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

Yag 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 Yag Dii appear to run counter to these standardly accepted generalisations. There are several series of pronouns in Yag 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:\footnote{The form glossed CM is a clause-final particle.}

1I 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.

†The form glossed CM is a clause-final particle.
(1) Nán ba’ad ø ’qı [ moo ’čn da bi töö bà ka vi bi man work (he,)] say for what friend his.LD other that sb-he, ask him.LD [ bà ’ii stíwá ’ulá]?

The worker, asked why his.LD friend asked him.LD that he.2LD repay him.2LD i did remodel the IOU.’ (corrected version of Bohnhoff, 1986, 119)

(2) 

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 gà [ [ sëy ‘ii làà teé] bà bín hô høn Papa, (he,) knows time he.2LD goes when, that he.LD will see thing Múusá wòó]

Moses his
‘Papa, knows that when he.2LD goes, he.LD’ll see Moses’s thing.’

(L. Bohnhoff, p.c.)

(4) Bà’á ø gà [ [ kóó ‘ii lúu ní s’i] bà míń hô høn Papa, (he,) knows time he.2LD leave NEG even, that I.will see thing Múusá wòó]

Moses his
‘Papa, knows that even if he.2LD doesn’t leave, I’ll see Moses’s thing.’

(L. Bohnhoff, p.c.)

(5) 

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 şohñi hegai ceeoj c ’am şoşa.
    that woman 3-perf hit that man SS cry
    ‘The woman, hit the man and she, cried.’

b. Hegai ’uuvi ’a-t ’am şohñi hegai ceeoj ku-t (hegai ceeoj) ’am şoşa.
    that woman 3-perf hit that man DS that man cry
    ‘The woman hit the man, and he, (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

Yağ 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 mı́ 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 1incl.pl form ba...vi 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
Table 1: PRON and SUBORD pronouns, from Bohnhoff (1986, 107, 109, 110).

<table>
<thead>
<tr>
<th></th>
<th>PRON subject</th>
<th>PRON object</th>
<th>PRON possessive</th>
<th>SUBORD subject</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.sg</td>
<td>-n/mí</td>
<td>-n/mí</td>
<td>míí</td>
<td>'ân</td>
</tr>
<tr>
<td>1incl.dual</td>
<td>ba</td>
<td>ba</td>
<td>bàn</td>
<td>ba</td>
</tr>
<tr>
<td>2.sg</td>
<td>-m/mó</td>
<td>-m/mó</td>
<td>móó</td>
<td>'ân</td>
</tr>
<tr>
<td>3.sg</td>
<td>ø</td>
<td>-wu</td>
<td>wòò</td>
<td>`à</td>
</tr>
<tr>
<td>1excl.pl</td>
<td>vó</td>
<td>vó</td>
<td>vóó</td>
<td>'òó</td>
</tr>
<tr>
<td>1incl.pl</td>
<td>ba...ví</td>
<td>ba ví</td>
<td>bà ví</td>
<td>ba...ví</td>
</tr>
<tr>
<td>2.pl</td>
<td>ví</td>
<td>ví</td>
<td>víí</td>
<td>'í</td>
</tr>
<tr>
<td>3.pl</td>
<td>vû</td>
<td>vû</td>
<td>vòò</td>
<td>'íí</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>PRON:</th>
<th>main clauses</th>
<th>subordinate clauses</th>
</tr>
</thead>
<tbody>
<tr>
<td>imperfective-factative, perfective-factative</td>
<td>indirect quotation, comparison clauses, causal adjuncts (“because...”) introduced by moo, ‘until’ adjuncts</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SUBORD:</th>
<th>main clauses</th>
<th>subordinate clauses</th>
</tr>
</thead>
<tbody>
<tr>
<td>imperfective-hortative</td>
<td>indirect order, relative clause, temporal/locative/conditional clause, purpose clause, concessive clause, causal adjunct introduced by ka or bà</td>
<td></td>
</tr>
</tbody>
</table>

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 làa kaalí
you.PRON go to.town
‘You go to town.’ (Bohnhoff, 1986, 107)

(8) Indirect quotation:

... bà mó làa kaalí
that you.PRON go town.to
‘... that you go to town.’ (Bohnhoff, 1986, 107)
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:

'Ành   là̀ kaalí
you.SUBORD.must go town.to
‘Go to town!’

(Bohnhoff, 1986, 108)

(12) Temporal/locative/conditional:

Tòw/sè’èy/ya  ành   là̀ kaalí  tée
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) ‘Ành   là̀ 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  ð   ’qd  bà’á  [[ sè’èy  bà   ’à  fií ya
Motherì (sheì) says.to Fatherj time that he.SUBORDj returns comes
bábí  tée]  bà  bíí  dé̀ ì  dubbi] field.from then] that she.LOGì cook yam.CM
‘Motherì says to Fatherj that when hej returns from the field, sheì will cook
the yams.’

(corrected version of Bohnhoff, 1986, 122)
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’a  `∅ ’q’d bà’a [] sè’èy bà ’à fîî ya
Motherₐ (sheⱼ) says to Fatherⱼ time that heⱼ returns comes babbi tće bà ’à dòò ḅubbì field from then] that heⱼ.SUBORD cook yam.CM
‘Motherₐ says to Fatherⱼ that when heⱼ returns from the field, heⱼ should cook the yams.’ (corrected version of Bohnhoff, 1986, 122)

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

<table>
<thead>
<tr>
<th></th>
<th>PRON subject</th>
<th>PRON object</th>
<th>PRON possessive</th>
<th>SUBORD subject</th>
<th>LD subject</th>
<th>LD object</th>
<th>LD possessive</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.sg</td>
<td>-n/mí</td>
<td>-n/mí</td>
<td>míí</td>
<td>'ãh</td>
<td>bí</td>
<td>-n/mí</td>
<td>míí</td>
</tr>
<tr>
<td>1incl.dual</td>
<td>ba</td>
<td>ba</td>
<td>bàá</td>
<td>ba</td>
<td>bí</td>
<td>ba</td>
<td>bàá</td>
</tr>
<tr>
<td>2.sg</td>
<td>-m/mó</td>
<td>-m/mó</td>
<td>móó</td>
<td>'ãh</td>
<td>bí</td>
<td>bi</td>
<td>bii</td>
</tr>
<tr>
<td>3.sg</td>
<td>Ø</td>
<td>-wú</td>
<td>wóó</td>
<td>'à</td>
<td>bí</td>
<td>bi</td>
<td>bii</td>
</tr>
<tr>
<td>1excl.pl</td>
<td>vó</td>
<td>vó</td>
<td>vóó</td>
<td>'ôo</td>
<td>bí</td>
<td>vó</td>
<td>vóó</td>
</tr>
<tr>
<td>1incl.pl</td>
<td>ba...ví</td>
<td>ba...ví</td>
<td>bàá...ví</td>
<td>ba...ví</td>
<td>bí</td>
<td>ba...ví</td>
<td>bàá...ví</td>
</tr>
<tr>
<td>2.pl</td>
<td>ví</td>
<td>ví</td>
<td>víí</td>
<td>'i</td>
<td>bí</td>
<td>bi</td>
<td>bii</td>
</tr>
<tr>
<td>3.pl</td>
<td>vu</td>
<td>vu</td>
<td>vôó</td>
<td>'îu</td>
<td>bí</td>
<td>bi</td>
<td>bii</td>
</tr>
</tbody>
</table>

Table 4: Distribution of LD and SUBORD subject pronouns.

<table>
<thead>
<tr>
<th></th>
<th>main clauses</th>
<th>subordinate clauses</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUBORD only:</td>
<td>imperfective-</td>
<td>relative clause,</td>
</tr>
<tr>
<td></td>
<td>hortative</td>
<td>concessive clause,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>temporal/locative/</td>
</tr>
<tr>
<td></td>
<td></td>
<td>conditional clause</td>
</tr>
<tr>
<td>LD only:</td>
<td>indirect quotation,</td>
<td>subordinate</td>
</tr>
<tr>
<td></td>
<td>desiderative</td>
<td>adject introduced by ka or bà</td>
</tr>
<tr>
<td>both SUBORD and LD allowed:</td>
<td>indirect order, purpose clause, causal</td>
<td></td>
</tr>
</tbody>
</table>

Bohnhoff (1986, 112) shows that the LD domain consists of indirect quotations, subordinate desiderative clauses, indirect orders, purpose clauses, and causal adjectives 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.

(18) Vú hî [ bi làa kaalí]  
they.PRON₁ want they.LD₂ go town.to  
‘They want to go to town.’ (corrected version of Bohnhoff, 1986, 113)

(19) Vú hî [ ‘iùu làa kaalí]  
they.PRON₁ want they.SUBORDᵵᵳ 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) Ná’á ṃó ḍá bá’á [Múúša bá ṃó ḍá bá bín híí]
Mother, (she) says.to Father Moses that (he) says that he.LD wants to go to town

‘Mother, tells Father that Moses says that *she/he* 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) Yógb yú kó ’á’á [bá hà́j bi ancestor.spirits, they.PRON, attack grandmother because.she refuses them.LD food]

‘Ancestor spirits, they attack grandmother because she refuses them food.’
(Bohnhoff, 1986, 115)

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

(22) yú híí [bí mbà’á kan yúú bíí nu]
they.PRON, want they.LD sit with head their.LD CM

‘They want to sit with their 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 Yág 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 kalbà constitute a LD domain (example 21), while causal constructions with moo, as in example (23), do not:

(23) Cause with moo:

\[ \text{Vù yaa bi mà”ñ lùù ’ú, [ moo vu ’ê] they.PRON come, they.LD grab.him leave CM, because they.PRON say bà yàŋŋë] that.he crazy.CM}

‘They came to take him away, because they 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 kalbà 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:

\[ \text{Bà’â Ÿ no’øy hághá [ bi hê pûggí] Father, (he.î) bends down he.LD 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:

\[ \text{Bà’â Ÿ ’ê [ bà bíî láá kôddî] Father, (he.î) says that he.LD.î will go forest.to}

‘Father says that he 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) ... và qo đ Yêsù: “Bà’a, i nii vô ô bà’ vó ya, moo òo they say-to Jesus: Sir, the.elder, our (he,1) send us come, so we qo vì bìi màn bà vô’n dân kí biil né. say.to you that.he.LD1 is.worthy that you enter house his.LD..in NEG.

Mo vồo nò mà, bìi yaa kan fôô bìi ní yê nò. for that CM then that.he.LD1 come.NEG with his.LD1 NEG here CM

Âmáa bà i ô moo yqê dágá sì’, nán bìi yê bância và bìi but that you say word cheek one only man his.LD1 this that.he heals CM

Moo bì Âm, bà bìi kíd i nii bìi vô tòggù, for he.LD1 too that he.LD1 hear.to the.elder his.LD1 plural ear.CM bà bìi dì kan süôóze bìi bà kú kíd bi tòg màa væ that he.LD1 is.there with soldier his.LD1 that they hear him.LD1 ear this pl âm.

too

Bà iì qo dágá: “Âm làâ ú” tée, bàn làâ. that if.he.2LD1 say-to one: You go CM if, that.he go

Bà iì qó toô: “Âm yaa ú” tée, bàn yaa. that if.he.2LD1 say-to another: You come CM if, that.he come

Bà iì qo nán ba’ad bìi: “Âm kó hên yê nó” tée that if.he.2LD1 say-to man work his.LD1: You do thing this CM, then bàn ká ú.

that.he do CM

‘...they say to Jesus: “Sir, our elder, has sent us to you, to say to you that he1 isn’t worthy for you to enter his1 house. That’s why he1 hasn’t come here himself1. But even if you simply say a single word, his1 worker will be healed. For he1 too says that he1 takes orders from his1 superiors; that he1 has his1 soldiers that take orders from him1, too. That if he1 says to one: “Go!” then he will go. That if he1 says to another: “Come!”, he will come. That if he1 says to his1 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 Yag 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 ∅) 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 ‘àù [SUBORD] subjects are used instead of mì [PRON] subjects in certain clauses, so ‘ìì [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
Table 5: Subject pronouns of all four pronoun types, from Bohnhoff (1986, 107,113,120).

<table>
<thead>
<tr>
<th></th>
<th>PRON</th>
<th>SUBORD</th>
<th>LD</th>
<th>2LD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.sg</td>
<td>-n/mí</td>
<td>'àn</td>
<td>bi</td>
<td>'àn</td>
</tr>
<tr>
<td>1incl.dual</td>
<td>ba</td>
<td>ba</td>
<td>bi</td>
<td>'aa</td>
</tr>
<tr>
<td>2.sg</td>
<td>-m/mó</td>
<td>'àrn</td>
<td>bi</td>
<td>'íi</td>
</tr>
<tr>
<td>3.sg</td>
<td>∅</td>
<td>'à</td>
<td>bi</td>
<td>'íi</td>
</tr>
<tr>
<td>1excl.pl</td>
<td>ví</td>
<td>'òo</td>
<td>bi</td>
<td>'òo</td>
</tr>
<tr>
<td>1incl.pl</td>
<td>ba...ví</td>
<td>ba...ví</td>
<td>bi</td>
<td>'aa...ví</td>
</tr>
<tr>
<td>2.pl</td>
<td>ví</td>
<td>'í</td>
<td>bi</td>
<td>'íi</td>
</tr>
<tr>
<td>3.pl</td>
<td>vu</td>
<td>'ùu</td>
<td>bi</td>
<td>'íi</td>
</tr>
</tbody>
</table>

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 ’íi [2LD] may also be heard.” Example (27) shows 2LD as the subject of a relative clause within the LD domain:

(27) ... ví qod í kíi àgà: “Àkàw ∅ ∅ [ lig [ bà íi you say.to the.one house self: “Teacher, (he) say house that he.2LD i lá hen láíí páska kan waa duuí bii vu wuú máá] bà eat thing eating Easter with child following his.LDi plural there when, that.it di té lá?]” 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’e à bá ’íi lá fíí ya babbí tée] bá (shei) says time that she.2LDi goes returns comes field.from when, that bií dúo gbókii she.LDi will cook pigeon ‘Shei said that when she.2LDi returned from the field, she.LDi 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:
Such an approach will not produce the right result for 2LD, however. We cannot ensure that the antecedent of 2LD appears at least two clauses removed by requiring 2LD to be free in the LD domain but bound in a larger domain, since 2LD can appear whether or not there is a potential binder in the LD domain, as shown in (3) and (4).

## 6 The LD requirement

LFG binding requirements are generally stated in terms of binding equations as shown in (31), where \( \uparrow \) is the f-structure of the pronoun, and \( \uparrow_\sigma \) is the semantic structure corresponding to \( \uparrow \):

\[
\text{\uparrow_\sigma \text{ANTECEDENT}} = (\text{GF}^* \text{GF_{pro}} \uparrow) \text{GF_{ant}} \text{\sigma}
\]

This constraint requires the pronoun \( \uparrow \) to appear at the end of the binding path \( \text{GF}^* \text{GF_{pro}} \). The antecedent of the pronoun bears the grammatical function \( \text{GF}_{\text{ant}} \).

The binding equations can be further specialised to encode particular binding requirements. For example, the binding equation in (33) uses the off-path constraint \( \neg(\text{\rightarrow TENSE}) \) to require the pronoun to find its antecedent in the minimal finite domain containing the pronoun; the off-path constraint prevents the path through the binding domain from passing through an f-structure with the attribute TENSE:

\[
\text{\uparrow_\sigma \text{ANTECEDENT}} = (\text{GF}^* \text{GF_{pro}} \uparrow) \text{GF_{ant}} \text{\sigma} \\
\neg(\text{\rightarrow TENSE})
\]
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):

\[
\begin{array}{c}
\text{SUBJ} \\
\text{LD-ANT} + \\
\text{COMP} \quad [\ldots \text{LD pronoun} \ldots]
\end{array}
\]

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:

\[
(\uparrow^* \text{ANTECEDENT}) = ((\uparrow\sigma \text{ANTecedent})\uparrow) \text{SUBJ}
\]

1. the clause containing the LD antecedent is the smallest clause marked with \text{LD-ANT} that also contains the pronoun (the path through the binding domain to the LD pronoun may not pass through a clause with \text{LD-ANT} marking);

2. the clause containing the LD pronoun cannot have \text{LD-ANT} marking (LD pronouns are not bound by a clausemate);

3. the \text{ANTECEDENT} of the pronoun is the \text{SUBJ} of a clause with \text{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 \text{LD-ANT}-marked clause, and the LD pronoun may not appear in an \text{LD-ANT}-marked clause): the antecedent must be the subject of the \text{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 ⟨SUBORD, +⟩.

(37)  

\[ \text{SUBJ} \ [\text{SUBORD PRONOUN}] \]

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)  

\[ \text{(SUBJ ↑) SUBORD} \]

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)  

\[ \text{V} \text{ SUBORD} \text{ PRON} \text{ want \SUBORD j go \LD to \town} \]

‘They want others to go to town.’ \(\) (Bohnhoff, 1986, 114)

(40)  

\[ \text{V} \text{ SUBORD} \text{ PRON \bi \SUBORD j go \LD to \town} \]

‘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) Bàbụma m vi [ moo ’ênh pêñ vuñ tid waa bii gbọ Rabbit_i asks for what first they.PRON_j hold child his.LD_i leave mammé màalà?] [ ’i yê màa, bà vuñ sóó ’ú] [ bà water.in Q? this here focus that they.PRON_j fake CM, that ’iin sóò waa bii pi bì dóg ’yê’ yê 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 fake 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) $\begin{array}{c}
\text{SUBJ} \quad \text{LD antecedent} \\
\text{LD-ANT} + \\
\text{COMP} \quad \text{ld} \\
\text{LD-DOMAIN + } \\
\text{COMP subord} \\
\text{SUBJ [2LD]}
\end{array}$

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:

$$
(\uparrow_\sigma \text{ANTECEDENT}) = (( \text{GF}^* \quad \text{SUBJ} \quad \uparrow ) \quad \text{SUBJ})_\sigma \\
\neg(\rightarrow \text{LD-ANT}) \\
\neg(\leftarrow \text{LD-DOMAIN}) \\
(\leftarrow \text{LD-ANT}) \\
1 \\
\neg(\leftarrow \text{LD-ANT}) \\
2 \\
\neg(\leftarrow \text{LD-ANT}) \\
3
$$

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 subj-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):

\[
(\uparrow_{\sigma} \text{antecedent}) =
\begin{pmatrix}
(\text{subj}) & (\text{subj}) \\
(\rightarrow \text{LD-ant}) & (\rightarrow \text{LD-ant}) & (\leftarrow \text{LD-ant}) & (\leftarrow \text{LD-ant}) & (\rightarrow \text{subord})
\end{pmatrix}_{\sigma}
\]

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)

\[
f_1^{\text{subj}} [\text{LD antecedent}] \\
\text{ld-ant} + \\
... f_2^{\text{gf}} f_3^{\text{gf}} f_4^{\text{subj [2LD]}} [\text{subord} +]
\]

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 \( f_1 \) and the f-structure labelled \( f_2 \) will be the same, with only one clause (the exclusion zone \( f_3 \)) 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**

Yağ 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


Langdon, Margaret and Munro, Pamela. 1979. Subject and (Switch-)Reference in Yuman. *Folia Linguistica* 32, 321–344.

