

A VERY LONG-DISTANCE ANAPHOR?

Mary Dalrymple
University of Oxford

Proceedings of the LFG11 Conference

Miriam Butt and Tracy Holloway King (Editors)

2011

CSLI Publications

<http://csli-publications.stanford.edu/>

Abstract

Yɔ̀g Dii (Niger-Congo/Adamawa-Ubangi, Cameroon; also called Duru) has a complicated pronominal system, originally described by Bohnhoff (1986), with four series of pronouns whose distribution is determined by their grammatical function and the type of clause in which they appear. One series seems to exhibit an otherwise unattested form of non-locality: at least one clause must intervene between the pronoun and its antecedent, and the presence or absence of coreferent phrases in the intervening clause does not affect its appearance or distribution. The nature of the relation between this very long-distance pronoun and its antecedent seems to violate otherwise well-established notions of locality of anaphoric relations and, indeed, of grammatical dependencies more generally. We present an analysis of the binding requirements of this anaphor that relies on features associated with different parts of its binding domain, and compare our analysis to alternatives which involve the specification of extended paths.

1 Locality in grammar

It is generally assumed that languages do not have grammatical dependencies that are exclusively nonlocal – there are no grammatical dependencies that operate at a minimal distance of two clauses away, for example (Fitzpatrick 2002, Sag 2008, among many others). In the context of anaphoric binding patterns, this assumption amounts to the claim that anaphors never ignore their local context. This is the *Locality Condition* of Dalrymple (1993), stated as: “binding constraints ... always refer to local elements, never exclusively to nonlocal ones”, and the *subset principle* of Manzini and Wexler (1987) for anaphoric binding domains, stating that smaller potential binding domains are always properly contained in larger ones.

Anaphoric binding patterns in Yɔ̀g Dii appear to run counter to these standardly accepted generalisations. There are several series of pronouns in Yɔ̀g Dii, one of which, glossed 2LD in the following, requires a very long-distance binder. In example (1), 2LD appears as the subject of a subordinate clause (*he repay the IOU*) which is itself contained within a subordinate clause (*his friend asked him that he repay the IOU*); 2LD must be bound by the main clause subject, two clauses away:¹

[†]I am grateful to Lee Bohnhoff (personal communication, October 1991) for providing corrections to mistakes in transcriptions and indices for the examples taken from Bohnhoff (1986), additional examples, and helpful comments on the patterns discussed here. For comments on this paper, I am grateful to Ash Asudeh, Miriam Butt, Tracy Holloway King, Jean-Marie Marandin, the audience at “Ling Lunch” Paris Diderot, June 2011, and the audience at LFG11, particularly Louisa Sadler and Doug Arnold.

¹The form glossed CM is a clause-final particle.

- (1) Nán ba'ad Ø 'ò [moo 'èh dà bì tóó bà ka vì bi
 man work (he_i) say for what friend his.LD_i other that sb-he_j ask him.LD_i
 [bà 'ìi súúwù 'íá]]?
 that he.2LD_i repay.him_j CM-Q
 'The worker_i asked why his.LD_i friend asked him.LD_i that he.2LD_i repay
 the IOU.' (corrected version of Bohnhoff, 1986, 119)

- (2)
$$\left[\begin{array}{l} \text{PRED} \text{ say} \\ \text{SUBJ} [\text{worker}]_i \\ \text{COMP} \left[\begin{array}{l} \text{PRED} \text{ ask} \\ \text{SUBJ} [\text{his.LD}_i \text{ friend}] \\ \text{OBJ} [\text{LD}]_i \\ \text{COMP} \left[\begin{array}{l} \text{PRED} \text{ repay} \\ \text{SUBJ} [2\text{LD}]_i \end{array} \right] \end{array} \right] \end{array} \right]$$

The 2LD pronoun can be used whether or not there is a coreferential pronoun in the intervening clause. As in (1), there is a coreferential subject in the immediately higher clause in example (3), but not in the equally acceptable example in (4):

- (3) Bà'á Ø gàà [[sèy ìi làà tée] bà bíń hò hẹn
 Papa_i, (he_i) knows time he.2LD_i goes when, that he.LD_i.will see thing
 Múúsà wòò]
 Moses his
 'Papa_i knows that when he.2LD_i goes, he.LD_i'll see Moses's thing.'
 (L. Bohnhoff, p.c.)

- (4) Bà'á Ø gàà [[kóó ìi lúu ní sị'] bà míń hò hẹn
 Papa_i, (he_i) knows time he.2LD_i leave NEG even, that I.will see thing
 Múúsà wòò]
 Moses his
 'Papa_i knows that even if he.2LD_i doesn't leave, I'll see Moses's thing.'
 (L. Bohnhoff, p.c.)

- (5)
$$\left[\begin{array}{l} \text{PRED} \text{ know} \\ \text{SUBJ} [\text{Papa}]_i \\ \text{COMP} \left[\begin{array}{l} \text{PRED} \text{ see} \\ \text{SUBJ} [\text{he.LD}_i/\text{I}] \\ \text{OBJ} [\text{thing}] \\ \text{ADJ} \left\{ \left[\begin{array}{l} \text{PRED} \text{ leave} \\ \text{SUBJ} [2\text{LD}]_i \end{array} \right] \right\} \end{array} \right] \end{array} \right]$$

Thus, the 2LD pronoun is an exceptionally long-distance anaphor, and seems to exemplify an exclusively nonlocal dependency: it must corefer with a subject at

least two clauses distant, and its distribution is not affected by the presence or absence of intervening potential binders.

Binding patterns for the 2LD pronoun may appear similar to familiar patterns of switch reference, where clauses are marked to indicate coreference between arguments, often subjects, of two different clauses. Haiman and Munro (1983) provide example (6) from Pima, citing Langdon and Munro (1979) and personal communication from Etheleen Rosero. The morpheme glossed SS enforces coreference between the subject of *cry* and the subject of *hit*, while the DS morpheme indicates that the subjects of the two verbs are not coreferent:

- (6) a. Hegai 'uuvi 'a-t 'am ʃohñi hegai ceoj c 'am ʃoʃa.
 that woman 3-perf hit that man SS cry
 'The woman_i hit the man and she_i cried.'
- b. Hegai 'uuvi 'a-t 'am ʃohñi hegai ceoj ku-t (hegai ceoj) 'am ʃoʃa.
 that woman 3-perf hit that man DS that man cry
 'The woman hit the man_i and he_i (the man) cried.'
- (Pima; Haiman and Munro, 1983, x)

Like other anaphoric processes, however, and unlike 2LD, switch-reference always operates locally: according to Haiman and Munro (1983, xiii), "there seem to be no languages ... in which switch-reference is marked *exclusively* between non-adjacent clauses. Thus, if a language has switch-reference marking between non-adjacent clauses, it will also mark switch-reference between adjacent clauses."

2 PRON and SUBORD pronouns

Yąg Dii has four distinct series of pronouns, each with a different distribution. First, there is a basic series of subject, object, and possessive pronouns, which we will gloss as PRON; Bohnhoff calls this the *mí* series, after the first person subject and object forms. Second, there is a series of SUBORD pronouns for use in subject position of certain main clauses as well as many subordinate clauses. The PRON and SUBORD pronoun paradigms are given in Table 1. The discontinuous *incl.pl* form *ba...ví* can be interrupted by the verbal complex (the verb or series of serial verbs and any object pronouns). Besides these forms, subject forms in the PRON series (but not the SUBORD series) can appear with suffixes indicating future or nonfuture tense. There are also emphatic forms corresponding to each member of the subject PRON, future subject PRON, and SUBORD series, and there is a separate series of possessive affixes for use with kinship terms; see Bohnhoff (1986) for further discussion of these forms. For present purposes, it will be sufficient to distinguish the members of the PRON and SUBORD series that are listed in Table 1.

The choice of PRON vs. SUBORD subject pronoun form depends only on the syntactic environment, and is not determined by requirements of coreference

	PRON subject	PRON object	PRON possessive	SUBORD subject
1.sg	-n/mí	-n/mí	míí	'àn
1incl.dual	ba	ba	bàà	ba
2.sg	-m/mó	-m/mó	móó	'àm
3.sg	∅	-wɛ	wòò	'à
1excl.pl	vó	vó	vóó	'òo
1incl.pl	ba...ví	ba ví	bàà ví	ba...ví
2.pl	ví	ví	víí	'ì
3.pl	wɛ	wɛ	wòò	'èu

Table 1: PRON and SUBORD pronouns, from Bohnhoff (1986, 107,109,110).

	main clauses	subordinate clauses
PRON:	imperfective-factative, perfective-factative	indirect quotation, comparison clauses, causal adjuncts (“be- cause...”) introduced by <i>moo</i> , 'until' adjuncts
SUBORD:	imperfective-hortative	indirect order, relative clause, temporal/locative/conditional clause, purpose clause, con- cessive clause, causal adjunct introduced by <i>ka</i> or <i>bà</i>

Table 2: Distribution of PRON and SUBORD subject pronouns, from Bohnhoff (1986, 107-108).

or noncoreference with an element of the main clause. The subject PRON and SUBORD pronouns are in complementary distribution, as detailed in Table 2. In examples (7)–(10), the basic PRON subject, object, and possessive pronouns are used:

- (7) Imperfective-factative:
Mó làà kaalí
you.PRON go to.town
‘You go to town.’
(Bohnhoff, 1986, 107)

- (8) Indirect quotation:
... bà mó làà kaalí
that you.PRON go town.to
‘... that you go to town.’
(Bohnhoff, 1986, 107)

(9) Mí hẹ ví 'ú
 I.PRON see you.Pl.PRON CM
 'I see you.' (Bohnhoff, 1986, 110)

(10) Mí hẹ lig mọ́ sù'ú
 I.PRON see house your.Sg.PRON already
 'I saw your house already.' (Bohnhoff, 1986, 110)

Examples (11) and (12) require the SUBORD subject form. We will refer to the domain in which the SUBORD form is used as the SUBORD domain:

(11) Imperfective-hortative:
 'Àm làà kaalí
 you.SUBORD.must go town.to
 'Go to town!' (Bohnhoff, 1986, 108)

(12) Temporal/locative/conditional:
 Tòw/sè'èy/ya 'àm làà kaalí téé
 if/when/where you.SUBORD go town.to demonstrative
 'If/when/where you go to town...' (Bohnhoff, 1986, 108)

SUBORD pronouns are found only in subject position; there is no separate SUBORD series of nonsubject pronouns.

The choice between PRON and SUBORD subject pronouns is not governed by requirements of coreference or noncoreference with an element of the main clause. Though the SUBORD domain is often a subordinate clause, a SUBORD pronoun is required as the main clause subject in examples (11) and (13):

(13) 'Àn làà kaaláa?
 must.I.SUBORD go town.to.Q
 'Must I go to town?' (Bohnhoff, 1986, 107)

In (14), the SUBORD pronoun appears within the complex clausal complement of the main verb *say*. It does not corefer with any argument in the immediately higher clause, though it is coreferent with an argument in the main clause:

(14) Nà'á ∅ 'òd bà'á [[sè'èy bà 'à fíí ya
 Mother_i (she_i) says.to Father_j time that he.SUBORD_j returns comes
 babbí téé] bà bín dèè dabbì]
 field.from then] that she.LOG_i cook yam.CM
 'Mother_i says to Father_j that when he_j returns from the field, she_i will cook
 the yams.' (corrected version of Bohnhoff, 1986, 122)

$$(15) \left[\begin{array}{l} \text{PRED } \text{say} \\ \text{SUBJ } [\text{mother}]_i \\ \text{OBJ } [\text{father}]_j \\ \\ \text{COMP } \text{ld} \left[\begin{array}{l} \text{PRED } \text{cook} \\ \text{SUBJ } [\text{LD}]_i \\ \text{OBJ } [\text{yams}] \\ \text{ADJ } \left\{ \text{subord} \left[\begin{array}{l} \text{PRED } \text{return} \\ \text{SUBJ } [\text{SUBORD}]_j \end{array} \right] \right\} \end{array} \right] \end{array} \right]$$

Example (16) is structurally similar to (14), but it contains two SUBORD pronouns. The SUBORD pronoun subject of *return* corefers with the SUBORD pronoun subject of *cook* in the immediately higher clause as well as with a nonsubject in the main clause:

(16) Nà'á ∅ 'òd bà'á [[sè'èy bà 'à fíí ya
 Mother_i (she_i) says.to Father_j time that he.SUBORD_j returns comes
 babbí téé] bà 'à dèè ðabbì]
 field.from then] that he.SUBORD_j cook yam.CM
 'Mother_i says to Father_j that when he_j returns from the field, he_j should
 cook the yams.' (corrected version of Bohnhoff, 1986, 122)

$$(17) \left[\begin{array}{l} \text{PRED } \text{say} \\ \text{SUBJ } [\text{mother}]_i \\ \text{OBJ } [\text{father}]_j \\ \\ \text{COMP } \text{ld} \left[\begin{array}{l} \text{PRED } \text{cook} \\ \text{SUBJ } [\text{SUBORD}]_j \\ \text{OBJ } [\text{yams}] \\ \text{ADJ } \left\{ \text{subord} \left[\begin{array}{l} \text{PRED } \text{return} \\ \text{SUBJ } [\text{SUBORD}]_j \end{array} \right] \right\} \end{array} \right] \end{array} \right]$$

Thus, the choice of PRON or SUBORD pronoun forms is determined in purely structural terms: clauses of particular types require the SUBORD form of the subject pronoun rather than the PRON form.

3 LD pronouns

Besides the PRON and SUBORD series, Yag Dii has a third series of bound pronouns which are used only in certain subordinate domains to corefer with a subject in a higher clause. We will gloss these pronouns with the label LD. Table 3 augments the patterns in Table 1 with the nonemphatic subject, object, and possessive LD forms. As with the PRON forms, LD forms can appear with suffixes for future and nonfuture tense, and there is a separate series of possessive forms for use with kinship terms; see Bohnhoff (1986) for the complete paradigms.

According to what Bohnhoff (1986, 112) calls the *reference condition*, the LD pronoun appears in a restricted set of subordinate clauses which we will call the LD

	PRON subject	PRON object	PRON possessive	SUBORD subject	LD subject	LD object	LD possessive
1.sg	-n/mí	-n/mí	míí	'àh	bi	-n/mí	míí
1incl.dual	ba	ba	bàà	ba	bi	ba	bàà
2.sg	-m/mó	-m/mó	móó	'àm	bi	bi	bìì
3.sg	∅	-w#	wòò	'à	bi	bi	bìì
1excl.pl	vó	vó	vóó	'òo	bi	vó	vóó
1incl.pl	ba...ví	ba ví	bàà ví	ba...ví	bi	ba ví	bàà ví
2.pl	ví	ví	víí	'ì	bi	bi	bìì
3.pl	w#	w#	vòò	'ùu	bi	bi	bìì

Table 3: Pronouns including LD forms, from Bohnhoff (1986, 107,109,110,113).

	main clauses	subordinate clauses
SUBORD only:	imperfective- hortative	relative clause, concessive clause, temporal/locative/conditional clause
LD only:		indirect quotation, subordinate desiderative
both SUBORD and LD allowed:		indirect order, purpose clause, causal adjunct introduced by <i>ka</i> or <i>bà</i>

Table 4: Distribution of LD and SUBORD subject pronouns.

domain, and must be bound by the grammatical subject of the clause immediately containing the LD domain, which we will call the LD antecedent. In (18), the LD domain is the subordinate clause *they go to town*, and the LD antecedent is the subject of the matrix verb *want*:

- (18) V# híí [bi làà kaalí]
they.PRON_i want they.LD_i go town.to
‘They want to go to town.’ (corrected version of Bohnhoff, 1986, 113)

Bohnhoff (1986, 112) shows that the LD domain consists of indirect quotations, subordinate desiderative clauses, indirect orders, purpose clauses, and causal adjuncts introduced by *ka* or *bà*. As shown in Table 4, there is some overlap between the LD domain and the SUBORD domain. Where either pronoun can be used, the LD pronoun is used when coreference with the LD antecedent is intended; when noncoreference is intended, the SUBORD form must be used. In contrast with example (18), the SUBORD form is used in example (19), since the pronoun appears in subordinate subject position and the subordinate clause is an indirect order, one of the environments in which the LD domain and the SUBORD domain overlap.

- (19) V# híí ['ù# làà kaalí]
they.PRON_i want they.SUBORD_{*i,j} go town.to
‘They want others to go to town.’ (Bohnhoff, 1986, 114)

Bohnhoff provides example (20) to show that the LD pronoun must be bound by the *closest* LD antecedent. The verbs *say* and *tell* both introduce an LD domain, since their complements are indirect quotations. However, example (20) is not ambiguous; the antecedent of the LD pronoun must be Moses, the closest eligible LD antecedent, and not Mother:

- (20) Nà'á ∅ 'òd bà'á [Múúsa bà ∅ 'ò [bà bih híí
 Mother_i (she_i) says.to Father Moses_j that (he_j) says that he.LD_{j,*i} wants
 lààlí kaalí]]
 to.go to.town
 'Mother_i tells Father that Moses_j says that *she_i/he_j wants to go to town.'
 (Bohnhoff, 1986, 118)

Unlike the SUBORD pronoun, whose appearance is restricted to subject position, the LD pronoun may appear as a subject, object, or possessor within the LD domain. In example (21), the object of the subordinate verb *refuses* is a LD pronoun whose antecedent is the subject of the matrix verb *attack*:

- (21) Yòṣb vḥ kó 'à'á [bà háj bi
 ancestor.spirits_i they.PRON_i attack grandmother because.she refuses them.LD_i
 nannè]
 food
 'Ancestor spirits_i, they_i attack grandmother because she refuses them_i food.'
 (Bohnhoff, 1986, 115)

In example (22), both the subordinate subject and the possessor of the object are LD pronouns:

- (22) vḥ híí [bi mbàà kan yúú bìi nu]
 they.PRON_i want they.LD_i sit with head their.LD_i CM
 'They_i want to sit with their_i head.' (= 'They want to be independent.')
- (Bohnhoff, 1986, 116)

The LD domain is not defined by properties typically associated with logophoricity, though its roots are likely based in an earlier logophoric system: Bohnhoff (1986, 112) observes that clauses constituting the LD domain “all may have been derived from underlying quotes”. Culy (1997) discusses the extension of logophoric marking from typical logophoric domains such as reported speech, thought, or perception to adjuncts such as purpose clauses and causal clauses, and proposes that this is the result of grammaticisation of an original logophoric system; this seems to be the case for Yag Dii. As Bohnhoff (1986, 113) notes, the LD domain “does not seem to be limited to contexts containing a performative verb, nor to a desiderative context, nor do such pragmatic/semantic notions as source/receiver of the information seem to govern the use of the series”. Nonsubject antecedents of LD pronouns are not permitted. Further, constructions that seem to have very similar

meanings vary as to whether they introduce a LD domain: for example, causal constructions introduced by *ka/bà* constitute a LD domain (example 21), while causal constructions with *moo*, as in example (23), do not:

(23) Cause with *moo*:

V_H yaa bi mà”_q lùù ’ú, [moo v_H ’_q
 they.PRON_i come, they.LD_i grab.him leave CM, because they.PRON_i say
 bà y_qq_qè]
 that.he crazy.CM
 ‘They_i came to take him away, because they_i said that he’s crazy.’

(Bohnhoff, 1986, 115-116)

The PRON pronoun is used as the subject of the subordinate clause *because they say that he’s crazy* because this is neither a LD domain nor a SUBORD domain: only causal constructions with *ka/bà* allow LD or SUBORD pronouns, not causal constructions with *moo*.

Morphosyntactically, the LD domain is usually marked either by the subordinator/complementiser *bà* or by the presence of a particular lexical predicate in the immediately higher clause; Culy (1997) discusses the importance of marking by particular complementisers in defining the logophoric domain in many languages. Subordinate purpose clauses seem to constitute an exception to this generalisation, since they do not contain special marking to indicate the LD domain, and need not appear with a particular predicate in the immediately higher clause; it may be that these are positionally encoded:

(24) Subordinate purpose clause:

Bà’á ∅ n_ə’_{əy} hághá [bi hò púggì]
 Father_i (he_i) bends down he.LD_i sees animal.CM
 ‘Father bends down to see the animal.’

(corrected version of Bohnhoff, 1986, 114)

Example (18) contains a subordinate desiderative clause, signalled by the presence of the verb ‘want’ in the matrix clause. Example (21) contains a causal adjunct with the subordinator/complementiser *bà*. Indirect quotations are also introduced by *bà*:

(25) Indirect quotation:

Bà’á ∅ ’_q [bà bíń láá kòddí]
 Father_i (he_i) says that he.LD_i.will go forest.to
 ‘Father_i says that he_i will go to the forest.’

(corrected version of Bohnhoff, 1986, 114)

In fact, indirect discourse may consist of a number of clauses, as in (26):

(26) ... vɛ ɔd Yésù: “Bà’á, í ní vóó 0 bà’ vó ya, moo òò
 they say-to Jesus: Sir, the.one elder_i our (he_i) send us come, so we
 ɔd ví bih màn bà vín dón kí bìlí né.
 say.to you that.he.LD_i is.worthy that you enter house his.LD_i.in NEG.

Mo wòò nò mà, bih yaan kan fọ́ọ̀ bì ní yẹ̀ nò.
 for that CM then that.he.LD_i come.NEG with body his.LD_i NEG here CM

Àmáa bà ì ọ̀ moo ɔɔɔ́ dágá sị́, nán bì yẹ̀ bàn zàà ọ̀.
 but that you say word cheek one only man his.LD_i this that.he heals CM

Moo bi áń, bà bih kíd í ní bì vɛ tógá,
 for he.LD_i too that he.LD_i hear.to the.one elder his.LD_i plural ear.CM
 bà bih dí kan sọ́ze bì bà kuu kíd bí tóg máa vɛ
 that he.LD_i is.there with soldier his.LD_i that they hear him.LD_i ear this pl
 áń.
 too

Bà ì ɔd dágá: “Àm làà ń” tée, bàn làà.
 that if.he.2LD_i say-to one: You go CM if, that.he go

Bà ì ɔd tóó: “Àm yaa ń” tée, bàn yaa.
 that if.he.2LD_i say-to another: You come CM if, that.he come

Bà ì ɔd nán ba’ad bì: “Àm kó hẹ̀n yẹ̀ nò” tée
 that if.he.2LD_i say-to man work his.LD_i: You do thing this CM, then
 bàn kó ń.
 that.he do CM

‘...they say to Jesus: “Sir, our elder_i has sent us to you, to say to you that he_i isn’t worthy for you to enter his_i house. That’s why he_i hasn’t come here himself_i. But even if you simply say a single word, his_i worker will be healed. For he_i too says that he_i takes orders from his_i superiors; that he_i has his_i soldiers that take orders from him_i, too. That if he_i says to one: “Go!”, then he will go. That if he_i says to another: “Come!”, he will come. That if he_i says to his_i worker, “Do this!”, he will do it.’ (L. Bohnhoff, p.c.)

Each clause in these multi-clause indirect discourse segments is marked with the subordinator/complementiser *bà*. We analyse these examples as subordination to an unpronounced main clause predicate, with only the subordinate LD domain realised. An alternative analysis might treat these in terms of a morphologically marked main-clause LD domain interpreted as indirect discourse (see Dimmendaal 2001 for more discussion). Under the second analysis, indirect discourse clauses as in (25) and (26) would differ from the other LD domains in that no syntactic relation would be required between the LD antecedent (which would not be syntactically present in the clause) and the LD pronoun; instead, indirect discourse would have to be analysed specially, as true logophoricity, different from the other

syntactically defined instances of the LD domain. For uniformity, and in the absence of evidence that the conditions governing these multi-clause examples are different from the other examples, we assume that subordination is involved, with an unpronounced main-clause predicate.

An orthogonal issue related to determination of the LD antecedent is raised by Bohnhoff's claim that the LD antecedent must be the *pronoun* subject of the immediately higher clause, which, on his analysis, is always present but sometimes unpronounced. This would make Yąg Dii a pronoun-incorporating language in the sense of Jelinek (1984): on this view, the subject of every clause is a (possibly unpronounced) pronominal, and what appears to be a full non-pronominal subject is treated as a dislocated topic or apposition to the pronoun. In fact, it is likely that full non-pronominal subject phrases are best treated as subjects and not topics or appositions, with unpronounced pronominal subjects posited only when there is no overt subject phrase (see Austin and Bresnan 1996 for a thorough discussion of differences between these two analyses and arguments against the Jelinek view). The choice between the two analyses does not affect the analysis of overt pronouns in the language, and so for clarity and consistency with Bohnhoff's presentation of examples, we include unpronounced pronouns (represented as \emptyset) in some examples, though we do not intend this as a claim that unpronounced pronominal forms are actually present in the structure.

4 Subordinate clause LD pronouns: 2LD

Our primary interest is a fourth series of pronouns which we will label 2LD, characterised above as the “very long-distance” series, as shown in Table 5. Like the LD series, 2LD pronouns appear in the LD domain and must corefer with the LD antecedent. Like the SUBORD series, they are used only in subject position of certain subordinate clauses within the LD domain. There are no 2LD object or possessive pronouns. As shown in examples (3) and (4), the 2LD pronoun neither requires nor disallows a coreferential pronoun in the intervening clause in the LD domain. 2LD is, then, an exceptionally long-distance anaphor, whose binding conditions seem to be exclusively nonlocal: it must appear as the subject of an embedded clause within the LD domain, and it must corefer with a subject outside the LD domain, at least two clauses distant, without imposing any binding requirements in the intervening clause.

There are clear morphological parallels between the 2LD series and the SUBORD series, as is evident from inspection of the paradigms in Table 5. Their distribution is also closely related; indeed, Bohnhoff (1986, 123) states that “in the same way that *'àn* [SUBORD] subjects are used instead of *mì* [PRON] subjects in certain clauses, so *'i* [2LD] subjects occur instead of *bi* [LD] subjects in (some of) those same grammatical contexts”. In fact, in light of additional data unavailable to Bohnhoff at the time the article was written, constraints on the distribution of 2LD seem to be very close or identical to those for the SUBORD subject pronoun: 2LD

	PRON	SUBORD	LD	2LD
1.sg	-n/mí	'àn	bi	'àn
1incl.dual	ba	ba	bi	'aa
2.sg	-m/mó	'àm	bi	'ii
3.sg	∅	'à	bi	'ii
1excl.pl	vó	'òo	bi	'òo
1incl.pl	ba...ví	ba...ví	bi	'aa...ví
2.pl	ví	'ì	bi	'ii
3.pl	vũ	'ùu	bi	'ii

Table 5: Subject pronouns of all four pronoun types, from Bohnhoff (1986, 107,113,120).

appears as the subject of a relative clause, temporal/locative/conditional clause, concessive clause, indirect order, purpose clause, or causal adjunct (though there are no available data that allow a determination of whether 2LD is limited to causal adjuncts introduced by *ka* or *bà*, as in the case of the SUBORD pronoun series). Bohnhoff does not provide examples of 2LD as the subject of a purpose clause or a causal adjunct, though he states that it can appear there. In fact, however, he notes (p. 121) that in some clauses, either the LD or the 2LD pronoun may appear: “Initial concessive and cause clauses for many speakers simply retain the *bi* [LD] forms, although some examples of *'ii* [2LD] may also be heard.” Example (27) shows 2LD as the subject of a relative clause within the LD domain:

- (27) ... ví ɔd í kíi àgà: “Àkàw ∅ ɔ [lig [bà ii
you say.to the.one house self: “Teacher_i (he_i) say house that he.2LD_i
lá hẹn lálí páska kan waa duulí bìi vũ wulí máa] bà
eat thing eating Easter with child following his.LD_i plural there when, that.it
dì tẹlá?”
is.there where?”
‘... you’ll ask the house owner: “The teacher asks, where is the house in
which he.2LD will eat the Easter meal with his disciples?”’ (L. Bohnhoff,
p.c.)

In (28), 2LD is the subject of a temporal adjunct clause in initial position within the LD domain:

- (28) ∅ 'ò [[sè`èy bà 'ii là fíí ya babbí tée] bà
(she_i) says time that she.2LD_i goes returns comes field.from when, that
bíí dèè gbokìi]
she.LD_i.will cook pigeon
‘She_i said that when she.2LD_i returned from the field, she.LD_i would cook
the pigeon.’

(Bohnhoff, 1986, 121)

Another example of a temporal/locative/conditional clauses with 2LD is given in (3), a concessive clause with 2LD subject is given in (4), and an indirect order with 2LD subject is given in (1).

5 Standard binding theory and 2LD

The following generalisations govern the distribution of the four Yag Dii pronoun series:

- (29) **PRON**: can bear any grammatical function, except for subject in SUBORD domain; noncoreferent with LD antecedent if in LD domain
- SUBORD**: must appear as subject in SUBORD domain; noncoreferent with LD antecedent if in LD domain
- LD**: must appear in LD domain; can bear any grammatical function (except for subject in SUBORD domain within LD domain); coreferent with LD antecedent
- 2LD**: must appear as subject in SUBORD domain within LD domain; coreferent with LD antecedent

The status of the parenthesised portion of the condition on LD reflects the uncertainty discussed at the end of the previous section: in at least some SUBORD clauses within the LD domain, either LD or 2LD can appear, but it is not clear whether the LD and 2LD pronouns are in free variation in all SUBORD domains.

It is not possible to capture the very long-distance nature of the binding constraints on the 2LD pronouns by means of standard binding-theoretic constraints. 2LD does not behave like a standard pronominal, in that it does not obey only a negative binding condition such as Binding Condition B (a pronominal must be free in its governing category: Chomsky 1981). 2LD pronouns require an antecedent in the same sentence, unlike pronominals, and cannot appear without an antecedent (setting aside the extended indirect discourse examples, which must be marked with *bà* and which we have proposed to treat as involving an unpronounced main clause subject and predicate). Of course, 2LD does not behave like a standard anaphor either, since it is not locally bound. To ensure the presence of a nonlocal antecedent, we might attempt to state the binding requirements for 2LD as a combination of a local noncoreference requirement (as we expect to find with pronominals) and a nonlocal coreference requirement (as we find with long-distance reflexives): that is, 2LD would be an overt pronominal anaphor, which must be locally free but bound in a larger domain, as originally suggested for Malayalam *taan* by Mohanan (1981) (see also Dalrymple 1993 and Kiparsky 2002). Mohanan (1981) provides examples (30a) and (30b) to show that *taan* must be bound, and example (30c) to show that the binder of *taan* may not be a coargument of the same predicate – that is, *taan* must be bound within the sentence in which it appears, but may not be locally bound:

$$(34) \left[\begin{array}{l} \text{GF}_{\text{ant}} \text{ [ANTECEDENT]} \\ \dots \quad \text{GF}^* \quad \dots \quad \text{GF}_{\text{pro}} \text{ [PRONOUN]} \\ \text{DOES NOT PASS} \\ \text{THROUGH AN F-} \\ \text{STRUCTURE WITH A} \\ \text{TENSE ATTRIBUTE} \end{array} \right]$$

For more discussion of LFG's binding theory, see Dalrymple (1993), Bresnan (2001), and Asudeh (2004).

For the Yag Dii LD pronouns, we propose that the clause that contains the LD domain and the LD antecedent – the **binding domain** for the LD pronoun – is marked with the attribute-value pair $\langle \text{LD-ANT}, + \rangle$. This marking is enforced by the predicate or construction which defines the subordinate domain as an LD domain (the main clause predicate whose complement is an indirect quotation, subordinate desiderative, or indirect order; the *ka* or *ba* marking on causal adjuncts; or the c-structure rule marking a subordinate clause as a purpose clause):

$$(35) \left[\begin{array}{l} \text{SUBJ [LD antecedent]} \\ \text{LD-ANT } + \\ \text{COMP [...LD pronoun ...]} \end{array} \right]$$

It is important to note that the LD-ANT domain is not what we have been calling the LD domain; instead, it is the clause containing both the LD domain and the LD antecedent. We can now state the requirements for the LD pronouns with some minimal modifications to the form of the standard binding equations:

$$(36) (\uparrow_{\sigma} \text{ ANTECEDENT}) = ((\quad \text{GF}^* \quad \quad \text{GF}_{\text{pro}} \quad \uparrow) \quad \text{SUBJ})_{\sigma}$$

$$\quad \quad \quad \neg(\rightarrow \text{LD-ANT}) \quad \neg(\leftarrow \text{LD-ANT}) \quad (\leftarrow \text{LD-ANT})$$

$$\quad \quad \quad 1 \quad \quad \quad 2 \quad \quad \quad 3$$

1. the clause containing the LD antecedent is the **smallest** clause marked with LD-ANT that also contains the pronoun (the path through the binding domain to the LD pronoun may not pass through a clause with LD-ANT marking);
2. the clause containing the LD pronoun cannot have LD-ANT marking (LD pronouns are not bound by a clausemate);
3. the ANTECEDENT of the pronoun is the SUBJ of a clause with LD-ANT marking.

This combination of constraints encodes the binding requirements for the LD pronoun, and enforces an appropriate degree of nonlocality. The antecedent of the LD pronoun is not a clausemate (since the antecedent must appear in an LD-ANT-marked clause, and the LD pronoun may not appear in an LD-ANT-marked clause): the antecedent must be the subject of the LD-ANT-marked clause which properly contains the LD domain in which the LD pronoun appears.

7 The SUBORD requirement

We propose that the SUBORD domain is marked as $\langle \text{SUBORD}, + \rangle$.

$$(37) \begin{bmatrix} \text{SUBJ} & [\text{SUBORD PRONOUN}] \\ \text{SUBORD} & + \end{bmatrix}$$

SUBORD pronouns require the presence of that feature, while PRON pronouns are incompatible with that feature. As with the LD feature, this marking is enforced by the main clause predicate or construction defining the domain as a SUBORD domain. We can then enforce the requirement for SUBORD pronouns to appear as subjects of SUBORD-marked clauses with the existential constraint in (38):

$$(38) ((\text{SUBJ } \uparrow) \text{SUBORD})$$

8 Constraints on 2LD

We now turn to the question of the binding requirements of 2LD. It is tempting to analyse 2LD as just the LD version of SUBORD: we have seen that 2LD pronouns appear in subject position of SUBORD clauses, just like SUBORD pronouns, and are bound by the LD antecedent, just like LD pronouns. However, more needs to be said in cases of overlap between the SUBORD and LD domains. Some clauses are both SUBORD and LD:

$$(39) \text{V}_{\text{H}} \quad \text{híí} \quad [\text{'ù}\text{H} \quad \text{làà kaalí}]$$

they.PRON_i want they.SUBORD_j go town.to
‘They want others to go to town.’ (Bohnhoff, 1986, 114)

$$(40) \text{V}_{\text{H}} \quad \text{híí} \quad [\text{bi} \quad \text{làà kaalí}]$$

they.PRON_i want they.LD_i go town.to
‘They want to go to town.’ (corrected version of Bohnhoff, 1986, 113)

The complement of the verb *want* is an indirect order, which is in the overlap between the SUBORD and LD domains. Both SUBORD pronouns (ex. 39) and LD pronouns (ex. 40) are allowed in this domain. This means that the subordinate clause subject position in these examples is SUBORD-marked, and the antecedent is in an LD-ANT-marked clause. If 2LD were simply required to appear in a SUBORD domain and to be bound by the LD antecedent, we would expect the 2LD pronoun to appear in (40). This is not possible, however: the LD pronoun and not the 2LD pronoun appears here.

It might appear that we could get around this problem by claiming that a clause cannot be LD and SUBORD at the same time, but this would lead to the incorrect prediction that LD and SUBORD pronouns cannot appear in the same clause. We do find LD and SUBORD pronouns in the same clause, however; an example is given in (41).

- (41) Bábàḡam Ø vì [moo 'èṅ pẹ́n vùn tíd waa bìì gbó
 Rabbit_i (he_i) asks for what first they.PRON_j hold child his.LD_i leave
 mammé máalá?] ['í yè máa, bà vùn sóó 'ú] [bà
 water.in Q? this here focus that they.PRON_j fake CM, that
 'èḡ sòò waa bìì pú bì dóg 'yà'ạ yè
 they.SUBORD_j.must look.for child his.LD_i give him.LD_i go.up now here
 no.]
 CM
 'Rabbit_i asks why they (Boar) held his_i child and let it fall in the water? (He
 says) that they faked it, that **they.SUBORD must look for his.LD_i child
 and give it to him.LD_i now!** (Bohnhoff, 1986, 118-119)

The generalisation seems to be that there is an 'exclusion zone' for 2LD in the topmost clause in the LD domain. 2LD is an exclusively long-distance anaphor, and its antecedent must appear at least 2 clauses away.

We propose to introduce an additional feature marking the top clause of the LD domain as an 'exclusion zone' for 2LD, and constrain the 2LD pronoun so as to prevent it from appearing there. We will use the feature LD-DOMAIN to mark the exclusion zone for 2LD in the LD domain:

$$(42) \left[\begin{array}{l} \text{SUBJ [LD antecedent]} \\ \text{LD-ANT +} \\ \text{COMP } ld \left[\begin{array}{l} \text{SUBJ ...} \\ \text{LD-DOMAIN +} \\ \text{COMP } subord \left[\begin{array}{l} \text{SUBJ [2LD]} \\ \text{SUBORD +} \end{array} \right] \end{array} \right] \end{array} \right]$$

Notice that this marking is still purely local to the predicate or construction defining the LD domain: the matrix clause containing the LD domain is marked with LD-ANT, and the LD domain itself is marked with LD-DOMAIN. No marking of more deeply embedded clauses or constraints involving purely nonlocal relations are required. We can now state the binding constraints for 2LD as follows:

- (43) Binding constraints for 2LD:

$$(\uparrow_{\sigma} \text{ ANTECEDENT}) = ((\text{GF}^* \text{ SUBJ } \uparrow) \text{ SUBJ })_{\sigma}$$

$$\begin{array}{ccc} \neg(\rightarrow \text{LD-ANT}) & \neg(\leftarrow \text{LD-DOMAIN}) & (\leftarrow \text{LD-ANT}) \\ 1 & (\leftarrow \text{SUBORD}) & 3 \\ & \neg(\leftarrow \text{LD-ANT}) & \\ & 2 & \end{array}$$

1. As with LD, the clause containing the 2LD antecedent is the **smallest** clause marked with LD-ANT that also contains the pronoun.
2. • Like the LD pronoun, the clause containing the 2LD pronoun cannot have LD-ANT marking (2LD pronouns are not bound by a clausemate).

- Like SUBORD pronouns, 2LD must appear as the SUBJ of a clause with SUBORD-marking.
- To enforce the nonlocal relation between 2LD and its antecedent, the 2LD pronoun cannot appear in a clause with LD-DOMAIN marking (i.e., it cannot appear in the highest clause in the LD domain).

3. Like the LD pronoun, the ANTECEDENT of the 2LD pronoun is the SUBJ of a clause with LD-ANT marking.

On this analysis, the nonlocal nature of 2LD's binding requirements fall out from a combination of purely locally specified features.

9 An alternative analysis

As suggested by Louisa Sadler (p.c.), an alternative way of analysing the binding requirements of 2LD is to directly encode the nonlocal nature of the binding relation. On this analysis, the binding equation associated with 2LD would be:

(44) Alternative binding equation for 2LD (to be rejected):

$$\begin{array}{ccccccc}
 (\uparrow_{\sigma} \text{ ANTECEDENT}) = & & & & & & \\
 ((& \text{GF}^* & & \text{GF} & & \text{GF} & & \text{SUBJ} & & \uparrow) & & \text{SUBJ} &)_{\sigma} \\
 \neg(\rightarrow \text{ LD-ANT}) & & \neg(\rightarrow \text{ LD-ANT}) & & 3 & & \neg(\leftarrow \text{ LD-ANT}) & & & & (\leftarrow \text{ LD-ANT}) \\
 1 & & 2 & & & & (\leftarrow \text{ SUBORD}) & & & & 5 \\
 & & & & & & 4 & & & &
 \end{array}$$

This constraint resembles the requirements for LD, except that the path delimiting the binding domain must contain at least three grammatical functions (GF GF SUBJ): this directly reflects the fact that there must be at least one clause intervening between 2LD and its antecedent. The 2LD pronoun is required by this constraint to appear in the following environment:

$$(45) \left[\begin{array}{l} \text{SUBJ [LD antecedent]} \\ \text{LD-ANT +} \\ \dots \text{GF}_1^* \dots \end{array} \right] f1 \left[\begin{array}{l} f2 \left[\begin{array}{l} \text{GF}_2 \\ \text{GF}_3 \end{array} \right] f3 \left[\begin{array}{l} \text{GF}_3 \\ \text{GF}_4 \end{array} \right] f4 \left[\begin{array}{l} \text{SUBJ [2LD]} \\ \text{SUBORD +} \end{array} \right] \end{array} \right] \left. \right]$$

The Kleene star in the portion of the path marked 1 means that that portion can be empty, so it will often be the case that the f-structure labelled $f1$ and the f-structure labelled $f2$ will be the same, with only one clause (the exclusion zone $f3$) separating 2LD from its antecedent. We can explicate the binding constraints in (44) as follows:

1. As with LD, the clause containing the 2LD antecedent is the **smallest** clause marked with LD-ANT that also contains the pronoun. This is enforced for the f-structure labelled f_2 by the constraint on the portion of the path marked 1.
2. The f-structure labelled f_3 may not be LD-ANT-marked. This is enforced by the constraint on the portion of the path marked 2. This component of the path is obligatory.
3. The portion of the path marked 3 is also obligatory.
4. The 2LD pronoun must bear the grammatical function SUBJ. The f-structure marked f_4 may not have LD-ANT-marking, and it must have SUBORD marking. This is enforced by the portion of the path marked 4.
5. As with the LD pronoun, the LD antecedent must be a SUBJ, and it must be in a clause with LD-ANT-marking.

This analysis has the advantage of requiring fewer features: only LD-ANT marking and SUBORD marking are required, and we do not need to appeal to additional features such as LD-DOMAIN. However, this advantage comes at the cost of allowing a nonlocal path with a minimal length of three grammatical functions. Given that this construction provides the only evidence we know of for grammatical nonlocality, we would prefer not to go down this slippery slope. Instead, we propose that nonlocality of this nature is forbidden in grammatical description, and that functional uncertainty paths are constrained by a general Locality Principle:

- (46) **Locality Principle:** Paths in functional uncertainty expressions are of length zero or more (Kleene star) or of length one or more (Kleene plus); no other options are available in grammatical description.

Given this Locality Principle, the alternative binding constraint for 2LD presented in (44) is disallowed.

10 Conclusion

Yag Dii presents a complicated picture for theories of anaphoric binding. The distributions of the PRON, SUBORD, and LD pronouns are not unexpected, given the general form of binding equations and the ability to mark domains with information about their syntactic properties. We have proposed that the distribution of the 2LD pronoun can be stated in local terms, by introducing additional features controlling the appearance of 2LD vs. LD at multiple levels of structure – governing a nonlocal relation by introducing a combination of local features to create a local ‘exclusion zone’ for 2LD. Our analysis obeys the Locality Principle, which we propose as a general principle for functional uncertainty paths in grammatical dependencies.

References

- Asudeh, Ash. 2004. *Resumption as Resource Management*. Ph. D.thesis, Stanford University.
- Austin, Peter and Bresnan, Joan. 1996. Non-Configurationality in Australian Aboriginal Languages. *Natural Language and Linguistic Theory* 14(2), 215–268.
- Bohnhoff, Lee E. 1986. Yag Dii (Duru) Pronouns. In Ursula Wiesemann (ed.), *Pronominal Systems*, pages 103–129, Tübingen: Gunter Narr Verlag.
- Bresnan, Joan. 2001. *Lexical-Functional Syntax*. Oxford: Blackwell.
- Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris Publications.
- Culy, Christopher D. 1997. Logophoric Pronouns and Point of View. *Linguistics* 35, 845–859.
- Dalrymple, Mary. 1993. *The Syntax of Anaphoric Binding*. CSLI Lecture Notes, number 36, Stanford: CSLI Publications.
- Dimmendaal, Gerrit J. 2001. Logophoric Marking and Represented Speech in African Languages as Evidential Hedging Strategies. *Australian Journal of Linguistics* 21(1), 131–157.
- Fitzpatrick, Justin M. 2002. On Minimalist Approaches to the Locality of Movement. *Linguistic Inquiry* 33(3), 443–463.
- Haiman, John and Munro, Pamela. 1983. Introduction. In John Haiman and Pamela Munro (eds.), *Switch Reference and Universal Grammar*, Amsterdam: John Benjamins.
- Jelinek, Eloise. 1984. Empty Categories, Case, and Configurationality. *Natural Language and Linguistic Theory* 2(1), 39–76.
- Kiparsky, Paul. 2002. Disjoint Reference and the Typology of Pronouns. In Ingrid Kaufmann and Barbara Stiebels (eds.), *More Than Words: A Festschrift for Dieter Wunderlich*, *Studia Grammatica*, No. 53, pages 179–226, Berlin: Akademie Verlag.
- Langdon, Margaret and Munro, Pamela. 1979. Subject and (Switch-)Reference in Yuman. *Folia Linguistica* 32, 321–344.
- Manzini, Maria Rita and Wexler, Kenneth. 1987. Parameters, Binding Theory, and Learnability. *Linguistic Inquiry* 18(3), 413–444.
- Mohanan, K. P. 1981. *Grammatical Relations and Anaphora in Malayalam*. Masters Thesis, MIT.
- Sag, Ivan A. 2008. Feature Geometry and Predictions of Locality. In Anna Kibort and Greville G. Corbett (eds.), *Features: Perspectives on a Key Notion in Linguistics*, Oxford: Oxford University Press.