

# Pronouns and Agreement: Systems Interaction in the Coding of Reference

Zygmunt Frajzyngier

## 1 The problems\*

The assumption that a category pronoun, with cross-linguistically homogeneous features exists does not explain the very significant differences in the properties of pronouns across languages. The first problem is that the pronominal subjects of embedded clauses have different binding properties. Consider complements of verbs of saying. In English (1), the pronoun *she* is coreferential with the subject of the matrix clause.

(1) because ^Sall 'said shèd 'like a "!!G\uinness you s/ee# (LL corpus)<sup>1</sup>

\* I would like to thank Aaron Broadwell for comments and constructive criticism of the previous version of this paper. Work on this paper was supported by an NSF grant *Grammars of Gidar, Mina, Lele, and East Dangla*. Data on Lele were gathered as part of an NEH supported work on Complex Sentence in Chadic.

<sup>1</sup> Whenever possible the data for analyses are drawn from corpora of natural discourse. In case of less familiar languages, elicited sentences are sometimes used in order to facilitate the understanding of the phenomena involved.

In certain cases, corpus data may contradict the *elicited* judgments of speakers. In such cases the corpus data are taken to represent the usage. The study of why the elicited judgment of speakers differs from their usage is beyond the scope of the present paper. What the speakers say they do with respect to language is quite often different from what they actually do. Such contradictions are of course not limited to language alone.

The quoted material is given in the transcription used in the source. The material quoted later from Nitsch (1961) is given in slightly modified transcription.

In Polish the situation is quite different. The pronoun *wuni* '3MPL' in the embedded clause of (2) encodes switch reference with respect to the subject of the matrix clause.

- (2) A wuni muvjili, že wuni maju  
 A.CONJ 3PL say:PAST:3PL COMP 3PL have:PRES:3PL  
 dosydz żyvności kedy xlebam stšylaju  
 enough food when bread:INSTR shoot:PRES:3PL  
 'And they<sub>1</sub> said that they<sub>2</sub> have enough food if they<sub>2</sub> use bread for  
 shooting.' (Nitsch 1960:245)

In Mupun in example (3) the pronominal subjects of the embedded clauses can only be coreferential:

- (3) wu/wa/mo sat nə ta d'i/d'e/d'u d'ee n-jos  
 he/she/they say COMP stop he/she/they stay PREP-Jos  
 'He<sub>1</sub>/She<sub>1</sub>/They<sub>1</sub> said that he<sub>1</sub>/she<sub>1</sub>/they<sub>1</sub> stopped over in Jos.'  
 (Mupun, Frajzyngier 1993:108)

In Mupun in example (4) the pronominal subjects of the embedded clause can only be switch reference:

- (4) wu/wa/mo sat nə wu/wa/mo ta d'ee n-jos  
 he/she/they say COMP he/she/they stop stay PREP-Jos  
 'He<sub>1</sub>/She<sub>1</sub>/They<sub>1</sub> said that he<sub>2</sub>/she<sub>2</sub>/they<sub>2</sub> stopped over in Jos.'  
 (Frajzyngier 1993:108)

The morpheme *she* in the embedded clause of (1) is considered a pronoun. Sentences (2) through (4) have exactly the same structure, hence there is no syntactic reason to consider the forms occurring there to be anything but pronouns as well. All of these pronouns, and also those that occupy the position of the subject of the matrix clause, share the property of coding person, gender, and number, but they vary in their binding properties. Thus instead of postulating a homogeneous grammatical category 'pronoun', it is necessary to differentiate among several categories of pronouns.

The traditional approach to pronouns cannot explain the differences in the interpretation of sentences (5) and (6):

- (5) yàá-dú nā bùy-dú kójò kò-tó  
 say-3F COMP break-3F hoe GEN-3F  
 'She<sub>1</sub> said that she<sub>2</sub> broke her<sub>1/2</sub> hoe.'
- (6) yàá-dú nā dú bùy kójò kò-tó  
 say-3F COMP 3F break hoe GEN-3F  
 'She<sub>1</sub> said that she<sub>1</sub> broke her<sub>1/2</sub> hoe.'  
 (Lele, East Chadic, Frajzyngier in progress)

The second problem is that the traditional approach to pronouns cannot explain why in the following sentence the second person masculine pronoun is used despite the fact that the language has a choice between second person masculine and second person feminine:

- (7) gaskiya, get kadan ka kə aak ʃe  
 truly (H.) past if (H.) 2M with pregnancy CONS  
 ba də mo pə ʃal d'ik n-ka  
 NEG PAST 3PL PREP marry PREP-2m  
 'Truly, in the past if you were pregnant they wouldn't marry you.'  
 (Mupun, Central Chadic, Frajzyngier 1993:88)

The third problem is whether there is a connection between the binding properties of pronouns and (a) the type of constructions in which they occur and (b) other reference coding means available in the language.

## 2 The goal and scope

The main goal of the present paper is to reanalyze the category pronoun as commonly understood in most syntactic theories. As a result of this reanalysis several types of pronouns are proposed for a taxonomy of pronouns. The paper presents two main hypotheses and several auxiliary ones: (1) The category pronoun in cross-linguistic typology consists of different types of morphemes, varying in their binding properties and in their functions. (2) The category pronoun is just one of several means of

encoding reference in language. The various coding means are complementary in that the functions encoded by one system are different from the functions encoded by another system.

The analysis is limited to four types of languages: languages without a system of coding arguments on the verb (agreement) and with only one set of pronouns, such as English; languages without an agreement system but with several sets of pronouns, such as Mupun; languages with one set of pronouns but with a rich agreement system (Polish); and language with one set of pronouns, with a split agreement system, and in which the agreement system and some pronouns are mutually exclusive within the same clause, such as Lele. The core of the argumentation is an analysis of the means of coding switch reference and coreference for third person arguments in complements of verbs of saying and in discourse. These two syntactic environments have been chosen because potentially they are the most opaque with respect to the reference of pronominal participants. In order to draw generalizations, it will be necessary to examine also other grammatical constructions.

The system of coding coreference and switch reference interacts in an interesting way with other functions performed by reference system. In the present paper I discuss the interaction with the coding of the *de dicto* domain.

The auxiliary hypotheses are as follows: (1) With respect to complements of verbs of saying, though not necessarily other constructions, which are potentially less opaque. I propose that if a language encodes the distinction between coreference and switch reference through a system of pronouns,<sup>2</sup> the following principle holds in complements of verbs of saying: If a language has a pronominal form encoding antecedent, the same form in the same syntactic position is deployed to code switch reference. A different form examples (2-4) or a different syntactic position examples (5-6) are deployed to code coreference. The importance of this hypothesis is that it provides a partial account for two elements of the taxonomy of pronouns: pronouns that are bound within a sentence and pronouns that are not bound within a sentence. It also provides an explanation for difference between sentences (5) and (6). This hypothesis contradicts the central intuition of Fiengo & May (1994) according to which 'if an expression is repeated, it follows that the two occurrences have the same semantic value, in the cases at hand, its reference, (Fiengo & May 1995:794).

<sup>2</sup> Languages may encode the distinction between cross-reference and switch reference through means other than pronouns. In Hua (Papuan) the means of encoding switch reference is through affixes to the verb (cf. Haiman 1983), cf. also Broadwell (this volume). The present paper is not concerned with these other mechanisms that languages may deploy.

There is a hierarchy of structures where the means of coding switch reference and coreference are deployed: first in clausal complements of verbs of saying, then in complements of cognitive verbs, and finally in conjoined sentences. Consistent with the proposed hypothesis, simple sentences in discourse may also be marked for switch reference.

The second hypothesis has to do with other coding means available in language: (2) If the verb encodes the subject of the clause (agreement), there exists a functional differentiation between the pronouns and the argument coding on the verb. Determination of which coding means carries which functions must be made for individual languages. In some languages pronouns carry pragmatic functions, one of which is switch reference (cf. (2)).<sup>3</sup> The coding of the subject (and other arguments) on the verb indicates referentiality, definiteness, and possibly other functions in the general domain of pragmatics. But it is theoretically possible that the distribution of functions between the coding on the verb and pronouns is different, and that pronouns encode coreferentiality within a sentence whereas coding on the verb encodes switch reference as in (5-6). Hence in addition to the taxonomy of pronouns we must have a taxonomy of head coding, in particular, a taxonomy of argument coding on the verb. Therefore, both taxonomies are proposed in the present paper.

The following are case studies organized according to the various functions the pronouns have with respect to cross-reference and disjoint reference coding.

### 3 Nondifferentiated pronouns

English has only one set of pronouns. Subject coreference is indicated in many constructions through the omission of subject pronouns, e.g.:

- (8) 4\_7\_0 <765 b> and his ^wife 'ran a" wVay# -  
4\_7\_0 <766 b> and ^had a 'child by an:other bl/oke# -(LLC)

The agreement system is limited to only one tense and one person. Sentential complements of verbs of saying must always have a subject; hence subject omission is not a viable means of encoding coreference. Omitting the embedded subject from the following sentence would result in an ungrammatical construction:

<sup>3</sup> The fact that pronouns in such languages carry a pragmatic function was already noted by Meillet (1937). Cf. also Dik (1989).

- (9) (and)) I said I haven't used a ^sewing-machine for/ !yyears# (LLC)

There have been some claims that pronominal subjects of an embedded clause are potentially ambiguous and could refer to the subject of the matrix clause, the addressee of the matrix clause, or a non-participant in the conversation, and the distinction between various types of binding is coded through prosodic means. Clements (1975:147) quoting work of Cantrill (1975) and Cooper (1976) points out that stress is not a reliable indicator of coreference, but that duration may be.

My own perusal of the London-Lund corpus of spoken English shows that in all cases where the third person pronominal subject of the embedded clause shares the features gender and number with the subject of the matrix clause, coreference is the norm. Here are some of the examples drawn from that corpus:

- (10) the ^president said she had :g\athered# ^from my -- ^some of my re:marks over :lunch that [@m] I ^wouldn't want to !live -- in the c/ollege#

In most cases the pronominal subject of the embedded clause is coreferential with the subject of the matrix clause if they agree in gender and number, e.g.:

- (11) 1\_4\_0 <1094 B> but "h\alf of them had t/ead it#  
1\_4\_0 <1095 B> and the ^others 'said they 'wanted :me to !t\alk a'bout it#
- (12) 1\_4\_0 <1112 B> \*((yeah)) at the "last\* 'faculty of \arts meeting#  
1\_4\_0 <1113 B> they ^said they'd \ask the \_provost#
- (13) 1\_5\_0 <919 A> ^and my p\arents were 'so fed 'up with me#  
1\_5\_0 <920 A> that they said ^hey w\ere 'fed up of supp\orting you#
- (14) 1\_6\_0 <725 B> I ^didn't s\ee him ((en/ough#))  
1\_6\_0 <726 B> but ^Gloria said she saw . was ^talking to 'him#
- (15) 1\_6\_0 <1136 B> ^said to 'me /once# -  
1\_6\_0 <1137 B> that he "f/elt#  
1\_6\_0 <1138 B> that ^[?]if - 'Doc :M\iddleton# .  
1\_6\_0 <1139 B> ((was)) ^coming ":d\own the 'stairs#

- (16) 1\_8\_0 <373 A> ^well ^hès just left {the ^skirting b\oard#}  
 {when he ^f\inished#}#  
 1\_8\_0 <374 A> he ^said he !\ouldn't 'do an'other 'five h\ours -  
 {which is ^all it n/eds#}# -
- (17) 1\_9\_0 <952 A> (laughs - ) ^they 'said they !werent {s\ure} of
- (18) 6\_8\_0 <117 B> +yes+ well ((the girl)) said that she was very  
 worried about his \*((several sylls))\* the :f\ind 'spot# .
- (19) well ^Nightingale :said that he . he :might want to/  
 :get a:way from . Lower N\etherhall you s/ee
- (20) 5\_8\_0 <570 b> and [@m] - . she ^just said \_well she just .  
 "!\idn't 'think she could 'make it on 'Friday n/ight#
- (21) 6\_8\_0 <107 B> they said that this they do [?]anyway they have a  
 lot of time together and they always have all their/  
 meals together - and they talk about all sorts of things -
- (22) 8\_4\_9 <586 A> and ^they !said \_theyd 'written it :d\own \*but#\*
- (23) 9\_2\_2 <173 A> ^does he {kn\ow} :Doctor S\owerbarn# -  
 9\_2\_2 <174 B> ^[j@ ? w] ^y\es#  
 9\_2\_2 <175 B> ^he !said he :d\id#

The third person embedded clause subject pronoun that shares the features gender and number with the matrix clause subject encodes switch reference only if it is in direct speech, which may be signaled by a number of prosodic or syntactic devices. Here is an example:

- (24) 4\_7\_0 <541 a> so the ^man said !!w\ell# .  
 4\_7\_0 <542 a> he was ^digging 'up !my r\oses#

A perusal of the Brown corpus of written English indicates virtually the same property. The third person pronoun of the embedded clause is coreferential with the subject of the matrix clause if they agree in gender and number, e.g.:

- (25) The jury said it did find that many of Georgia's registration . . .
- (26) The jurors said they realize "a proportionate distribution  
 of these funds might disable this program in our less populous counties".
- (27) Ratcliff said he expects to tell home folks in Dallas why he thinks  
 Berry's proposed constitutional amendment should be rejected.

One of the systematic means to encode switch reference in the LL corpus is to use a full noun phrase, as in the following:

(28) Alec had to ring me about examina<sub>tion</sub>\_business#/ this <sub>morning</sub># and <sub>he</sub> said <sub>he</sub> <sub>happened</sub> to <sub>mention</sub> that Oscar was a way for a couple of <sub>days</sub># - . <sub>but</sub> will be <sub>back</sub> [<sub>@:m</sub>] on <sub>Monday</sub> ((<sub>or</sub> there<sub>abouts</sub>#{#})) .

(29) I mean\* <sub>this</sub> . [<sub>@</sub>] <sub>Mallet</sub> said <sub>Mallet</sub> was [<sub>@</sub>] <sub>said</sub> something <sub>about</sub> [<sub>@</sub>] you know <sub>he</sub> <sub>felt</sub> it <sub>would</sub> be a good thing if [<sub>@</sub>] . if Oscar: <sub>went</sub>#

(30) 6\_7\_0 <678 a> remember Brockhouse had said that Frank Jones made a great mistake in not retiring from politics at the end of the First World War

In the whole LL corpus there are very few examples where the 'agreeing' pronominal subject of the embedded clause must be construed as disjoint reference. One is where the pronom of the embedded clause encodes the topic of the paragraph and the matrix clause preceding it, is a parenthetical remark:

(31) 10\_1\_0 <751 ra> and <sub>he</sub> <sub>never</sub> looked <sub>happy</sub>#

10\_1\_0 <752 ra> <sub>he</sub> <sub>never</sub> looked <sub>con:ent</sub># .

10\_1\_0 <753 ra> <sub>and</sub> as <sub>Norman</sub> : <sub>Yardley</sub> said <sub>he</sub> : <sub>spent</sub> his : <sub>time</sub>

playing <sub>back</sub># -

10\_1\_0 <754 ra> <sub>and</sub> <sub>paid</sub> the <sub>penalty</sub> in the <sub>end</sub># - - -

Based on this sample only, one could say that if the pronominal subject of the embedded clause shares the features of gender and number with the subject of the matrix clause, the third person pronom of the embedded clause encodes coreference with the subject of the matrix clause. But written English, especially artistic prose, contains data in which a subject pronom must be construed as encoding a switch reference. Compare the following fragment, which occurs at the beginning of a paragraph. In the second sentence of the fragment, the two pronom 'he' have different referents:

- (32) As an archer he was without equal. My grandfather said he could hit a moving buck at 150 yards, adding that he would not have liked to expose any part of himself in battle to a Bushman archer under 150 yard's range. But he not only hunted with bow and arrow.  
(van der Post 1986:10)

The conclusion for English is that there is no differentiation in the pronominal system or in the coding on the verb for encoding the distinction between switch reference and coreference. The absence of differentiation in the pronominal system and the absence of other means of coding reference explains the very frequent usage of direct speech to assure proper interpretation in conversational English, as represented by the LL corpus.

#### 4 Logophoric pronouns: the sentence as a binding domain

The following terms are necessary for understanding the argumentation: A pronominal antecedent is a pronoun that occurs in the matrix clause in a complex sentence or in a preceding clause in discourse. A logophoric pronoun is a pronoun that is bound by an argument of the matrix verb of saying (this formulation is more restrictive than the one found in Hagège 1974). This argument may be the subject or the addressee of the matrix clause; hence a clause may contain a logophoric subject or a logophoric addressee (Frajzyngier 1985, 1989).

In many languages from different families the logophoric pronouns have a different form from that of the matrix pronouns. The pronouns identical with the matrix pronouns are used to encode switch reference. Compare the following data from Ubangi languages spoken in the Central African Republic (data from Cloarec-Heiss 1969:61ff) with terminology as presently understood):

Table 1

Language	Matrix pronoun	Logophoric	Switch reference
Banda	ɔ̃	ɔ̃nɛ	cɛ
Ngbaka	xɛ̃	xɪ̃	xɛ̃
Gbandili	xá	xɪ̃	xā

In Ewe (Kwa branch, Niger-Congo) the antecedent third person pronoun is *e*. The logophoric pronoun is *yè* (Clements 1975:142).

In Igbo (Niger-Congo) the antecedent third person is *ó* and the logophoric is *yá*. In Gokana (Cross-River, Benue-Congo, Niger-Congo) the antecedent is *áè* preceding the verb, and the logophoric marker is *E* suffixed to the verb (Hyman & Comrie 1981:69).

Mupun (West Chadic), an SVO language, does not have a system of argument coding on the verb. In Mupun there exist two series of logophoric pronouns, one bound by the subject of the matrix clause and the other by the addressee of the matrix clause. Both types of logophoric pronouns have different forms from that of antecedents. If the complement clause pronouns are identical with the antecedent, they are not bound by the antecedent. The following table gives the form of logophoric and switch reference pronouns in Mupun:

Table 2

Person	Antecedent subject	Log. subject	Antecedent addressee	Log. addressee	Switch reference
3M	wù	ɗí	wùr	gwàr	wù
3F	wà	ɗè	wàr	páa	wà
3PL	mō	ɗū	mò	nūwā	mō

Examples (tones unmarked):

- (33) wu/wa/mo sat nə ta ɗi/ɗe/ɗu ɗee n-jos  
 he/she/they say COMP stop he/she/they stay PREP-Jos  
 'He<sub>1</sub>/She<sub>1</sub>/They<sub>1</sub> said that he<sub>1</sub>/she<sub>1</sub>/they<sub>1</sub> stopped over in Jos.'
- (34) wu/wa/mo sat nə wu/wa/mo ta ɗee n-jos  
 he/she/they say COMP he/she/they stop stay PREP-Jos  
 'He<sub>1</sub>/She<sub>1</sub>/They<sub>1</sub> said that he<sub>2</sub>/she<sub>2</sub>/they<sub>2</sub> stopped over in Jos.'  
 (Frajzyngier 1993:108)

Examples of coding logophoric and switch reference addressees:

- (35) n-sat n-wur nə gwar ji  
 1SG-say PREP-3SG COMP 3SG come  
 'I told him<sub>1</sub> that he<sub>1</sub> should come.'
- (36) n-sat n-wur nə wur ji  
 1SG-say PREP-3SG COMP 3SG come  
 'I told him<sub>1</sub> that he<sub>2</sub> should come.' (Frajzyngier 1993:112)

The two sets of pronouns in Mupun are deployed mainly in complements of verbs of saying. In other environments, including complements of other verbs, only the pronouns from the antecedent set are used and consequently the grammatical system allows for ambiguous interpretation. The ambiguity must be resolved by the context, as in the following sentences:

- (37) amma kat ba me mat ta tok ká mo kas  
 but (H.) when NEG QUANT woman fall greet CONJ 3PL NEG  
 d'ag mo seet dak ʃe mo n-dəm ter a yil  
 then 3PL depart just CONS 3PL FUT-go spend the night PREP bush  
 'But if there is no woman who comes across and greets them they will just  
 go and spend the night in the bush.' (Frajzyngier 1993:509)
- (38) kat wur naa ngo d'ə ba wur ʃal kas  
 when 3M see man REL NEG 3M strong NEG  
 ʃe wur dəm pe mənə kə wur put d'i  
 CONS 3M go place ANAPH COMP 3M leave COMP  
 'When it [the leopard] sees a man who is not strong then it goes in that  
 direction so that it can get out.' (Frajzyngier 1993:511)

The facts presented by languages with logophoric pronouns indicate that two types of pronouns exist, one bound within a sentence and the other not and that the antecedent is bound within the discourse.

##### 5 Switch reference pronouns and the function of agreement

Polish has subject pronouns differentiated for three persons and three genders in the singular and three persons and two genders in the plural. The past tense of the verb

differentiates among three persons and three genders in the singular and three persons and two genders in the plural. The present and the future forms encode person and number but not gender. Because there is such a rich system of coding on the verb, subject pronouns can be deployed for other functions. I propose that one of these functions is switch reference. Here is an illustration and the evidence for the hypothesis.

The following sentence has a subject pronoun before the embedded clause verb in the future tense. If the pronoun were omitted, the subject of the embedded clause would be identical with the subject of the matrix clause:

- (39) Očec móvi, ze ně vé,  
 father say:3SG:PRES COMP NEG know:3SG:PRES  
 cy uona přystaě.  
 whether 3F agree:3SG:FUT  
 'The father<sub>1</sub> says that he<sub>1</sub> does not know whether she will agree.'  
 (Nitsch 1960:138)

In the following sentence, the third person subject in the apodosis clause is marked by the third person plural pronoun *wuni*. The context clearly indicates that this pronoun encodes a subject different from the third person plural subject of the protasis clause:

- (40) To bylo tutaj jak Šfydy vojovali,  
 it was here when Swede:PL fight:3PL:PAST  
 to wuni tam šaěec wusypali  
 DEM 3PL there rampart pour:PAST:3PL:M  
 'It was here, when Swedes (1) were fighting, they (2) built a  
 rampart over there.'

In the following sentence, all clauses except the last one have the same subject. The nominal subject is present in the first clause but is coded on the verb in the remaining clauses, except for the relative clause, where the subject is different and is marked by the noun *Šfydy* 'Swedes':

- (41)    *ji*    *p'jekaže* *po-vyskrobyvali*                   *s*    *korytuf*  
           CONJ bakers   PL-scrape:FREQ:PAST:3PL:M   from   trough:PL:GEN  
           *ji*    *wup'jekli*           *xlip*   *ji*    *vystšeli*   *tam*   *gdzie*  
           CONJ bake:PAST:3PL:M bread   CONJ shoot:PAST:3PL:M   there   where  
           *te*   *Šfydy*   *mieli*                    *ten*   *wobus*  
           DEMSwedes have:PAST:3PL:M DEM camp.<sup>4</sup>  
           'bakers scraped the troughs clean, baked the bread [from the scraps], and  
           shot [the bread] in the direction where the Swedes had their camp.'

The above sentence is followed in the text by the following sentences whose third person subject is the pronoun *wuni*. The use of this pronoun is motivated by the fact that there is again a change of subject from the preceding sentence. The embedded clause also has the third person plural *wuni* because again there is a change of subject. In the third clause there is no pronoun because the subject is coreferential with the subject of the preceding clause:

- (42)    A           *wuni*   *muvjili,*           *že*                    *wuni*   *maju*  
           A.CONJ 3PL   say:PAST:3PL   COMP 3PL   have:PRES:3PL  
           *dosydz* *žywności*   *kedy*   *xlebam*            *stšylaju*  
           enough food    when   bread:INSTR   shoot:PRES:3PL  
           'And they<sub>1</sub> said that they<sub>2</sub> have enough food if they<sub>2</sub> use bread for  
           shooting.' (Nitsch 1960:245)

In the following fragment, the first sentence has third person plural subject. In the second sentence there is a switch of subject to third person masculine singular, and accordingly this is done with the third person pronoun:

<sup>4</sup> Demonstratives (DEM) in this and other non-literary texts mark previous mention in discourse, the function similar to that of the English definite marker.

- (43) Se drudzy napili ji zaro posnyli.  
 REFL other:PL drink:PAST:PL:M CONJ soon fall asleep:PAST:PL:M  
 że nic wo ničem nie vědzieli.  
 COMP nothing about nothing NEG know:PAST:PL:M  
 'Those others drank, and soon fell asleep, not knowing anything around them.'
- (44) A wun cisnuw te košyčysko, . . .  
 A.CONJ 3M throw:PAST:3:M DEM basket  
 'And he threw away that basket . . .' (Nitsch 1960:242)

Consider now a fragment of text consisting of two sentences. The first sentence has three clauses, each with different subject. The subject of the first clause is coded only on the verb, and it is coreferential with a preceding subject. The second clause has the subject *batog* 'whip'. The third clause has the subject *xtož* 'somebody'. The second sentence has the subject identical with the subject of the first clause. But since there was another subject in between, the subject of the third clause is marked by the pronoun *on*: '3M,SG':

- (45) I jak wyšoł s košćóła już  
 CONJ when leave:PAST:3MSG from church already  
 batoga ni bylo, xtož ukrat.  
 whip:GEN NEG be:3N somebody steal:PAST:3MSG  
 'When he left the church, the whip was not there. Somebody stole it.'
- I on zaklon:  
 CONJ 3M curse:PAST:3MSG  
 'And he cursed.' (Nitsch 1960:145)

Omission of *on* in the above sentence would result in an interpretation whereby the person who stole the whip was also the person who cursed. The interpretation resulting from the omission of the pronoun is the evidence for the switch reference function of the pronoun.

A perusal of spoken texts in Nitsch (1960) shows that every usage of third person pronouns *on* '3MSG', *ona* '3FSG', *one* '3FPL' and *oni* '3MPL' before a verb is a case of subject switch reference. The same switch reference function of pronouns can be observed in contemporary literary Polish:

- (46) tak zrobiłby każdy z nas, ale to jest  
 so do:3M:HYP each PREP IPL but this is  
 za proste. On wolał stanąć do raportu.  
 too simple. 3SG:M prefer:PAST:3SG:M stand:INF to report  
 'Each of us would have behaved like that. But that is too simple. He  
 preferred to report for reprimand' (Sources)<sup>5</sup>
- (47) kiedy ja stąd wyjadę, on będzie miał ...  
 when 1SG from:here leave:FUT:1SG 3SGM FUT:3SG have:3SG:M  
 'When I leave, he will have ...' (Sources)
- (48) Wy mówcie, a on będzie słuchał  
 2PL speak CONJ 3MSG FUT:3SG listen:3SG:M  
 'you speak, and he will listen' (Szpotkański, Poemat o Szmaciaku)

The third person pronoun does not necessarily encode switch reference with respect to the subject of the preceding clause. It may encode switch reference with respect to the potential antecedent. Thus the pronoun may be used to disambiguate between the third person masculine and the third person feminine antecedent, as in the following example:

<sup>5</sup> Literary Polish examples are taken from samples used in preparation of Kurcz et al. (1990), *Słownik frekwencyjny polszczyzny współczesnej*. The electronic media through which I had access to the material, courtesy of Zygmunt Saloni, do not allow identification of the source of every sample. These examples are later identified as Sources. I had access to Szpotkański's poem also through the electronic media; hence no publisher or page references are indicated. Because of understandable space limitations the number of examples in this paper is kept to a bare minimum.

- (49) odkażd jednak kierownik skupu Bolek Domagała  
 since however manager acquisition:GEN B. D.  
 został oficjalnym narzeczonym Stasi Rosłoń poczuł  
 become:3M:PAST formal fiancé S. R.-GEN feel:PAST:3MSG  
 się tak pewnie, że jawnie przechwalał się  
 REFL so sure COMP overtly boast:PAST3MSG REFL  
 przed ludźmi, jak on to nabiera  
 in front people:INSTR how 3M DEM cheat:PRES:3SG  
 głupich urzędników.  
 stupid:PL:ACC clerk:PL:ACC  
 'When the manager of acquisitions B.D. became the formal fiancé of  
 S.R., he openly boasted of how he hoodwinks stupid clerks.' (Sources)

We can now compare the English and the Polish coding systems with respect to reference in the same syntactic constructions. In Polish, argument coding on the verb indicates coreference with the preceding subject. The subject pronoun preceding the verb encodes switch reference. In spoken English the subject pronoun preceding the verb encodes coreference. Switch reference in spoken English is coded through the use of a full noun or through the deployment of direct speech. The use of pronouns in English corresponds thus to the coding on the verb in Polish rather than to the use of pronouns in Polish.

## 6 Split-agreement system and syntactic coding of logophoricity

Lele is an East Chadic, SVO language spoken in the Kelo district, Chad Republic.<sup>6</sup> The language has an interesting split-pronominal system whereby the first and second person pronouns in matrix or independent clauses occur before the verb but the third person masculine, feminine, and plural pronouns occur after the verb. The following table illustrates pronominal marking on the verb *é* 'go':

<sup>6</sup> All data on Lele are from my own fieldnotes and work in progress. The published work on Lele (Weibegué 1992 and Weibegué & Palayer 1982) does not address the problems of the binding of pronouns.

Table 3

	Singular	Example	Plural	Example
1	ŋ-	ŋè	nī INCL	nī-è
2M	gi-	gi è	ngú	ngú-è
2F	mē-	mé è		
3M	-dī	è-dī	-gé	è-gé
3F	-dú	è-dú	Indefinite human gé	gé è

If the verb begins with a consonant, the first and second person pronouns are not prefixed to the verb but occur as separate words. The pronouns following the verb are suffixes and therefore can be considered to be a system of coding on the verb, 'agreement' in traditional terminology. The interest of Lele is first, that it provides evidence that the pronominal coding on the verb is not an agreement system in the sense of an element in the clause triggering the presence of some formal characteristics on another element, and second, it encodes logophoricity through syntactic and not morphological means.

### 6.1 Coding on the verb is not an agreement system

Natural discourse in Lele provides evidence that the coding on the verb is not an agreement system. The third person singular pronouns are not used in the same clause in which an overt nominal subject is used, e.g.:

- (50)    *gilkínín è sùk liqda*  
           Gilkinin go market yesterday  
           'Gilkinin went to the market yesterday.'

In the texts gathered there is no single third person subject suffix occurring in the same clause in which there is an overt nominal subject. Attempts to insert a pronominal subject in clauses occurring in natural discourse resulted in ungrammatical constructions. In the following sentence, taken from a narrative, there is no third person subject pronoun:

- (51) kùrmbàlò né bè gē kāyā máàni tù  
king do BEN 3PL thing so all (Fr)  
'The king made things so for them.' (i.e. as they asked)
- (52) \*kùrmbàlò né-dī bè gē kāyā máàni tù  
king do-3M BEN 3PL thing so all (Fr)  
'The king made things so for them.' (i.e. as they asked)

Even when an elicited sentence was situated within a natural discourse, it was not acceptable with a nominal subject and with a subject pronoun:

- (53) tērēn-dī né bè-gē kāyā máàni tù  
daughter-3M do BEN-3PL thing so all (Fr)  
'His daughter made things so for them.' (i.e. as they asked)
- (54) \*tērē-n-dī né-dú bè gē kāyā máàni tù  
daughter-3M do-3F BEN 3PL thing so all (Fr)  
'His daughter made things so for them.' (i.e. as they asked)

So the coding on the verb of the third person singular masculine and feminine subject in Lele provides powerful evidence for the independence of this coding means. It is not triggered in any way by the nominal subject present in the clause.<sup>7</sup> The coding of the subject on the verb is not an agreement system. This poses a natural question about the function of the coding on the verb. We have already seen in Polish that the subject marking on the verb encodes coreference with the immediately preceding subject, either within the same clause or in a preceding clause. Given this fact and similar facts in other IE languages, one would expect that marking of the subject on the verb also encodes coreference. But that is not the case for third person singular subjects. The

<sup>7</sup> In data obtained through elicitation, a clause with a nominal subject may optionally also have the third person subject suffix added to the verb, e.g.:

- (a) cāngé ãm-dī lùmba  
canige domesticate-3M horse  
'Canige broke a horse.'
- (b) gilkimín e-dī kasugú lɔdà  
Gilkimin go-3F market (local) yesterday  
'Gilkimin went to the market yesterday.'

These two sentences are examples of elicited data being different from the ones found in a corpus.

third person plural suffix on the verb may, however, cooccur with the third person nominal subject. This fact will be explained later in this paper.

### 6.2 Logophoricity in Lele

Complements of verbs of saying in Lele provide the evidence that the third person singular pronoun when coded on the verb cannot be bound within a sentence. That is most clearly evident when the subjects of the matrix and the embedded clause are different, e.g.:

- (55)    η-yàá    ná    bòy-dú    kójò    kò-tó  
           1SG-say    COMP    break-3F    hoe    GEN-3F  
           'I said that she broke her hoe.'
- (56)    η-yàá    ná    bòy-dí    kójò    kè-i  
           1SG-say    COMP    break -3M    hoe    GEN-3M  
           'I said that he broke his hoe.'

When the subject pronouns of the complement of the verb of saying follow the verb, i.e., when they have the same syntactic position as pronominal antecedents, they encode switch reference as in (5). If the pronominal subjects of the complement clause precede the verb, they encode coreference as in (6). Thus the coding in Lele supports the hypothesis that different means are deployed to code coreference and the same means are deployed to code switch reference. In addition to examples (5-6), which were elicited, compare the following examples, taken from natural discourse:<sup>8</sup>

- (57)    tōrmō-η    nā    dú    òdè    ná    bà-tó    ɖè  
           girl-DEF    COMP    3F    go    CONJ    father-3F    NEG  
           sé    kāl    è-gé    ná    kāmḍā    sùbù  
           got up    pass    go-3PL    CONJ    wives    three  
           'The girl said that she would not go with her father, she passed [from  
           her father to the man] and he went with three wives.'  
           (lit. 'they went with three wives')

<sup>8</sup> In Lele, as in many Chadic languages (cf. Frajzyngier 1996) the verb of saying is most often omitted in natural discourse. The complementizer or an auxiliary is the evidence that the predicate of the clause is a verb of saying.

- (58) ntóormò nā dú ódè ná bà-tóo ɖè  
 girl COMP 3F leave CONJ father-3F NEG  
 sē è ná bāyndí-ŋ  
 INCEPT go CONJ man-DEF  
 'The girl refused to go with her father and instead went with the man.'  
 (lit. 'the girl [said] that she would not go with her father . . .')

The third person pronoun may not occur before the verb except to code logophoricity. The position before the verb results in an ungrammatical sentence if logophoricity is ruled out:

- (59) \*ŋ̄-yàá ná dú bòy kójò kò-tó  
 1SG-say COMP 3F break hoe GEN-3F  
 'I said that she broke her hoe.'
- (60) \*ŋ̄-yàá ná dí bòy kójò kè-i  
 1SG-say COMP 3M break hoe GEN-3M  
 'I said that he broke his hoe.'

The third person masculine logophoric pronoun is *ì* rather than *dí* and like the third feminine it precedes the verb; it is cliticized to the preceding complementizer, e.g.:

- (61) dàì lē gòl kām̄dā-ì sùbù tu né gō láy. wà  
 3M also see wives-3M three all COP REL beautiful very  
 nā-ì bēè-ì tāmá mání pínà ɖè  
 COMP-3M give.FUT-3M(DAT) wife ANAPH one NEG  
 'He himself saw that all his three wives were very beautiful, and said that he would not give one of them to him.'

## 7 Reference system in the de dicto domain

The purpose of the present section is to demonstrate (1) that there exists a connection between logophoricity and hypothetical mood and (2) that languages make a systematic distinction between de re and dicto addressees (cf. Frajzyngier 1991). Taken together the two hypotheses indicate that one must also take into consideration, whether the

reference is in the domain de re or domain de dicto, as well as the syntactic and discourse domains of binding. The evidence for (1) is provided by the fact that in languages in which there exists a formal distinction between logophoric and nonlogophoric coding, the means deployed for logophoric reference are also deployed for the coding of hypothetical. The explanation for this fact is that if a reference is logophoric, it is to an entity that exists in the domain of discourse, not necessarily in the domain of reality. The hypothetical modality is never in the domain of reality. I illustrate the connection between the means of encoding logophoricity and hypothetical coding in two languages in which the means are formally different: Mupun and Lele.

### 7.1 Logophoricity and the hypothetical in Mupun

Although most of the logophoric pronouns in Mupun occur in the embedded clause when the main clause has a verb of saying, there are examples of sentences in which the main verb is not a verb of saying but that have logophoric pronouns in the embedded clause:

- (62) kat puun la reep nə mo cin ko d'un mbə sin  
 when parent girl DEF 3PL do as 3PL.L FUT give  
 la reep nə n-kenken kas  
 girl DEF PREP-very fast NEG

'If the parents of the girl act as if they will not give away the girl fast . . .'

- (63) kat la reep ɕeer am kaa la mis ɕe la mis nə man nə  
 if girl pour water on boy CONS boy DEF know COMP  
 paa pə dem d'in mənə  
 3F.L PREP like 3M.L then

'If a girl pours water on the young man, then he knows that she loves him.'  
 (Frajzyngier 1993:117)

In the two examples above, and in other similar ones, the main clause represents a hypothetical event. And the hypothetical event is treated as an event that belongs to the domain de dicto rather than de re; hence the participants are referred to through logophoric pronouns.

## 7.2 Logophoricity and the hypothetical in Lele

Coding of the third person plural subject on the verb through suffixation in Lele indicates a known, referential subject, e.g.:

- (64) han è-gé dāwù-gē ná kámdà sò  
 when go-3PL gather-3PL CONJ women two  
 'When they went it was a group with two wives.'

Third person plural is coded by the suffix *ge*. In the second verb in the sentence the subject marker is omitted, e.g.:

- (65) déená nòjú-tò kínyé sē òm'-gé tū gō bàl  
 then relative-3F DEM-PL INCEPT catch-3PL goat REL castrated  
 nè òb ná kārā  
 HABIT ask ASSOC people  
 'Then, her relatives caught a large castrated goat and challenged the people.'
- (66) dà tógú tóŋ ní kārā kòŋdirè kūsī-gē sùbù  
 LOC village certain LOC people young men body-3PL three  
 'In a village there were three young men.'
- (67) yā-gé nā  
 say-3PL COMP  
 'They said thus:'
- (68) è-gé cáani bāyndí-ŋ kàlò sè yír-ì  
 go-3PL bush man-DEF snake got up bite-3M  
 'When they went into the bush, a snake bit him.' (the man)

The presence of the two types of configuration in Lele allows the encoding of the indefinite human subject (Frajzyngier 1982), equivalent of English *they*, French *on*, German *man*, and Spanish third person plural coding on the verb. The coding of the hypothetical is achieved through the position of the third person plural pronoun *ge* before the verb, e.g.:

- (69) *dū h-aŋ bà nā ge bē-ì kùrmbàlò tāmá-ì*  
 3M DEM CONTR COMP INDEF give-3M chief wife-3M  
*bà na-ì wèl ná ùndò hà kúr ná wèl*  
 CONJ COMP-3M sleep CONJ 3F till place COMP day  
*nā gē dígr-ì ná kólò tūmádū-ì ná d'árínìŋ d'è*  
 COMP INDEF kill-3M CONJ reason death-3M CONJ anger-3M NEG  
 'The first one said that if he is given the chief's wife to sleep with till the  
 daybreak, then he can be killed, because death does not worry him.'
- (70) *gē yà bē kùrmbàlò*  
 INDEF tell BEN chief  
 'The chief was informed.'
- (71) *gē dígrì-gē sò*  
 INDEF kill-3PL two  
 'Two were killed.'
- (72) *dài gō sùbù gē dígrè-ì è wèl ná kùrmbàlò tāmá-ì*  
 3M REL three INDEF kill:FUT-3M go sleep CONJ chief wife-3M  
 'The third one, who was going to be killed, went to sleep with the chief's  
 wife.'

The logophoric reference and the hypothetical are both elements of the *de dicto* domain (Frajzyngier 1991). Compared to the domain of reality, the *de dicto* domain has often fewer semantic distinctions. Thus if a language has a gender distinction in the domain of reality it might not have a gender distinction in the domain *de dicto*. This is the case in Polish and Lele, and it explains why in example (7) in Mupun, where instead of the biologically expected second person feminine, the narrator (who was a woman) used the second person masculine.

## 8 Implications

The proposed analyses reveal that the category pronoun has different binding properties across languages. These properties are in a complementary relationship with other coding means available in the language. The proposed analyses call for a taxonomy of pronouns based on their binding domains. Along with a taxonomy of pronouns we must have a taxonomy of other means of coding reference in the language. Case studies in the present paper suggest a taxonomy that contains the

following elements: pronouns that are bound within a sentence but not within a clause (logophoric pronouns); switch reference pronouns, which are bound by an antecedent in discourse but not by the immediately preceding antecedent (spoken, nonliterary Polish); and pronouns that may be bound either within the same sentence or by an immediately preceding antecedent in discourse (English).

Another major implication of the analyses is that what is commonly called 'agreement' is a coding means with a number of functions. One of these functions is the coding of coreference with the immediately preceding subject. But there are also split 'agreement' systems, in which some pronouns encode coreference and others encode switch reference.

In order to fully understand the binding of pronouns or lack thereof, one must take into consideration the semantic modalities encoded in the grammatical structure of the language. In the present paper only two modalities have been discussed, the *de dicto* and *de re* modality. But it is conceivable that some languages may have other distinctions.

## References

- Broadwell, Aaron. (this volume). Binding Theory and Switch-Reference.
- Cantrall, William R. 1975. Favored structures and intonational limitations. *Papers from the parasession on functionalism*. Chicago: Chicago Linguistic Society. 36-46.
- Clements, George N. 1975. The logophoric pronoun in Ewe: Its role in discourse. *Journal of West African Languages* 10. 141-177.
- Cloarec-Heiss, France. 1969. *Les modalités personnelles dans quelques langues oubanguiennes. Discours direct-Discours indirect*. Paris: SELAF.
- Cooper, William E. 1976. Speech timing of coreference. *Proceedings of the 1st International conference on speech, acoustics, and signal processing*. Philadelphia: I.E.E.E.
- Dik, Simon. 1989. *The theory of functional grammar*. Dordrecht: Foris.
- Fiengo, Robert & Robert May. 1994. *Indices and identity*. Cambridge, Mass: The MIT Press.
- Fiengo, Robert & Robert May. 1995. Response to William's review of *Indices and identity*. *Language* 71. 794-801.
- Frajzyngier, Zygmunt. 1982. Indefinite agent, impersonal, and passive: A functional approach. *Lingua* 58. 267-290.
- Frajzyngier, Zygmunt. 1985. Logophoric systems in Chadic. *Journal of African Languages and Linguistics* 7. 23-37.
- Frajzyngier, Zygmunt. 1989. Three types of anaphora. In Isabelle Haik & Laurie Tuller (eds). *Current Progress in African Linguistics*. Dordrecht: Foris. 194-216.
- Frajzyngier, Zygmunt. 1991. The de dicto domain in language. In Elizabeth C. Traugott & Bernd Heine (eds). *Approaches to Grammaticalization*. Volume I. Amsterdam / Philadelphia: John Benjamins. 219-251.
- Frajzyngier, Zygmunt. 1993. *A Grammar of Mupun*. Berlin: Reimer.
- Frajzyngier, Zygmunt. 1996. *Grammaticalization of the Complex Sentence: A case study in Chadic*. Amsterdam & Philadelphia: John Benjamins. Complementary Series to Studies in Language.
- Frajzyngier, Zygmunt. in progress. A syntax of Lele.
- Hagège, Claude. 1974. Les pronoms logophoriques. *Bulletin de la Société de Linguistique de Paris* 69.1. 287-310.
- Haiman, John. 1983. On some origins of switch reference marking. In John Haiman & Pamela Munro (eds). *Switch reference and universal grammar*. Amsterdam/Philadelphia: John Benjamins. 105-128.
- Hyman, Larry M. & Bernard Comrie. 1981. Coreference and logophoricity in Gokana. *Studies in African Linguistics, Supplement* 8. 69-72.
- Kurcz, Ida, Andrzej Lewicki, Jadwiga Sambor, Krzysztof Szafran, & Jerzy Woronczak. 1990. *Słownik frekwencyjny polszczyzny współczesnej*. Cracow: Polska Akademia Nauk, Instytut Języka Polskiego.
- Meillet, Antoine. 1937. *Introduction à l'étude comparative des langues indo-européennes*. Paris: Hachette.

- Nitsch, Kazimierz. 1960. *Wybór polskich tekstów gwarowych*. 2nd ed. Warsaw: Państwowe Wydawnictwo Naukowe.
- van der Post, Laurence. 1986. *The lost world of the Kalahari*. San Diego: Harcourt Brace.
- Weibegué, Christophe. 1992. *Système phonologique du Lélé*. Mémoire en vue de l'obtention de la Licence ès Lettres, Université du Tchad.
- Weibegué, Christophe & Pierre Palayer. 1982. *Lexique Lele-Français*. Sarh: Centre d'études linguistiques.

### List of abbreviations

ANAPH	Anaphor
COMP	Complementizer (morpheme)
A.CONJ	Counterexpectation conjunction
CONS	Consecutive
CONTR	Contrastive focus marker
FUT	Future
GEN	Genitive
H.	Hausa
INDEF	Indefinite human subject
INSTR	Instrumental
LLC	London-Lund (corpus)
LOG	Logophoric
NEG	Negative
PAST	Past tense
PL	Plural
PREP	Preposition
PRES	Present