomenon very thinly sprinkled through the languages of the world. As always, it is going to make a difference here how one defines the phenomenon, since a larger set of languages might have politeness markings of various kinds (e.g. Korean) that could be akin to allocutive marking. But as a coherent phenomenon of allocutive *agreement*, it seems very likely that only a small percentage of languages of the world have it.

An admittedly strong/extreme/provocative view that helps underline what is at stake here is the so-called (strong) uniformity hypothesis, sometimes taken to be part of Chomsky's Minimalist Program (Chomsky 2001: 2, Miyagawa 2010, Sigurdsson xxx, etc). This view hypothesizes that there are essentially no syntactic parameters, but only a universal grammar plus morphological variation at PF. According to this view, all languages have essentially the same syntax. In some very important domains, that may be true, for example when it comes to the basic stock of categories (especially lexical categories; see Baker 2003) and the principles by which they are merged together to make larger phrases. Similarly, it may not be implausible to think that all languages have *wh*-movement of one form or another (Huang 1982, Stowell xxx, many others). One can also make a case for simple case marking and agreement being universal, covert if not overt, as in Vergnaud's famous GB-era hypothesis (Chomsky 1981, Vergnaud xxxx, although I personally being led away from this; see, for example, Diercks 2012, Baker 2015). But whatever one thinks of these classic cases, the strong uniformity hypothesis begins to strain not only belief but even the imagination when it comes to the sorts of rarer constructions that we have sampled from here. Taken literally, it would amount to saying that all languages have a rich SR system like Shipibo, along with a system of logophoric pronouns like Ibibio, along with upward C-agreement like Lubukusu—not to mention things like noun incorporation and serial verb constructions. This begins to make one's head spin to think about, and more and more so as one exposes oneself to more of the rare constructions that are attested in languages of the world.

1.2 An analogy from comparative anatomy

To shed some light on how we might think about the paradox that rare syntactic constructions create for notions of UG, consider an analogy taken from what we know about biology. The analogy focuses on the comparative anatomy of mammals. I believe that it is true that most mammals do not have wings. I, for example, do not (sadly). In fact, only bats have them. Furthermore, most mammals do not have flippers for swimming. But whales and seals do. Arms are also a rare thing for mammals to have. But we humans have them (compensating somewhat for the lack of wings). Finally, horses have front legs, whereas these other types of mammals do not. So there are rare anatomical features of mammals, much as there are rare syntactic constructions in natural language.

However, *forelimbs* are not at all a rare thing for mammals to have. In fact, *all* of them have them: bats, seals, whales, humans, horses, mice, and so on. Furthermore, we are told that mammal forearms all have essentially the same syntactic structure. By this I mean that they have the same number of bones, which are connected to one another in the same ways. This common syntactic/skeletal structure then takes on notably different functions in different animals. There is a theory of “universal mammal” (UM) skeletons that applies just as well to bats and seals as to

humans and horses. The wings of a bat are not counterexamples to this theory of UM, nor do they fall outside its scope. On the contrary, they are good exemplars of the theory, properly understood.

It could be that something similar holds for certain kinds of rare syntactic constructions in natural language. Suppose that the same syntactic “skeleton” underlies all five of the phenomena I inventoried in section 1.1, as well as a few more. Upward complementizer agreement is undeniably a rare feature of language, and so is switch-reference. But now suppose that they are both manifestations of a structure in which a complementizer licenses a null nominal which is controlled by the matrix subject and with which the complementizer enters into relationship of Agree. (All this gets unpacked below.) Then that underlying structure is not as rare or geographically limited as the two surface constructions are when considered separately. Indeed, if there are other surface realizations of this underlying structure too, it might even be universal in a meaningful sense. This is the highest level thesis of this book.³

1.3 Some motivating surfacy similarities

Of course, a unifying view such as this one is more plausible if the phenomena being unified have something readily observable in common, even before we start “dissecting” them in earnest. Mammalian forelimbs do have such similarities. Whether wings or flippers or front feet or arms, there are two of them, and they are symmetrically placed protruding from the upper/front torso of the animal. The constructions that I have listed here are arguably also in the same ballpark when viewed in the right way.

For starters, all five of the constructions listed in section 1.1 are what I like to call “funny things complementizers do to relate to noun phrases around them.” This description holds most obviously of upward C agreement seen in (3), in that one funny thing that Cs can do to relate to NPs around them is agree with them. It also holds fairly straightforwardly for switch-reference constructions like in (1) once one recognizes that that the outermost affix on the verb in a head-final languages like Shipibo (and most other languages with SR) is often the realization of a C-like head. Then SS and DS markers can be taken to be complementizers that relate funnily to the subject NPs just above and below them by indicating whether they are coreferential or not. Allocutive marking also falls under this description if one believes, following Speas and Tenny (2003), that the person to which a sentence is addressed can be represented syntactically by a null pronoun in the periphery of the sentence (see also Oyharcabal 1993 for an earlier although less general version). Then allocutive marking can be seen as the result of C relating to this special NP near to it by agreeing with it.

Logophoric pronouns and indexical shift are less obvious cases of complementizers relating the NPs around them in funny ways. After all, no complementizer or clause-peripheral morpheme was mentioned in the preliminary characterization of those two constructions. But the notion of a complement clause was referred to in the characterizations of these constructions, and complement clauses have complementizers. Indeed, it turns out that in languages like Ibibio which complementizer is present partially determines whether logophoric pronouns are possible

³ Note however that my claim is that the underlying structure is relatively common and widespread, not necessarily that it is clearly attested in every language.

in the complement clause or not.⁴ For example, the normal declarative complementizer *ke* does allow a logophoric pronoun inside the complement clause to be used to refer to the matrix subject, as in (6a), but the complementizer *naña* used with perception complements does not, as in (6b).

- (6) a. Okon a-ma-a-kit **ke** Emem a-ma-a-yip ebot ímò.
 Okon 3.SG-PST-3.SG-see that Emem 3.SG-PST-3.SG-steal goat LOG
 ‘Okon_i saw that Emem stole his_i goat.’
- b. Okon a-ma-a-kit **naña** Emem a-yip ebot ọmọ/*ímọ.
 Okon 3.SG-PST-3.SG-see how Emem 3.SG-steal goat his/*LOG
 ‘Okon_i saw Emem steal(ing) his_i goat.’

This then does qualify as another funny thing that complementizers do to relate to the NPs around them: the complementizer *ke* in Ibibo enables a special pronoun in its complement to refer to the subject just above *ke*. The situation with indexical shift is similar; for example, in Mishar Tatar null first person pronouns can refer to the matrix subject (which need not be first person) when the embedded clause is headed by the verbal complementizer *diep* (historically related to the verb ‘say’), as in (7a), but not with other forms of complementation, as in (7b) (Podobryaev 2014; see also Shklovsky and Sudo 2014 on related Uyghur).

- (7) a. Alsu [pro kaja kit-te-m **diep**] at’-tx? (p. 84)
 Alsu where go.out-PST-1SG that say-PST.3SG
 ‘Which place did Alsu_i say that I_{i,sp*} went?’
- b. Marat Alsu-ga [(minem) kil-gän-em-ne] at’-tx. (p. 89)
 Marat Alsu-DAT I.GEN come-NMLZ-1.SG-ACC say-PST.3SG
 ‘Marat_i said to Alsu that I_{*i,sp*} came.’

It is not outlandish, then, to think that these five constructions might have something to do with each other, just as it is not outlandish to think that the forelimbs of different kinds of mammals might be theoretically related.

A more specific property that the five constructions have in common that encourages the idea that they are related is that they center on CPs in complement positions—not CPs in other kinds of positions, such as relative clauses. For example, one can have an agreeing C in addition to an ordinary C in a complement clause in Ibibio ((8b)) but not in a relative clause ((8a)).

- (8) a. Okon a-sak a-yem [ngwet odo [(**a-bo/*a-te*) se [(ami) ng-k-i-ño Eno]].
 Okon 3.SG-PROG 3.SG-see book the 3.SG-C /3.SG-C REL I 1.SG-PST-1.SG-give Eno
 ‘Okon is looking for the book that I gave to Eno.’
- b. Okon á-ké-n-dòkkó [(a-bo/a-te) ké [Emem í-máá-ghá Eno]].
 Okon 3.SG-PST-1.SG.O-tell 3.SG-C/3.SG-C that Emen 3.SG-like-NEG Eno

⁴ The connection between a particular complementizer and the possibility of logophoric pronouns is apparently even closer in Ewe, according to Clements (1975: 157): logophoric pronouns are possible only inside a CP headed by the complementizer *be*. See chapter 4 for discussion.

‘Okon told me that Emem does not like Eno.’

Similarly, logophoric pronouns which refer to the closest c-commanding subject are licensed in complement clauses in Ibibio, but not normally in relative clauses, as shown in (9).

- (9) a. *Okon a-ma-a-duok ngwet odo se imo i-k-i-dep.
 Okon 3.SG-PST-3.SG-lose book the REL LOG 3.LOG-PST-3.LOG-buy
 (‘Okon_i lost the book that he_i bought.’) (Ok with ... *anye a-ke-dep* ‘he 3.SG-PST-buy’)
- b. Okon a-ma-a-dokko ayin omò ke imò i-ma-i-yip ngwet.
 Okon 3.SG-PST-3.SG-tell son his that LOG 3.LOG-PST-3.LOG-steal book
 ‘Okon_i told his_{i,k} son that he_i stole the book.’

In the same way, shifted ‘I’ can refer to the matrix subject when it is in a complement clause in Magahi, but not when it is in a relative clause, as shown in (10). (This is shown for two different forms of relative clause in Magahi in (10b,c).)

- (10) a. Santee kalpanaa kark-ai [ki ego sudar laiki hamraa-se biah kart-ai].
 Santee imagine do-3.NH that one-CL beautiful girl me-INS marriage do-3.NH
 ‘Santee_i imagines that a beautiful girl will marry me_{i,sp}*.’
- b. Santee ego sudar laiki-ke baare-me kalpanaa kark-ai je hamraa-se biah kart-ai.
 Santee one-CL beautiful girl-GEN about-LOC imagine do-3.NH REL me-INS marry do-3.NH
 ‘Santee_i imagined (thought) about a beautiful girl who will marry me_{i,sp}*.’
- c. Santee, ego sudar laiki je hamraa-se biah kart-ai, okra baare- kalpanaa kark-ai.
 Santee one beautiful girl REL me-INS marry do-3.NH her about-LOC imagine do-3.NH
 ‘Santee_i imagined (thought) about a beautiful girl who will marry me_{i,sp}*.’

As for switch-reference, distinctive same subject marking is allowed on a small range of complement clauses in Shipibo, as in (11). However, it is never allowed on relative clauses. Hence no SS marker like *-ax* or *-i/kin* is seen in (12), but only the invariant nonfinite affixes *-a* (perfective), *-ai* (imperfective), and *-ti* (infinitive) (Valenzuela 2003). (In (12b) the understood subject of both the matrix clause and the relative clause are the same; in (12a) the two subjects are different.)

- (11) Maria-nin=ra [kenti be-**kin**] peo-ke.
 Maria-ERG=EV pot bring-SS.ERG start.TR-PFV
 Maria started to bring the pots.
- (12) a. Jain-xon-ribi=ra a-kan-ti iki, [kawin [jan-n yawa raka-n-**ti**]]....
 there-ERG-also=EV do.TR-3.PL-INFAUX rush.mat 3-LOC peccary.ABS lying-TR-INF
 ‘Also they have to make rush mats [(for people) to place the peccaries on].’ (PV: 477)
- b. [Ja [wexa-anan-**a**] joni rabé] raká-kan-ai westiora oxé. (PV: 471)
 that cut-RECP-PFV.PTCP man two lying-3.PL-IPFV one moon
 ‘The two men [who had cut each other] lay for one month.’

Finally, allocutive marking is allowed on relative clauses in Magahi, as in all finite clauses. However, *shifted* allocutive marking, where the agreement on the embedded verb reflects the relationship of the agent of the matrix clause to the goal of the matrix clause rather than that of the speaker to the addressee, is not allowed in relative clauses ((13a)). In contrast, this sort of shifted allocutive marking is allowed in CP complements in Magahi ((13b)).

- (13) a. Santee-aa baabaa-ke khabar kahk-au je Ram okraa kahk-**au/*o**.
 Santee-FM grandfather-DAT word tell-3.NH.NH.AL REL Ram him.ACC tell-NH.AL/*H.AL
 ‘Santee_i told grandfather the message that Ram told him_i.’ (said to a peer)
 (Bad is honorific /o/ reflecting Santee’s relationship to his grandfather.)
- b. Santee-aa baabaa-ke kahk-au ki Ram Sita-ke dekh-au/o hal.
 Santee-FM grandfather-DAT told-3.NH.NH.AL that Ram Sita-ACC saw-NH.AL/H.AL was
 ‘Santee told grandfather that Ram saw Sita.’ (said to a peer)
 (OK is honorific /o/ reflecting Santee’s relationship to his grandfather.)

So all five constructions show a parallel contrast between complement clauses, which are directly related to the main verb, and relative clauses that modify the object, which are only indirectly related to the main verb.

Another specific property that these five constructions have in common is that they all take place in fully finite CP complements, not nominalized clauses. For indexical shift, this has been pointed out by Shklovsky and Sudo (2014) for Uyghur (see also Mishar Tatar in (7) above); it is also true for Magahi. Note that in (14b) the complement of ‘say’ not only includes the nominalizing affix *-lik*, but it also bears accusative case and its subject bears genitive case—all signs of nominalization, and all ways in which (14b) is different from (14a). Correlated with this, (14b) does not allow a shifted indexical reading of ‘I’, whereas (14a) does.

- (14) a. Ahmet [men ket-tim] di-di (S&S 383 (4b))
 Ahmet 1.SG leave-PST.1SG say-PST.3
 ‘Ahmet_i said that I_{i,*sp*} left.’
- b. Ahmet [mening kit-ken-lik-im-ni] di-di.
 Ahmet 1.SG.GEN leave-REL-NMLZ-1SG-ACC say-PST.3
 ‘Ahmet_i said that I_{*i,*sp*} left.’

Similarly, in Ibibio logophoric pronouns are possible in CP complements like (15b) but not in comparable true derived nominalizations like (15a). Note that the “subject” of the derived nominal in (15a) follows the nominalized verb, as possessors of nouns do in this language, whereas the subject of the clause in (15b) precedes the verb in accordance with the language’s basic SVO word order.⁵

⁵ However, logophoric pronouns are possible in gerundival complements in Ibibio, which have a blend of nominal and verbal properties; see chapter 5. How nominal is too nominal for these rare constructions is a question I do not go into yet.

- (15) a. Okon i-kit-te [n-dudue eka ɔmɔ/*imɔ].
Okon 3.SG-see-NEG NMLZ-commit.fault mother his/*LOG
'Okon_i does not see his_i mother's mistake/fault.'
- b. Okon i-kit-te [ke eka imɔ a-ma-a-due].
Okon 3.SG-see-NEG that mother LOG 3.SG-PST-3.SG-commit.fault
'Okon_i does not see that his_i mother committed a fault.'

In addition, here is no possibility of adding an upward agreeing C to a nominalized/gerund-like constituent in Ibibio, whereas these can be added to normal CPs. This contrast is shown in (16).

- (16) a. Okon a-ma a-tre [(a-bo) u-koot ngwet].
Okon 3.SG-PST-3.SG-stop 3.SG-C GER-read book
'Okon stopped reading the book.'
- b. Okon á-ké-n-dòkkó [(a-bo/a-te) ké [Emem í-máá-ghá Eno]].
Okon 3.SG-PST-1.SG.O-tell 3.SG-C/3.SG-C that Emen 3.SG-like-NEG Eno
'Okon told me that Emem does not like Eno.'

Similarly, (17) shows that there is no allocutive marking on nominalized (infinitive and participial) complements in Magahi, analogous to what can be seen in complement clauses in examples like (13b) (Alok 2021). (In this case the allocutive marking on the complement clause would presumably resume the marking that is possible on the matrix verb—one of the two options attested in (13b)—since there is no goal argument of the matrix verb to support a shifted version of allocutive marking.)

- (17) a. Santeeaa [jaa-yel-*au/*o/*ain chaha h-au/o/ain (Infinitive)
Santee go-INF-NH.AL/H.AL/HH.AL want be.3.NH-NH.AL/H.AL/HH.AL
'Santee wants to go.'
- b. Ham okaraa dhekke-se bach-l-i-ain/o/au. (participle)
I 3.SG.NH.ACC see.PTCP-INS avoid-PFV-1.SG-HH.AL/H.AL/NH.AL
'I avoided seeing him.'

Finally for switch-reference, Imbabura Quechua has SR markers on subjunctive complement clauses, but not on nominalized ones, as shown in (18). (18a) has a nominalized complement of the verb 'want' as shown by the fact that the complement as a whole is marked for accusative case whereas its object does not need to be marked accusative (Cole 1982; Herman 1985; Cole and Hermon 2011). This version does not have SR marking. In contrast, (18b) has a nonnominalized subjunctive complement, as shown by the fact that the complement as a whole does not bear accusative case and the object inside it must be accusative. This version is marked for SR; in (18a) the form is SS *-ngapaj* rather than DS *-chun* because the wanter is the same as the desired seer (see (20b) for a DS analog). There is a systematic complementarity in this, in that no clause is marked for both SR and morphological case in Quechua.

- (18) a. [Aycha-(ta) miku-na-ta] muna-ni. (infinitive)
meat-(ACC) eat-NMLZ-ACC want-1.S

‘I want to eat meat.’ (Hermon 1985:25)

- b. Muna-y-man [ñuka mama-ta riku-ngapaj]. (SS subjunctive, IQ)
 want-1.S-COND my mother-ACC see-SBJV.SS
 ‘I want that I see my mother; I want to see my mother.’ (Cole 1982: 37)

So all five phenomena are things that happen in nonnominal CPs rather than in nominalized constituents.

One further important thing that four of the five constructions have in common is that they are subject-oriented rather than object oriented in ways that may seem surprising from a theoretical perspective. For example, when the verb selecting CP with an agreeing C has an (indirect) object as well as a subject, the C agrees with the subject, not the object, even though the object seems to be structurally closer to C. This is shown in (19) from Lubukusu (see also (3) from Kinande). Previous literature on upward C-agreement like Diercks (2013) and Carstens (2016) wrestles with how to explain this.

- (19) a. **Ba-ba-ndu** ba-bol-el-a Alfredi **ba-li** a-kha-khil-e
 CL2-CL2-people CL2.S-say-APPL-FV CL1.Alfred CL2-that CL1.S-FUT-conquer-FV
 ‘The people told Alfred that he will win..’
- b. **Alfredi** ka-bol-el-a ba-ba-ndu **a-li** ba-kha-khil-e.
 CL1.Alfred CL1.S-say-APPL-FV CL2-CL2-people CL1-that CL2.S-FUT-conquer-FV
 ‘Alfred told the people that they will win.’

Similarly, SR marking on a complement clause shows whether the embedded subject is coreferential with the subject of the matrix clause, not with the object of the matrix clause, even though the object seems to be closer to the SR head in the complement clause. This can be seen in (20) from Imbabura Quechua: (20a) in which the lower subject is coreferential with the matrix subject has the proximate/SS marker *ngapaj*, but (20b) in which the lower subject is coreferential with the matrix object does not; it must have the obviate/DS marker *-chun* instead. This is a disanalogy between SS-marked clauses and controlled infinitives that Hermon (1985: 122-124) struggles with.

- (20) a. Muna-y-man [ñuka mama-ta riku-ngapaj]. (SS subjunctive)
 want-1.S-COND my mother-ACC see-SBJV.SS
 ‘I want that I see my mother; I want to see my mother.’ (Cole 1982: 37)
- b. Juan-da kunvinsi-rka-ni [(pay) Kitu-man ri-**chun**]. (DS subjunctive)
 Juan-ACC persuade-PST-1.S he.NOM Quito-to go-SBJV.DS
 ‘I persuaded Juan that he go to Quito.’ (not **ri-ngapaj*, Hermon 1985:123)

One can also envision a connection here with the fact that logophoric pronouns are subject-oriented in that they can refer to the subject of the matrix clause, but not to the object of the matrix clause, as seen in (21) from Ibibio. Note that ordinary third person pronouns in languages like English are not required to refer to a subject in this way.

- (21) Okon á-ké-dòkkó Edem ké Emem í-maá-ghá ímò. (Ibibio).
 Okon 3.SG-PST-tell Edem that Emem 3.SG-like-NEG LOG
 ‘Okon_i told Edem_k that Emem does not like him_{i,k}.’

In the same way, a first person indexical pronoun ‘I’ in Magahi can refer to the matrix subject of a verb like ‘tell’, but not to the matrix goal.

- (22) Santee-aa Banteeaa-ke kahl-ai ki ham tej h-i.
 Santee-FM Bantee-FM-DAT told-3.NH that I intelligent be-1
 ‘Santee_i told Bantee_k that I_{i,*k,sp*} am intelligent.’

The odd construction out for this last similarity is allocutive marking. If anything, this seems to be object-oriented (anti-subject-oriented) rather than subject-oriented, in that allocutive marking on the embedded verb can change when the status of the person referred to by the indirect object changes, as shown in (23).

- (23) a. Santee-aa **Bantee-aa-ke** kahk-ai ki Ram-ke Sita-se baat kareke chah-**au**.
 Santee-FM Bantee-FM-DAT told-3.NH that Ram-DAT Sita-INS talk do.INF should-NH.AL
 ‘Santee told Bantee that Ram should talk to Sita.’ (or *kahk-ain*, HH.AL, to a teacher).’
- b. Santee-aa **baabaa-ke** kahk-ai ki Ram Sita-ke dekh-l-**o** ha-l.
 Santee-FM grandfather-DAT told-3.NH that Ram Sita-ACC saw-H.AL be-PFV
 ‘Santee told grandfather that Ram saw Sita.’ (or *kahk-au*, NH.AL, said to a peer)
- c. Santee-aa **profesar saaheb-kekahk-ai** ki Ram Sita-ke dekh-l-**ain** ha-l.
 Santee-FM professor HH-DAT told-3.NH that Ram Sita-ACC saw-HH.AL be-PFV
 ‘Santee told the professor that Ram saw Sita.’ (said to a peer)

So four of the five “funny things that Cs do to relate to the NPs around it” are subject-oriented in ways that may not be theoretically expected, whereas the fifth may be anti-subject-oriented.

There are, then, some readily accessible similarities between the five constructions that render it not-implausible that they would have a partially unified analysis—that they might be different functions or adaptations of the same underlying structural “skeleton”.

1.4 Overview of the analytical framework

I now sketch upfront my high-level hypothesis about what that the structural skeleton consists of, which is the unifying threads that tie together the different topics discussed in this work.

The first crucial assumption is that some Cs (or better C-like heads in an exploded left periphery, after Rizzi 1997) license a pronoun-like DP in the periphery of the clause headed by C. These DPs are rather hard to detect by conventional means in that they are (a) obligatorily phonologically null, (b) have minimal lexical semantic content, and (c) do not create islands for the extraction of material out of the CP (in contrast, say, to *wh*-phrases in the CP periphery, which can create *wh*-islands). Because they are hard to detect by these means, I sometimes refer to them as “ghostly DPs”, helping give the discussion a bit of color. I also refer to them as operators, having in mind null operators in the syntactic sense, similar to the null DPs that

undergo *wh*-movement in post-Chomsky (1977) analyses of many constructions (although the ghostly DPs posited here are base-generated in the CP periphery, rather than arriving there by movement). All five constructions involve such ghostly DPs in the C-space, I claim. This is at the root of why the five constructions are possible in CP complements but not in derived nominals or nominalized clauses: nominal constituents do not have the C-like heads that license these ghostly DPs. At the same time, some of the differences between the constructions can be attributed to exactly which head in the C-space licenses the ghostly DP, and to exactly what formal features the ghostly DP has. For example, the ghostly DPs involved in indexical shift and allocutive marking have first and second person features; the ghostly DPs involved in logophoric constructions have a language-particular [+log] feature, the ghostly DPs involved in so-called indexiphor constructions (Deal 2020) have both first person and [+log] features, and the ghostly DPs involved in simple upward C-agreement and switch-reference have no intrinsic features of their own.

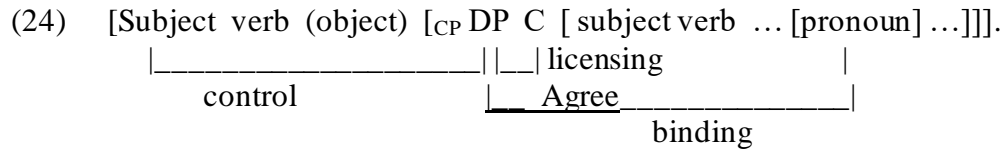
The second crucial assumption is that any of these ghostly operators may, and in many cases must, undergo obligatory control, such that they are bound by a designated argument of the verb that selects the CP as its complement. As such, the ghostly operators are somewhat analogous to PRO, the necessarily null DP that is licensed by infinitival Tense in English and plenty of other languages (see Landau 2013 for an overview). This assumption helps to explain why the five phenomena happen in complement clauses but not (for example) in relative clauses: the extra nominal projections that intervene between CP and the matrix verb in a relative clause construction block obligatory control, causing some of the ghostly operators to crash (those that do not have interpretable features of their own) and others to show different behavior as to their possible antecedents (as in the classical distinction between obligatory and nonobligatory control). I also claim that the involvement of obligatory control is what causes four of the five constructions to be subject-oriented in the sense outlined in (19)-(22): the subject and not the object can control the ghostly DP operator because of a condition governing which what fine-grained multi-layered thematic role(s) the controller must have in order to match that of the controllee—a property of obligatory control that can be seen for ordinary PRO in the notorious phenomenon of “control shift.”

The third crucial assumption, which almost goes without saying, is that the ghostly DP operators can bind pronouns inside the TP complement of the C that licenses them, as long as the pronoun matches/is compatible with the ghostly DP in features. Technically this can happen in any of the constructions, but it is most noticeable in the logophoric pronoun constructions and the indexical shift constructions, because logophoric pronouns and indexicals are specified as needing to be bound by a ghostly DP operator with the proper features. One variant of this has a ghostly DP that is an A-binder rather than an A-bar binder, resulting in long-distance anaphor constructions like the one found in Japanese. Another variant is one in which the ghostly DP is allowed to add features to its bindee, resulting in indexiphors/monstrous agreement constructions like the one found in Telugu (Messick 2022) or Donno So (Deal 2020). I also argue that certain kinds of crossover effects can emerge in logophoric constructions, analogous to those that appear when a *wh*-operator binds pronouns and variables of different kinds.

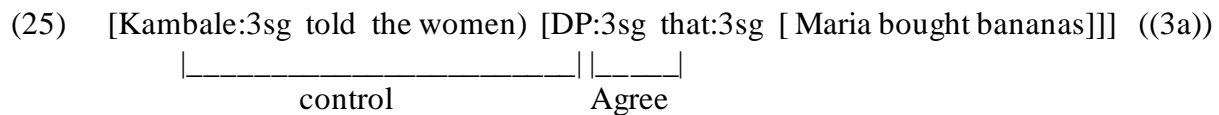
Finally, I assume that C can enter into Agree with the ghostly operators in the CP periphery—either the very same C-head that licenses the ghostly operator in the first place, or a nearby one. An obvious result of this in some languages is that phi-features that the ghostly DP has intrinsically or gets from its controller are transferred to C; this happens in upward C agreement constructions and in allocutive constructions. A less obvious result of this is that some

languages create a pointer from C to the ghostly DP (“Agree-Link” in the terms of Arregi and Nevins 2012) and via the ghostly DP to its controller. This is an ingredient in some switch-reference constructions, I claim.

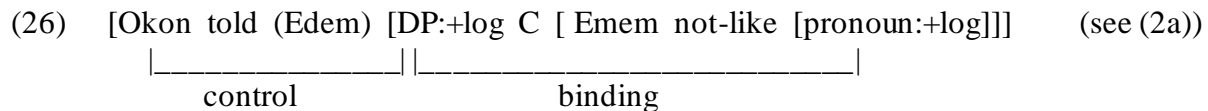
The basic skeletal template for all these constructions, then, is summarized in (24), following in essence Koopman and Sportiche’s (1989) approach to logophoricity in Abe.



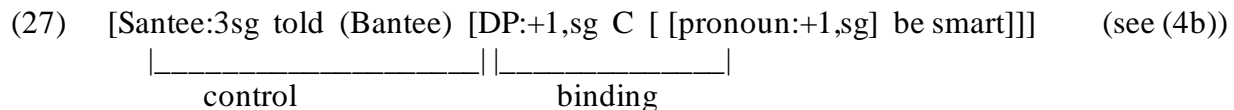
The building blocks of this template—DP-licensing, control, binding, and Agree—are all provided by UG, analogous to the way that “universal mammal anatomy” provides a template of bones which have become expanded or contracted in different mammals to facilitate different functions. For example, upward C-agreement involves licensing a featureless ghostly DP, control of that DP by the matrix subject, and Agree holding between C and the ghostly DP. (The ghostly DP may also bind a pronoun in the CP complement, but nothing special happens with that.)



The logophoric pronoun construction involves C licensing a [+log] DP, which is controlled by the matrix subject. C does not agree with this DP (in any language known to me, although this shouldn’t be impossible) but the DP can bind special [+log] pronouns in its c-command domain. Indeed [+log] pronouns must be bound by this operator.



Similarly, the indexical shift construction involves licensing a [+1st] DP in the CP of an embedded clause, which is again controlled by the matrix subject. Again C does not usually agree with this ghostly DP (although it apparently can in Dargwa; see chapter 3), but it can bind [+1st] pronouns inside its c-command domain. Indeed, [+1st] pronouns must be bound by this operator—by an instance in the root clause if not by one in the embedded clause. (Note that complete phi-feature matching is not required for an argument of the matrix clause to control a ghostly DP operator, if the ghostly DP has some intrinsic features.)



Variants of this construction result in indexiphoric constructions in languages like Telugu and long-distance anaphoric constructions in languages like Japanese. Then finally switch-reference marking on complement clauses in languages like Imbabura Quechua, Washo, and Choctaw

involve licensing a featureless ghostly DP which is controlled by the matrix subject. C then undergoes Agree-Link but not Agree-Copy with both the ghostly DP operator and the highest DP inside its c-command domain (the embedded subject). These Agree-links are interpreted as coreference (or disjoint reference) at LF, resulting in a “same subject” (or “different subject”) construction.

- (28) a. [I want [DP C [I see my mother]]] (see (20a))
 _____|_____|_____|
 control Agree-Link (2x) → SS
- b. [I persuade Juan [DP C [he go to Quito]]] (see (20b))
 _____|_____|_____|
 control Agree-Link (2x) → DS

I must clarify, however, that this theory is proposed for SR only in complement clauses in the minority of languages that allow that. For languages that allow SR only or primarily in adjunct clauses, I do not use a null DP operator, but make use of direct Agree between the C of the adjunct clause and an NP argument in the matrix clause, as in Baker and Camargo Souza (2020), Arregi and Hanink (2022), and related work. Overall, switch-reference constructions are going to be integrated only up to a point with the other rare constructions discussed in this work.

One important further assumption for the overall framework is that the C-space in languages of the world can actually license pairs of ghostly DP operators, not just solitary ones; the CP-periphery can contain an object-like null DP as well as a subject-like one. Whereas the thematic-role-matching condition on obligatory control implies that only the matrix subject can control the subject-like ghostly DPs that I have presented so far, the same condition implies that only the matrix object can control these additional object-like ghostly DPs. This assumption allows us to fold allocutive agreement in a language like Magahi into the account. The fuller claim is that Magahi has two ghostly DPs in the periphery of finite CPs, a subject-like one that is [+1st] and an object-like one that is [+2nd]. Allocutive agreement arises when C licenses the [+2nd] null DP, the matrix object controls it, and a head near C (Fin, according to Alok 2020, 2021) agrees with it. This is sketched in (29) as a rough structure for (23c).

- (29) [Santee: told professor:HH [DP₁:+1 DP₂:+2,HH C:+2,HH [Ram saw Sita]]].
 _____|_____|_____|
 control _____| Agree

The innovation of an object-like ghostly DP allows for allocutive agreement, but goes far beyond it too. The same [+2nd] operator in Magahi also accounts for indexical shift of a second person pronoun like ‘you’, such that it refers to the goal object of a matrix verb like ‘tell’ in (30). Here the pronoun ‘you’ matches DP₂ in features (+2nd, NH) and is bound by it; the indirect object of ‘tell’ controls DP₂, so ‘you’ ends up referring Bantee.

- (30) Santee-aa Bantee-aa-ke kahk-ain ki Ram toraa dekh1-i-au hal.
 Santee-FM Bantee-FM-DAT told.3NH-HH.AL that Ram you.NH.ACC saw-1.S-NH.AL be.pfv
 ‘Santee_i told Bantee_k that Ram saw you_k.’ (said to a teacher)

Similarly, a few African languages have special addressee pronouns that appear in embedded clauses and refer to the goal of the matrix verb alongside their logophoric pronouns that appear in embedded clauses and refer to the subject of the matrix verb. Mupun is one such language, as shown in (31); here the addressee pronoun *gwar* can only refer to the goal of the saying event picked out by the matrix verb, just as the logophoric pronoun *Di* can only refer to the agent of that event. These addressee pronouns can be analyzed as pronouns that need to be bound by a [+log] object-like operator (DP₂ in a structure like (29), but with different features) that exists alongside the [+log] subject-like operator already discussed (DP₁ in a structure like (29)).

- (31) Datar sat n-Dapus nə d'i naa la reep gwar. (ZF: 125)
 Datar say P-Dapus that LOG see ASP? girl ADDR
 'Datar told Dapus that he (=Datar) saw his (=Dapus's) daughter.'

Moreover, one African language, Kipsigis, has double upward agreement on C: C can agree with both the matrix subject, as in Lubukusu, and the matrix object (Diercks and Rao 2019).⁶ This possibility follows given that Kipsigis also has a pair of ghostly DP operators and C in this language can undergo multiple Agree, collecting phi-features from both of these operators.

- (32) a. Ko-i-mwaa- an **i-le-ndʒan** ko-Ø-it layok.
 PST-2SG.S-tell-1SG.O 2SG-C-1SG PST-3.S arrive children
 'You (sg) DID tell me that the children arrived.'
- b. ko-i-mwaa- tʃi **a-le-ndʒi** ko-Ø-it layok
 PST-1SG.S-tell-3.O 1SG-C-3 PST-3.S arrive children
 'I DID tell him/her/them that the children arrived.'

So most of the constructions have subject-oriented and object-oriented twins, although the object-oriented versions are generally much rarer and seem to depend on the subject-oriented version being present in the same language.⁷

1.5 Brief methodological remarks

The design of this investigation and the way it is carried out is not intended to be especially innovative, but rather a further example of how generative linguistics has been fruitfully pursued for years. However, some comments on two general issues may be of some use: how I picked the languages I focus on and my sources for those languages, and how I see the relationship between syntax and semantics in this domain.

⁶ Here I follow Diercks and Rao's (2019) discussion. In contrast, Dreimel and Kouneli (2002) argue that IE is still synchronically the verb 'say' in examples like (32). See Chapter 2 xx for some discussion.

⁷ The one construction that I do not have an object-oriented version of is switch-reference. It is conceivable that the super-rare subject=object construction found in the Panoan language Amahuaca (Clem to appear) is an instance of this, but I do not investigate that here. Another possibility is that the fact that SR alone among these constructions happens freely on adjunct clauses inhibits the development of an object-oriented version.

I also haven't discussed a subject-oriented version of allocutive agreement. That may exist in Dargwa (xx); see chapter 3 for discussion. Allocutive agreement may be more common than speaker agreement for a simple functional reason: it is usually obvious who is speaking, and what their gender is, and one does explicitly not honor oneself.

1.5.1 Languages and sources

First on languages and sources. For the most part, this is not the kind of topic that can be pursued by pulling out descriptive grammars on a wide range of languages and seeing what they happen to say about these topics. In part, that is because plenty of languages do not have any of the five constructions, and fewer still will have more than one. In addition, these are phenomena that take place primarily in embedded clauses, hence only in sentences with a relatively high degree of complexity. Such complex sentences are usually described only rather incompletely in descriptive grammars. To carry out this kind of project, then, one really wants access to sophisticated native speakers of languages that have the relevant constructions (sophisticated in that they are able to concentrate on and retain examples of a certain complexity, not in being indoctrinated into a particular theory). Fortunately, for all but one of the five constructions outlined in this chapter, I have had the opportunity to work with native speaker linguists in detailed fashion over the course of several years, feeding back what I have learned about one construction in one language when asking questions about another construction in another language. Therefore, each subtopic has a strong home base in a particular language—hopefully one that is typical and/or especially revealing concerning the phenomenon in question. The primary languages and sources are as follows:

- (33) a. Upward C-agreement: Ibibio (Willie Willie), Lubukusu (Justin Sikuku), Kinande (Philip Mutaka)
- b. Allocutive agreement: Magahi (Deepak Alok)
- c. Indexical shift: Magahi (Deepak Alok), (also a little Sakha, Nadya Vinokurova plus some MA students in Yakutsk)
- d. Logophoric pronouns: Ibibio (Willie Willie), Yoruba (Oluseye Adesola), a little Edo (O.T. Stewart)
- e. Switch reference: Yawanawa (*Livia Camargo Souza), Shipibo (three consultants in Lima, 2012).

A pattern in this is that the people above have almost all been students or postdocs at Rutgers University, where it has been my privilege to work with them closely in various capacities. It is not for nothing, then, that this book is dedicated to my students; it would have been impossible to pursue this project without their involvement and inspiration. The one person here who is not a native speaker of the relevant language is Livia Camargo Souza, but she has worked on Yawanawa for many years—since before she came to Rutgers—has made many trips to the Amazon to study it, and has developed relationships where she can ask speakers questions remotely when needed. Switch-reference is also a special case in that the SR languages that are most relevant to this study are ones in which SR marking is found on complement clauses as well as on adjunct clauses. That is true in a limited respect in Yawanawa and Shipibo which will help us to break into the topic in an effective way, but once that is done I will be more dependent on secondary sources for this topic than for the others. In addition, there are two constructions that are related logophoricity and indexical shift that I have not picked out for separate discussion in this overview, but for which I have had similar access to native-speaker linguists of strategic languages: these are the indexiphoric/monstrous agreement construction found in Telugu (Sreekar Rogatham) and the long-distance anaphor construction found in Japanese (Shiori Ikawa). Even if the grand unification project that I am attempting is not deemed successful, I hope that my individual descriptions and analyses will be found to have value.

Of note here is the fact that two of these languages are special in having not just one but two of the rare constructions. As such, they provide very special opportunities to see how the constructions

compare and interact within the very same language. One of these is Magahi, which has both allocutive agreement and indexical shift. These were serendipitous, and the two constructions turn out to be deeply interdependent, both using the same pair of ghostly DP operators, as hinted at above. The other is Ibibio, which has both upward C-agreement and logophoric pronouns. This fact motivated me to reconnect with Willie Udo Willie via the network of Ken Safir's Afranaph project. In this case, the two constructions turn out to be largely independent of each other, although parallel, using different ghostly DP operators at different levels of the CP space. Magahi and Ibibio together thus constitute the heart of the heart of the book.

Beyond these few languages where there has been the opportunity to do significant primary research, the most useful sources have been articles or books (chapters) which study one of the constructions in some other language in significant detail. In each case, I compare the results that have been achieved by me and my collaborators with others reported in the literature on the various construction, usually the "classics" from the literature (e.g. Ewe for logophoric pronouns; Amharic, Zazaki, and Nez Perce for indexical shift; Basque for allocutive marking, and so on). This gives some sense of what is and is not stable about a particular construction across (some of) the languages that are known to have it, but the "sampling" is not very systematic or complete. My emphasis has been to consider cases that are relatively well described and analyzed (since those are relatively few) rather than pursuing typological or areal balance. More generally, one could imagine a unifying project proceeding at the three levels in (34)

- (34) a. Studying construction A in language A' .
 b. Studying construction A in languages A'', A''', A*,
 c. Studying construction A in language A', B in B', C in C', ...

Indeed, there is some material at all three levels in this work. But the emphasis is on the first level ((34a)) and the third level ((34c)), with investigation at the second level ((34b)) being more haphazard and opportunistic. The hope here is that the languages that I can focus on in depth listed in (33) are fairly representative of the phenomena in question, either by blind luck or because UG is such that almost any instance of one of these phenomena is a good instance of the phenomenon. If so, then, not much will be lost in moving fairly quickly—although not immediately!—from the level of (34a) to the level of (34c). In proceeding this way, I am assuming that if one spends all one's time and energy on (34a) and (34b), one might miss something interesting and important on the level of (34c). Time will tell to what degree this research strategy is vindicated. (As my old bridge partner used to say: "You pay your money, and you take your chances.")

1.5.2 Syntax and semantics

Finally, I say something about how I see the relationship of syntax and semantics in this domain (and indeed in most domains). The five constructions that I have mentioned have gotten uneven amounts of attention from specialists in the two subfields. Some of them have been discussed more by syntacticians than by semanticists. Unsurprisingly, that is particularly true for the two constructions in which agreement plays a prominent role: upward C-agreement and allocutive agreement. Others of the topics have been discussed more by semanticists than by syntacticians. Unsurprisingly, that is particularly true for the two constructions that feature bound pronouns: indexical shift and logophoricity (at least if one includes LD anaphors with the latter). At least one of these topics, switch-reference, has been the subject of something of a turf war between syntactically-oriented approaches (e.g. Finer 1984, 1985) and semantically-oriented approaches

(e.g. Stirling 1993, McKenzie 2012). I take this analytic diversity to be part of the “fun” of working in this area. One basic assumption that is relevant to staying oriented in all of this is that no one group owns any of these topics. We should not say that syntacticians own upward C-agreement, or that semanticists own indexical shift. It is obviously true that any construction in any natural language has both a syntactic structure and a semantic interpretation. Therefore, it can obviously be studied from both perspectives. In particular cases, one kind of analysis may be relatively trivial, while the other kind may be deep and illuminating. But in general we will only find that out by pursuing both kinds of analysis for every construction, and what looks trivial from within one frame of reference can prove to be deep and illuminating from another frame of reference.

Pursuing this a little further, not only does every natural language construction have a syntax and a semantics, in my experience the two are often parallel to one another to a large extent. When that is the case, a descriptive generalization, or even an explanation, can be stated in either terms. It is still, in a significant sense, the same generalization/explanation, given the homomorphic relationship between the two subdisciplines in this domain. Sometimes the generalization makes more sense conceptually when couched in one set of terms rather than the other, such one has a stronger sense of why it holds. That is great when it happens, but I take it to be a secondary priority, less urgent than finding new generalizations and explanations stated in whatever terms. We can also expect there to be some cases in which a syntactic account of a construction and a semantic account are complementary, not homomorphic, where one can explain some features of the construction about which the other is relatively silent. That is great too. Occasionally there will be actual conflicts, where a syntactic account and a semantic account are genuinely different, and one works better than the other. But I take that to be a less common situation, to be resolved by doing both our syntax and semantics better. In the ultimate scheme of things, every natural language construction must have both a syntax and a semantics, and they must be consistent with one another, by the definition of what a language is.

If that is the right general lay of the land, how do I position myself within it? The answer is: as an unabashed syntactician reasonably (but not perfectly) literate in the semantic issues. Therefore, in this investigation I lead with the syntax and with generalizations and (partial) explanations phrased in the syntactic idiom. It is what I do, what I’ve been trained to do, and what I profess to be reasonably good at. There is a general flow in this work from the more syntactically-studied topics (chapter 2 on upward C-agreement, chapter 3 on allocutive agreement) toward the more semantically-studied topics (chapter 4 on indexical shift, chapter 5 on logophoricity), as I start in my comfort zone and see where that leads. (Although, somewhat ironically, I claim that the ghostly DP that C agrees with in cases of upward agreement makes a relatively clear semantic contribution—more so indeed than some others that have gotten more attention in the semantics literature.) In pursuing this study, I want to emphasize the similarities and interconnections across the five constructions, and a bit beyond them. Suppose, then, that I start with constructions thought to be syntactic, give them a syntactic analysis, and go on to argue that constructions thought to be semantic are similar to them in important ways, such that they should be explained in using the same theoretical resources or risk failing to capture a generalization. Is this being aggressive or imperialistic for syntax in a way that other researchers should feel threatened or offended by, or consider misguided? Hopefully not. Please keep in mind that in asserting a generalization or explanation in syntactic terms, I do not mean to be denying that there is also a generalization or explanation in semantic terms that may be parallel to the syntactic version, that may be even deeper than the syntactic version, or that may overlap

with it and cover things that the syntactic version does not. We should all spin the best yarns we can, working out how they fit together as we go. I also do not in general present any worked out formal semantics for the syntactic structures that I propose for the various constructions. That might seem quaint and retro of me in the 2020s, almost like I was trained in the 1980s. However, my experience has been that it is never impossible to provide a reasonable semantics that goes with a reasonable syntax for a certain construction. One does what one has to do, and formal semantics is not such a tightly constrained enterprise that it cannot find a way, as I understand the situation. To me, it seems fairly clear what the syntactic representations I present are supposed to mean in informal approximate terms; hopefully it will be clear to others as well. Perhaps someone will even feel motivated to work out the semantic details at some point. But me doing that would be like a dog challenging a dolphin to a swimming contest: it would not be impossible, probably, for the dog to get to the other end of the course, but it would be relatively slow, clumsy, and not the dog doing what it does best (it has the wrong sort of forelimbs, as it happens). I take my strengths to be syntax, collecting data on interactions from less studied languages, and constructing a big picture. So I concentrate on doing that, rather than on details about the syntax-semantics interface in this domain.

1.6 What is where and midlevel theoretical results

The basic organization of the rest of the book flows out of what I have said so far. The next six chapters are each organized around one of the crosslinguistically rare constructions that has been introduced in this chapter, with a general movement from constructions which have been more discussed in the syntax literature toward constructions that have been discussed as much or more in the semantics literature. Chapter 2 takes up the topic of upward C-agreement in Central African languages, perhaps the most purely syntactic topic of them all (or perhaps not). Chapter 3 considers allocutive agreement, which can function as another form of upward C-agreement in Magahi where (unlike Northern dialects of Basque) allocutivity is freely marked on complement clauses as well as root clauses. Chapter 4 turns to indexical shift, focusing again on Magahi; this is a natural next step given that allocutive agreement and indexical shift interact in Magahi in systematic ways. Chapter 5 discusses logophoric pronouns, with an emphasis on Ibibio—another natural step given that logophoric pronouns in CP complements in West African languages refer to the matrix subject in much the same way that shifted first person indexicals do in Magahi and other languages with indexical shift. This chapter also compares logophoric pronouns in Magahi with long distance reflexives in Japanese, which are often said to be in some sense logophoric elements. Chapter 6 extends the discussion to indexiphoric constructions, not exemplified here so far. These are constructions in which what seems to be a logophor or long-distance anaphor (or bound pronoun) triggers first person agreement on the embedded verb; as such, it combines elements of indexical shift and logophoricity in what looks like a hybrid form. (35) is an example from the Dravidian language Telugu.

- (35) Raju [tanu parigett-əə-nu ani] cepp-əə-Du. (Messick 2022)
 Raju 3.SG run-PST-1SG that say-PST-3.M.SG
 ‘Raju_i said that he_i ran.’

Chapter 7 then considers whether switch-reference belongs in this family of constructions. I argue that the answer is sometimes: SR on adjunct clauses typically does not involve a ghostly

DP operator controlled by the matrix subject, but SR on complement clauses does in the subset of languages that permit that. Finally, chapter 8 takes up a more theoretical topic: it takes a closer look at the principles of obligatory control by which an argument of the matrix verb controls a ghostly DP at the periphery of its CP complement, a notion that is used quasi-descriptively in chapters 2-7. There I pursue a unified analysis of this phenomenon and normal cases of the control of the PRO subject of infinitival clauses, even though the choice of which argument of the matrix verb controls the null DP seems different in an important range of cases.

This chapter, and indeed the larger book, is organized more on basis of the constructions being analyzed than on the midlevel theoretical discoveries being made. There are some theoretical discoveries and innovations, however, and I would be happy for readers to find them and appreciate them. Here is an overview of the main ones.

One midlevel result concerns agreement. Three of the six constructions involve full-fledged agreement including the transfer of phi-features: upward C-agreement, allocutive agreement, and the indexiphor construction in (35). Comparison of these constructions reveal that they are often subject to an additional condition on what can control the ghostly DP operator: in two of the three constructions, the controller must itself be the goal of an agreeing T. I call this the T/Agree condition. I claim that this constraint shows us something about how Agree works. In particular, it testifies to Arregi and Nevin's (2012) distinction between Agree-Link, which creates a pointer between a probing functional head and its NP/DP goal, and Agree-Copy, which in a distinct step actually transfers phi-features from the goal to the probe. The main discussion of this is at the end of chapter 2, with a brief reprise in chapter 3 and an extension in chapter 6.

Another midlevel result has to do with the typology of ghostly DP operators: what kinds of features they can have, and how this affects their behavior in ways that go beyond just determining the features of the pronouns that they bind. The most important distinction is between ghostly DPs that have interpretable features and those that do not. I argue that ghostly DPs that do not have interpretable features need to undergo obligatory control almost immediately (on the next phase) or they run afoul of the principle of Full Interpretation. In contrast, ghostly DPs that have interpretable features can survive outside of contexts of obligatory control—e.g. in adjunct clauses, subject clauses, relative clauses, and root clauses—and even when they are in a context that requires obligatory control they allow a somewhat broader range of controllers. This topic is touched upon in most of the chapters, but the first key discussion is toward the end of chapter 3 (comparing the goal of allocutive agreement to the goal of upward C agreement) and the fullest discussion is toward the end of chapter 5, by which point we have a critical mass of examples large enough to discern some significant patterns.

The operator-pronoun binding relationship gets some significant theoretical attention in chapters 5 and 6. Chapter 5 discovers some new types of crossover violations in some languages with logophoric pronouns, where the same ghostly DP operator may or may not be able to bind both a logophoric pronoun and an ordinary pronoun, depending on the exact feature values of the items involved and the details of the c-command relationship. This is potentially relevant to the formulation of the Crossover principle. Chapter 6 then argues that a ghostly DP operator can actually add phi-features to the pronoun it binds, creating complex phi-feature bundles that result in indexiphoric constructions like (35).

The biggest theoretical challenge and opportunity is saved for last and gets a chapter of its own (chapter 8). This concerns the fundamental nature of the control relationship between an argument of the matrix verb and the ghostly DP operator in the periphery of its CP complement.

In some important ways, this is analogous to the relationship between PRO and its controller in ordinary cases of obligatory control. However, there are some differences too, particularly as to which argument of the matrix verb is chosen to be the controller of the null DP in the complement clause. I take these challenges to be opportunities, and consider what a Generalized Control Theory could look like that has both ordinary control theory and what is needed for ghostly DP operator constructions as special cases. This is the most theoretical ambitious piece of the unified analysis of this set of crosslinguistic constructions.

Let's get on with it then!

References

- Alok, Deepak. 2020. *Speaker and addressee in natural language: Honorificity, indexicality and their interaction in Magahi*. New Brunswick, NJ: Rutgers University dissertation.
- Alok, Deepak. 2021. The morphosyntax of Magahi addressee agreement. *Syntax* 24(3). 263–296.
- Alok, Deepak & Baker, Mark. 2018. On the mechanics (syntax) of indexical shift: Evidence from allocutive agreement in Magahi. Manuscript, Rutgers University.
- Alok, Deepak & Haddican, Bill. (to appear). The formal heterogeneity of allocutivity. *Glossa*.
- Antonov, Anton. 2015. Verbal allocutivity in a crosslinguistic perspective. *Linguistic Typology* 19:55-85.
- Arregi, Karlos, and Nevins, Andrew. 2012. *Morphotactics: Basque auxiliaries and the structure of Spellout*. Oxford: Oxford University Press.
- Austin, Peter. (1981). Switch-reference in Australia. *Language* 57: 309-34.
- Baker, Mark. (2008). *The syntax of agreement and concord*. Cambridge: Cambridge University Press.
- Baker, Mark. (2003). *Lexical Categories: Verbs, Nouns, and Adjectives*. Cambridge: Cambridge University Press.
- Baker, M. and L. Camargo Souza (2020). "Agree without agreement: Switch-referent and reflexive voice in two Panoan languages." Natural Language and Linguistic Theory.
- Carstens, Vicki. (2016). Delayed Valuation. *Syntax* 19. 1-42.
- Chomsky, Noam. (1977). On *wh*-movement.
- Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in language*, ed. Michael Kenstowicz, 1-52. Cambridge, Mass.: MIT Press.
- Clements, G. N. (1975). "The logophoric pronoun in Ewe: its role in discourse." Journal of West African Languages 3: 19-37.
- Cole, Peter. (1982). *Imbabura Quechua*. Amsterdam: North-Holland.
- Cole, Peter. (1983). Switch-reference in two Quechua languages. In John Haiman and Pamela Munro (eds.), *Switch-reference and universal grammar*, pp. 1-15. Amsterdam: John Benjamins.
- Cole, Peter and Gabriella Hermon. (2011). Nominalization and case assignment in Quechua. *Lingua* 121: 1225-1251.
- Culy, Christopher. (1994). "Aspects of logophoric marking." Linguistics 32: 1055-1094.
- Deal, Amy Rose. 2020. *A theory of indexical shift: Meaning, grammar and crosslinguistic variation*. Cambridge, MA: MIT Press.
- Diercks, Michael. 2013. Indirect Agree in Lubukusu complementizer agreement. *Natural Language and Linguistic Theory* 31:357-407.

- Finer, Daniel. (1984). The formal grammar of switch-reference. Ph.D Dissertation. University of Massachusetts Amherst.
- Finer, Daniel. (1985). The syntax of switch-reference. *Linguistic Inquiry* 16: 35-55.
- Hagege. (1984).
- Hermon, Gabriella (1985). *Syntactic modularity*. Dordrecht: Foris
- Jacobsen, William. (1967). Switch-reference in Hokan-Coahuiltecan. In Dell Hymes and W. Bittle, (eds.), *Studies in Southwestern Ethnolinguistics*, pp. 238-63. The Hague: Mouton.
- Koopman, Hilda, and Dominique Sportiche (1989). "Pronouns, logical variables, and logophoricity in Abe." *Linguistic Inquiry* 20(4): 555-589.
- Landau, Idan. (2013). *Control in generative grammar*. Cambridge, Cambridge University Press.
- McKenzie, Andrew. (2012). The role of contextual restriction in reference-tracking. Ph.D. Dissertation, University of Massachusetts Amherst.
- McKenzie, Andrew. (2015). A survey of switch-reference in North America. *International Journal of American Linguistics* 81: 409–448.
- Messick, Troy. (2022). On apparent pronominal feature contradictions: Shifty agreement in Telugu. *Syntax*.
- Miyagawa, Shigeru. 2012. Agreements that occur mainly in main clauses. In Lobke, Albrecht & Haegeman, Liliane & Nye, Rachel (eds.), *Main clause phenomena: New horizons*, 19–112. Amsterdam: John Benjamins.
- Miyagawa, Shigeru. 2017. *Agreement beyond phi*. Cambridge, MA: MIT Press.
- Oyharçabal, Beñat. 1993. Verb agreement with non arguments: On allocutive agreement. In Huale, José & Ortiz de Urbina, Jon (eds.), *Generative studies in Basque linguistics*, 89–114. Amsterdam: John Benjamins.
- Podobryaev, Alexander. 2014. *Persons, imposters, and monsters*. Cambridge, MA: MIT dissertation.
- Schlenker, Philippe. 1999. *Propositional attitudes and indexicality: A cross-categorical approach*. Cambridge, MA: MIT dissertation.
- Schlenker, Philippe. 2003. A plea for monsters. *Linguistics and Philosophy* 26(1). 29–120.
- Shklovsky, Kirill, and Sudo, Yasutada. 2014. The syntax of monsters. *Linguistic Inquiry* 45:381-402.
- Speas, Margaret, and Tenny, Carol. 2003. Configurational properties of point of view roles. In *Asymmetry in grammar*, ed. Anna Maria Di Sciullo, 315-344. Amsterdam: John Benjamins.
- Stirling, Lesley (1993). *Switch-reference and discourse representation*. Cambridge: Cambridge University Press.
- Rizzi, Luigi. 1997. The fine structure of the left periphery. In Haegeman, Liliane (ed.), *Elements of grammar*, 281–338. Dordrecht: Kluwer.
- Valenzuela, Pilar (2003). Transitivity in Shipibo-Konibo Grammar”. Ph.D Dissertation. University of Oregon.