

**PRONOMINAL AFFIXES, THE BEST OF BOTH WORLDS:  
THE CASE OF YURAKARÉ<sup>1</sup>**

By RIK VAN GIJN

*Radboud University Nijmegen and Max Planck Institute for Psycholinguistics Nijmegen*

## ABSTRACT

Pronominal affixes in polysynthetic languages have an ambiguous status in the sense that they have characteristics normally associated with free pronouns as well as characteristics associated with agreement markers. This situation arises because pronominal affixes represent intermediate stages in a diachronic development from independent pronouns to agreement markers. Because this diachronic change is not abrupt, pronominal affixes can show different characteristics from language to language. By presenting an in-depth discussion of the pronominal affixes of Yurakaré, an unclassified language from Bolivia, I argue that these so-called intermediate stages as typically attested in polysynthetic languages actually represent economical systems that combine advantages of agreement markers and of free pronouns. In terms of diachronic development, such ‘intermediate’ systems, being functionally well-adapted, appear to be rather stable, and it can even be reinforced by subsequent diachronic developments.

## 1. INTRODUCTION

Many authors regard the distinction between anaphoric pronouns and agreement markers as a scalar one (cf. Siewierska 2004: 121; Corbett 2006). On this continuum, pronominal affixes in polysynthetic languages typically take a middle position: they share some characteristics with agreement markers, and others with pronouns. The differences between the two extremes on the continuum apply at several levels. In a morphological sense, anaphoric pronouns are typically free elements, whereas agreement markers are typically bound inflectional elements. Syntactically, anaphoric pronouns and agreement markers have opposite binding conditions (agreement markers must have their antecedent within the same domain, pronouns outside their domain). Another difference is that whereas pronouns can generally index all available case roles in the language, agreement markers are normally restricted to a single case role (the absolutive or nominative: cf. Corbett 2003: 172). Finally, there are differences in semantics and pragmatics, such as the fact that agreement markers (but not free pronouns) are fully compatible with all kinds of referential status. Whether the antecedent is definite or indefinite, generic or not, should not have any consequences for an agreement marker in a purely syntactic agreement system. Semantic distinctions, like animacy and person, that have been

<sup>1</sup>I thank the speakers of Yurakaré who have taught me their language for sharing their knowledge with me. I would furthermore like to thank Grev Corbett, Michael Cysouw, and an anonymous reviewer for commenting on earlier drafts of this paper. All remaining errors are mine. The research reported in this paper was made possible by grants from Prof. Pieter Muysken’s Spinoza project Lexicon & Syntax, the University of Surrey, the DoBeS foundation, and the Netherlands Organization for Scientific Research, for which I am grateful.

shown to interact with some agreement systems (cf. Siewierska 2004: 148–62) should also have no consequences for a purely syntactically determined agreement marker.

This paper is dedicated to exploring the intermediate field between personal pronouns and syntactic agreement markers by studying the pronominal affixes of Yurakaré, an unclassified language spoken in central Bolivia. Yurakaré has affixed subject and object person markers that have most characteristics of agreement markers, except for the fact that the object prefixes can distinguish many case roles (contrary to prototypical agreement markers) and that both subject and object affixes allow for their antecedent to be left unexpressed within the same clausal domain. I argue that the system in Yurakaré (and in many other polysynthetic languages) is rather economical, and may be more advantageous than systems that do not index grammatical relations on the verb and systems that have syntactic agreement markers.

The paper is organised as follows. First, I sketch the contours of the pronoun–agreement continuum by defining the extremes, and discussing the differences between them (Section 2). In Section 3, I introduce argument marking in Yurakaré, and in Section 4, which forms the bulk of the paper, I study the pronominal affixes of the languages in the light of the discussion in Section 2. In Section 5 I argue for the economical nature of the system in Yurakaré, and in Section 6 I hypothesise what this may mean for diachronic developments of person-marking systems. I conclude with a summary in Section 7.

## 2. THE PRONOUN–AGREEMENT CONTINUUM

At first sight, pronouns and agreement markers seem to be rather different. A pronoun can be defined as ‘special case of the more general linguistic category “proforms”, i.e. (semantically empty or nearly so) function words that replace (lexical content-bearing) syntactic units of a particular category’ (Saxena 2006: 131).<sup>2</sup> Functionally, personal pronouns ‘are used for denoting speech roles and for indicating their involvement in the events or states that the sentences in which they occur describe’ (Bhat 2004: 9). Demonstrative pronouns can fulfil the second function of personal pronouns, and in addition they may encode contextual and/or inherent characteristics of the referent.

Agreement can be defined as a ‘systematic covariance between a semantic or formal property of one element and a formal property of another’ (Steele 1978: 610). Canonical agreement, i.e. the most extreme definition of agreement (Corbett 2006), is redundant rather than informative, syntactically simple (i.e. it can be captured by simple and general rules), and is encoded by affixal morphology.

A number of differences between canonical pronouns and agreement markers follow from the above definitions. They cover several linguistic levels: morphophonology, syntax, semantics and pragmatics.

A first obvious difference between canonical pronouns and agreement markers regards their morphological status. Agreement markers are canonically bounded (affixal) inflectional elements. Pronouns, given the fact that they ‘replace’ nouns, are expected to be free elements, like nouns.

Syntactically, one of the differences between pure pronouns and agreement markers is the fact that they have opposite binding domains. The antecedent of a personal pronoun must be outside the clause, whereas the antecedent of an agreement marker must be within the same clause (or any syntactic environment where the agreement occurs, but we focus on the clause here).

<sup>2</sup>Saxena mentions that the problem of first and second person pronouns can be solved by saying that ‘the replaced NPs are the names of the speaker(s) or addressee(s), etc.’.

Another argument with respect to syntax is related to case marking. One of the principles used by Bresnan & Mchombo (1987) and Corbett (2003) to differentiate anaphoric (pronoun-like) agreement from syntactic agreement is to consider case roles. The idea behind this criterion is that ‘in the undisputable cases of verb agreement, we find that normally only one case role can be indexed. This may be the subject, or it may be the absolutive argument’ (Corbett 2003: 172). Free pronouns normally have the full range of case-marking possibilities (cf. Bhat 2004, who considers case-role expression one of the core functions of pronouns).

Demonstrative pronouns (which are often the only pronominal device to refer to third persons) typically mark some contextual (deictic) information about their referent. Agreement markers, being syntactically (and semantically) simple, apply independently of the situational context. Personal pronouns are more like agreement markers than demonstrative pronouns in this respect, since personal pronouns normally do not indicate any of the identifying characteristics of their referent either (cf. Bhat 2004: 9).<sup>3</sup>

A final difference that I will mention here is pragmatic in nature. Pronouns, like nouns, can be marked for pragmatic status, especially emphasis. Agreement markers, being inflectional morphology, are not expected to be able to encode information status.

When maximally contrasted like this, it seems unlikely that pronouns and syntactic agreement – the latter being very rare cross-linguistically (see e.g. Siewierska 1999; Corbett 2006) – can be confused. However, polysynthetic languages pose a problem to the distinction between these two phenomena. These languages tend to have bound pronominal markers that are like personal or demonstrative pronouns in some respects, and like agreement markers in others. In fact, a continuum can be drawn up between free pronouns at one end and syntactic agreement markers at the other, and pronominal affixes in polysynthetic languages will be at different points between these extremes.

In the next two sections, I present a discussion of the bound pronominal markers of one such polysynthetic language, Yurakaré (Van Gijn 2006), and examine them in the light of the pronoun–agreement opposition.

### 3. PARTICIPANT INDEXING IN YURAKARÉ

Yurakaré is an unclassified language spoken in central Bolivia in one of the westernmost fringes of the Amazon rainforest by approximately 2,500 speakers. It has an elaborate system of marking arguments on the predicate by means of pronominal affixes in combination with applicative constructions. It marks at least one and at most three arguments on the verb by means of prefixes (objects) and suffixes (subject), according to a nominative–accusative alignment.<sup>4</sup> The features involved in participant marking on the predicate are person, number and, to a lesser extent, collectivity. Pronominal affixes are partially formally related to the forms of the independent pronouns with which they may but need not co-occur. Apart from the head-marking strategies, Yurakaré also has a number of postpositional enclitics indicating the case role of a participant: locative, directional, ablative, instrument (with extensions to path and reason) and comitative. These postpositional enclitics attach to the rightmost element of the noun phrase. When an argument is marked on the verb, it cannot at the same time be marked by a postposition on the antecedent. The strategies are in complementary distribution, and often there is no choice between them. I will not discuss the postpositions here, since the paper focuses on the head-marked participants.

<sup>3</sup>First and second persons may also encode gender distinctions (cf. e.g. Siewierska 2004: 106), though it is a marginal pattern, cross-linguistically. I thank Michael Cysouw for bringing this point to my attention.

<sup>4</sup>There are some predicates that seem to have an ergative alignment, but these predicates are of the experiencer type (cf. Van Gijn in press), which have a cross-linguistic tendency to deviate from the standard transitive pattern in a language (cf. Malchukov 2005).

Strictly speaking, Yurakaré independent personal pronouns distinguish two persons and two numbers:<sup>5</sup>

- (1) sĕĕ<sup>6</sup> 1SG  
 mĕĕ 2SG  
 tuwa 1PL  
 paa 2PL

Reference to 3SG is made by means of demonstrative pronouns, which can be pluralised by means of the marker =w<sup>7</sup> to refer to 3PL.

- (2) ana this  
 ati that  
 naa that (further away)

These distinctions correspond to the following cross-reference markers for subjects:

- (3) -y 1SG -tu 1PL  
 -m 2SG -p 2PL  
 -∅ 3SG -∅=w<sup>8</sup> 3PL

Direct objects are cross-referenced by the following set of prefixes:

- (4) ti- 1SG ta- 1PL  
 mi- 2SG pa- 2PL  
 ∅-/ka- 3SG ma- 3PL

Subject and direct object markers are simply attached to the predicate stem:

- (5) ma-bobo-y  
*3PL-hit/kill-1SG.S*  
 'I hit/kill them.'

The choice between forms zero or *ka-* for 3SG object is determined by collectivity, which is a partly transparent semantic feature in Yurakaré (cf. Van Gijn 2004). Collective singular nouns (i.e. bounded mass) are cross-referenced with the marker *ka-*. All other direct objects are unmarked:

<sup>5</sup>The following abbreviations are used in this paper: 1 first person, 2 second person, 3 third person, ABL ablative, ACC accusative, BEN benefactive, CMT commitment, DEF definite, DEM demonstrative, DO direct object, DSC discontinuative, EXH exhortative, FMR former, FRS frustrative, GO goal, HAB habitual, HYP hypothetical, IC involuntary comitative, IMP imperative, INDEF indefinite, INS instrument, INT intentional, IPF imperfective, IRR irrealis, JUS jussive, M masculine, MAL malefactive, MDG medial degree, MOM momentaneous, NEG negation, P possessive, PL plural, PRN pronoun, PRS personal assessment, REF referential, REP reportative, RES resignative, S subject, SG singular, SS same subject, THM theme, VC voluntary comitative.

<sup>6</sup>The written vowel ě represents the low front vowel /æ/.

<sup>7</sup>The morphological status of the plural marker is rather unclear: it seems to be in a process of changing from a clitic to a suffix.

<sup>8</sup>I assume a pronominal zero element because of the nominal plural marker in 3PL subject marker, which needs a (pro)nominal element to have scope over, thus also capturing the analogy with the free pronouns.

- (6) a. *bějti samma*  
*see:1SG.S water*  
 ‘I see water.’  
 b. *ka-bějti samma*  
*3SG-see:1SG.S water*  
 ‘I see a river/lake.’

Apart from the direct object, five more types of object can be distinguished on semantic and formal grounds. These objects are cross-referenced by the pronominal affixes given in (4), but without the collective/non-collective distinction: all third person singulars are cross-referenced with *ka-*. In addition to the pronominal affixes, all but one type of object require an additional marker, which can be called an applicative marker (see Table 1). The involuntary comitative object is minimally distinguished from the direct object, the only difference being the third person singular, which in the case of the involuntary comitative is always marked *ka-*, regardless of collectivity:

Table 1. Yurakaré applicatives

Meaning	Marker	Applies to
involuntary comitative	∅	intransitives
voluntary comitative	vowel change	intransitives, transitives
benefactive	-n-	intransitives, transitives
malefactive	-la-	intransitives, transitives
goal	-y-	intransitives

- (7) a. *tiya-y ti-chata* (direct object)  
*eat-1SG.S 1SG-food*  
 ‘I ate my food.’  
 b. *ka-mala-y ti-chata* (involuntary comitative)  
*3SG-go.3SG-1SG.S 1SG-food*  
 ‘I took my food.’ (lit. ‘I went with my food’)

The vowel change for the voluntary comitative yields the following paradigm:

- (8) *tě-* 1SG      *tu-* 1PL  
*mě-* 2SG      *pu-* 2PL  
*ku-* 3SG      *mu-* 3PL

I do not go into the details of the semantics of the different applied objects here; for that the reader is referred to Van Gijn (2005; 2006: 149–60).

#### 4. PRONOMINAL AND AGREEMENT-LIKE CHARACTERISTICS OF YURAKARÉ PERSON MARKERS

In this section, I discuss the characteristics and behaviour of the Yurakaré pronominal affixes in the light of the pronoun–agreement continuum described in Section 2. I look at a number of areas of the language where differences between pronouns and agreement markers are expected to become apparent: morphological boundedness, binding domains, case marking, referentiality, pragmatic status, and semantic aspects of the antecedent.

#### 4.1. Morphological boundedness

An obvious first observation on the Yurakaré verbal person markers which makes them look rather like agreement markers is that they are bound elements. Corbett (2006: 13) proposes the following condition for canonical agreement markers (targets) with respect to their morphological form:

(9) *Target*

Bound > free

Inflectional marking (affix) > clitic > free form

This condition says that, in canonical agreement, the marker on the target is a bound form rather than a free form – more precisely, an inflectional affix rather than a clitic or, even less canonically, a free form. The scales in (9) can possibly also be read (from right to left) as the diachronic path agreement markers typically have followed.

The person markers in Yurakaré are best analysed as affixes. They attach to the verb, either before (objects) or after (subjects) the stem; they also interact with word stress (as do positional clitics in Yurakaré), and they may enter into morphophonological processes with elements in their immediate morphological surroundings. They furthermore appear in fixed slots; this is especially clear with the object markers. The following template can be given for their ordering:<sup>9</sup>

- (10) *position* 3    2    1            0  
           MAL DO APPL. MRKR. ROOT  
           BEN GO  
           VC IC

The template should be read as follows: there are two positions for the object person markers before the root (positions 2–3), meaning that there can be at most two head-marked objects per predicate; the person markers referring to the direct object, goal and involuntary comitative appear closest to the root and are mutually exclusive. Further to the left, in position 3, the pronominal prefixes referring to the other three object types can appear. They are also mutually exclusive.<sup>10</sup>

Nevertheless, there is an idiosyncratic peculiarity about the 1SG in terms of morpheme ordering. It concerns the marker for the imperative plural. This marker normally appears attached to the inflected predicate before all object person markers,<sup>11</sup> but in combination with 1SG benefactives and voluntary comitatives there is a different order: the person marker precedes the imperative marker.

- (11) a. ti-pi-n-dyuju = chi

*1SG-IMP.PL-BEN-tell* = *FRS*

‘Tell me!’

<sup>9</sup>The terms for the applied objects here refer to the person marker only (or the portmanteau prefix in the case of the voluntary and the involuntary comitative); the applicative markers themselves are positioned after the cross-reference complex.

<sup>10</sup>I have one example of the malefactive with the benefactive, which is non-canonical in terms of both semantics and form. I consider it to be an exception.

<sup>11</sup>There is also a suffix variant of the imperative plural marker, but that is of no concern here.

- b. *të-pë-yuda*            *lacha*  
*1SG.VC-IMP.PL-help* *too*  
 ‘Help me too!’

This idiosyncrasy, however, is not enough to propose clitic status for the object markers.

#### 4.2. *Binding domains*

A syntactic difference between pronouns and agreement markers is the fact that they have opposite binding domains. The antecedent of a personal pronoun must be outside the clause, whereas the antecedent of an agreement marker must be inside the same clause (or any syntactic environment where the agreement occurs, but we focus on the clause here).

This difference in binding conditions gives us the following heuristic: ‘the (im)possibility of a free pronoun occurring in the same clause as the marker gives an indication as to the status as a verbal marker’ (Corbett 2003: 186). Siewierska (1999) sets up a tripartite typology of agreement markers on this basis, between pronominal, ambiguous, and syntactic agreement markers. In pronominal agreement systems, there is an absolute prohibition on the co-occurrence of the controller within the agreement domain (as e.g. in the Cariban language Macushi: Abbott 1991). At the other end of the continuum, there are languages with syntactic agreement, where the controller occurs obligatorily in the same domain as the agreement marker (e.g. in English). In between these two extremes of the continuum there are languages which allow for co-occurrence, but without it being required (ambiguous agreement).

Yurakaré is an example of the ambiguous agreement type. A count of a number of texts shows that subjects of intransitive predicates and direct objects are ‘dropped’ in around 60 per cent of cases. Subjects of transitive clauses and applied objects are both dropped in about 82 per cent of cases.<sup>12</sup> The use of free pronouns for head-marked arguments seems to differ from text to text. The smallest proportion of the use of free pronouns cross-referenced on the verb that I have found constitutes 2 per cent of all arguments, and 3.4 per cent of all overtly expressed arguments. The highest proportion is 32.7 per cent of all arguments, and 51.7 per cent of all overtly expressed arguments. Most, but not all, uses of overt pronouns can be explained by the fact that they express certain characteristics that cannot be expressed by the bound pronouns, such as referential and pragmatic status differentiation. Other factors may be related to the speech situation. In the following fragment, a woman is talking to her grandchild. She uses more free pronouns than the average, and this may be due to the fact that her grandchild is not very fluent in Yurakaré.

- (12) a. *ti-ëtějñu bay-tu = naja*            *putu ta-ø-ya = naja*  
*1SG-son go.EXH-1PL.S = DSC Putu say-3 = REP = DSC*  
 ‘My son, “let’s go”, Putu said.’
- b. **ku-bata-m**                    **mëë**            *mi-choo = bë?*  
*3SG.VC-go.EXH-2SG.S 2SG.PRN 2SG-uncle = MOM*  
 ‘Are you going with your uncle?’
- c. *mi-choo putu* **ku-tütü-n-tu**            **tuwa = bë?**  
*2SG-uncle Putu 3SG.VC-be-INT-1PL.S 1PL.PRN = MOM*  
 ‘Shall we stick with your uncle Putu?’

<sup>12</sup>Intransitive subjects were dropped in 295 of 472 cases (62.5%), direct objects were dropped in 164 of 285 cases (57.5%). Transitive subjects (i.e. subjects of transitive, extended intransitive and ditransitive verbs) were dropped in 447 of 550 cases (81.3%), applied objects in 275 of 329 instances (83.6%).

- d. numma = jsha bay-**tu**            **tuwa**  
*night* = *ABL*    *GO.EXH-1PL.S*    *1PL.PRN*  
 ‘Let’s go tomorrow.’
- e. lëtta puyni enseñe-**ni**            ta-buybu = 1a  
*one*    *sun*    *teach-INT:1SG.S*    *1PL-language* = *INS*  
 ‘I will teach our language for one more day.’
- f. tütü-n-**tu** = naja itta achaya **më** = ja            gana-cha-**m** uruju  
*be-INT-1PL.S-DSC*    *thing*    *so.that*    *2SG.PRN* = *THM*    *earn-JUS-2SG.S*    *Uruju*  
 ‘Let’s stay a bit, so that you earn some money, Uruju.’

Differences in the extent to which free forms are used in systems with obligatory person markers on the verb is an interesting but difficult topic for cross-linguistic research. For Yurakaré, further research is needed to establish the constraints on the use of free forms more accurately.

#### 4.3. Case marking

One of the principles used by Bresnan & Mchombo (1987) and Corbett (2003) to differentiate anaphoric agreement from syntactic agreement is to consider case roles. The idea behind this criterion is that ‘in the undisputable cases of verb agreement, we find that normally only one case role can be indexed. This may be the subject, or it may be the absolutive argument’ (Corbett 2003: 172). Free pronouns normally have the full range, or at least a considerable range, of case-marking possibilities. Therefore, if agreement markers have a range of case-marking possibilities comparable to the set found on free pronouns, they are best analysed as pronominal arguments. This is connected to the idea that canonical agreement is redundant rather than informative (Corbett 2003; 2006).

As we have seen, the pronominal affixes in Yurakaré can be marked for various distinctive case roles. This is an important argument for treating them as pronominal elements. However, the pronominal affixes of Yurakaré do not have the same case-marking possibilities as free pronouns. Grammatical relations in Yurakaré are either marked on the head (applicatives) or on the dependent (postpositions); for details see Van Gijn (2005). A schematic overview of the possibilities is given in Figure 1.

As can be seen in Figure 1, most cases that can be expressed on the head do not have an equivalent in dependent marking.<sup>13</sup> The only exception are the comitative-type cases. However, these are not entirely equivalent either, as head-marked comitatives have a formally marked two-way distinction (voluntary versus involuntary), while the corresponding comitative postposition does not make this distinction. Dependent-marked comitative and head-marked comitative do not co-occur. Either one or the other is used. The choice between the two is largely determined by topicality: the more accessible and central in the preceding discourse the referent, the more likely it is to be head-marked. With the voluntary comitative, there is an extra animacy requirement (cf. Van Gijn 2005).

Of course one could object that this simply concerns the locus of the case-relation marking, and does not mean that the pronominal prefixes themselves are case-marked, and that the applicative marker case-marks the controller (whether present or not). This is unlikely for two reasons. The obvious first one is that we can distinguish three different (though related) pronominal paradigms: direct object vs. voluntary comitative vs. other objects. In the forms

<sup>13</sup>There is partial overlap in functionality in some instances (e.g. directionals and benefactives can both indicate endpoints of movements, ablative and malefactive can both indicate sources of movement), but the point is that the head-marked and dependent-marked case roles have non-overlapping functions as well.

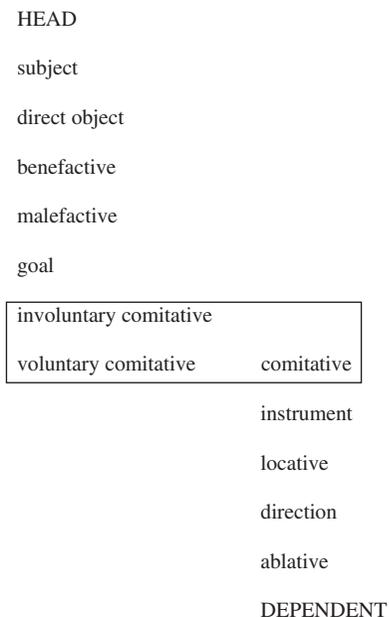


Figure 1. Head marking and dependent marking of case relations in Yurakaré

of the object prefixes themselves three different kinds of objects are distinguished at least paradigmatically. A second reason is the placement behaviour of the prefixes. I showed in (10) that there are strict ordering principles governing object markers, where direct object, involuntary comitative, and goal are closest to the root, and the other three (benefactive, malefactive, and voluntary comitative) are further away. The fact that the bound pronouns can be distinguished for so many case roles is a strong indication of the more pronominal nature of the markers, even though they cannot be equated with free pronouns in this respect.

#### 4.4. Referentiality and definiteness

Evans (2002: 16), studying object affixes<sup>14</sup> in Biniñ Gun Wok, notes that ‘bound object affixes in at least some polysynthetic languages pattern more like subject agreement in European languages than like free pronouns, in that they specify person and number information while remaining non-committal about reference and discourse status’.

Referentiality concerns the relation between a linguistic sign and what that sign stands for in the context of the utterance (a person, an object, etc.). Free personal pronouns and demonstratives ‘are regarded as “inherently” semantically referential with little or no descriptive content’ (Chen 2009: 1658). They are thus, prototypically at least,<sup>15</sup> associated with referential contexts. Definiteness encodes identifiability of a referent. Definite referents are expected to be accessible to the hearer, and easily activated in the hearer’s mind. Personal or demonstrative pronouns share these characteristics, and require activation or saliency in

<sup>14</sup>Evans specifically talks about object affixes in response to a claim made by Siewierska (1999) that object affixes never seem to make it to becoming agreement markers; but the quote is in principle applicable to bound subject markers in many polysynthetic languages as well.

<sup>15</sup>There are exceptions to this claim, e.g. non-referential *it* in English weather verbs, or pronouns in ‘donkey sentences’, non-referential uses of *you*: see Evans (2002: 17), Chafe (1994: 103–4), Mithun (2003: 240–41).

the hearer's mind probably even more than definite NPs. Pronouns, then, are also associated with definite contexts.

Canonical agreement markers, on the other hand, appear regardless of the referential or definiteness value of their antecedent. This falls under one of Corbett's (2006) general principles of canonical agreement: that agreement is syntactically simple, i.e. it is not sensitive to any conditions, and the appearance of agreement markers thus does not depend on contextual characteristics of the referent.

The bound pronominal affixes of Yurakaré, being obligatory, co-occur with antecedents that can have all kinds of referential and definiteness status. Definite and indefinite referents trigger the same person marking on the verb:

- (13) a. **ma**-bëbë-ni            ti-sse = w  
           *3PL-search-INT:1SG.S 1SG-child = PL*  
           'I am going to look for my children.'  
       b. **ma**-bëbë-ni            turibbinñu  
           *3PL-search-INT:1SG.S parrot:PL*  
           'I am going to look for some parrots.'

Even non-referential, generic antecedents are treated in the same way:

- (14) a. kormena **ma**-duya-jti-ø  
           *bee 3PL-sting-HAB-3*  
           'A bee stings [people].'  
       b. otto-tu            tuwa            ele = jsha            ta-jti-ø = w  
           *go.out-1PL.S 1PL.PRN ground = ABL say-HAB-3 = PL*  
           'They say [it is said] that we came out of the ground.'

This type of interpretation only seems to be possible when the predicate is marked for habitual aspect. However, the fact that the predicate is marked for habituality does not force the non-generic reading. Examples (14a, b) are ambiguous. Note, however, that the verb *-ta-* 'say' does not mark the addressee, whereas this is obligatory in referential contexts:

- (15) yita = la            ku-ta-ø = ya            a-tata  
           *good = CMT 3SG.VC-say-3 = REP 3SG.P-father*  
           '“Good!”, her father said to her.'

A non-habitual alternative for non-referential unspecified objects is with the use of the imperfective prefix *a-*. The direct object participant of transitive verbs in such cases is unmarked:

- (16) a-ense-jti-ø = w = ya            ana            yee = w  
           *IPF-drink-HAB-3 = PL = REP dem woman = PL*  
           'The women were drinking.'

In these constructions, the object of the verb is non-referential generic, and not expressed. It can also be combined with an explicit NP, which must be generic and grammatically singular, as can be observed in the following examples for the transitive verb *bëbë* 'search for' (Van Gijn 2006: 186).

- (17) a. a-bëbë-y  
*IPF-search.for-1SG.S*  
 ‘I am searching.’
- b. a-bëbë-y            petche  
*IPF-search.for-1SG.S fish*  
 ‘I am searching for fish.’
- c. \*a-bëbë-y            petche = w  
*IPF-search.for-1SG.S fish = PL*  
 ‘I am searching for fishes.’

Plural generic objects are cross-referenced on the verb:

- (18) ma-bëbë-y            petche = w  
*3PL-search.for-1SG.S fish = PL*  
 ‘I search(ed) for fishes.’

In imperfective contexts, referential objects are encoded by the malefactive applicative:

- (19) ka-l-a-bëbë-y            petche  
*3SG-MAL-IPF-search.for-1SG.S fish*  
 ‘I am searching for a (particular) fish.’

Related to referentiality and definiteness is the behaviour of quantifiers such as *nobody* and *every(body)*, where the presence of person markers co-indexed with these quantifiers is inconsistent with a bound pronoun interpretation, and consistent with an agreement interpretation. Quantifiers like *no one* and *everyone* function in a peculiar way in Yurakaré. A word for *no one* as such does not exist, but is formed as a combination of *lëtta* ‘one’ with the suffix *-mash* ‘medial degree’ and the negation marker *nish/nij* before the predicate. The combination has a free translation as ‘not even one’:

- (20) a. lëtta-mash nij buyusa-jti-ø yee  
*one-MDG NEG kiss-HAB-3 woman*  
 ‘No one kissed a woman.’
- b. lëtta-mash nij ka-yile bëshëë  
*one-MDG NEG 3SG-know stuff*  
 ‘No one knows things.’

In the second example it can be observed that the quantifier *no one* is cross-referenced on the verb (*ka-*). The verb *ayle* (here: ‘to know’) has a deviant participant-marking pattern with the experiencer encoded as an (involuntary comitative) object, the stimulus being the subject.

The quantifier ‘all’ *-mumuy-* or *-mumay-* functions as a noun in that it takes possessive prefixes. It can be translated as ‘all of it, us, you or them’, depending on the person value of the possessive prefix. This quantifying expression functions as an argument and is cross-referenced on the verb:

- (21) a. ma-mmuy shunñe = w ma-buyusa-jti-ø = w yee = w  
*3PL-all man = PL 3PL-kiss-HAB-3 = PL woman = PL*  
 ‘All the men kissed the women’

- b. nish ma-mmuy ma-yle-jti bëshëë  
 NEG 3PL-all 3PL-know-HAB stuff  
 ‘Not everyone knows about things.’

#### 4.5. Pragmatic status

Van Valin (2005: 73) posits a markedness hierarchy for the combination of type of referring expression and pragmatic status (topic or focus) (see Figure 2). Even though pronouns may not be the best candidates for focus marking, free personal and demonstrative pronouns can normally be specified for pragmatic status by stress, morphological marking, or position in the clause. Bound pronouns do not have the same optionality as free pronouns in this respect, but polysynthetic languages may distinguish between focused and non-focused antecedents/controllers in that person markers only appear in combination with a non-focused controller (Siewierska 1999).

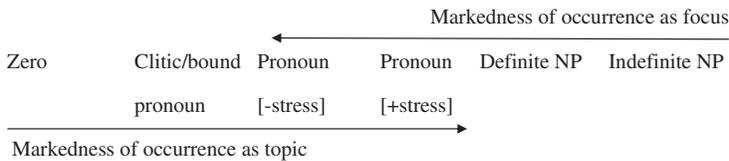


Figure 2. Markedness relations of referring expressions and pragmatic status (Van Valin 2005)

In Yurakaré, whether a referent is in focus or not has no consequences for person marking. The interaction with agreement can be tested by seeing what happens with questioned elements, typically in focus:

- (22) a. tětë-pshë = w = ri      ma-bobo-m = chi  
*what-thing = PL = RES.M 3PL-kill-2SG.S = FRS*  
 ‘What things did you kill?’  
 b. tětë-pshë ti-ja-n-wita-m  
*what-thing 1SG-3SG-BEN-arrive.SG-2SG.S*  
 ‘What did you bring us?’

In (22a), the questioned element is the direct object. As it is plural, it is cross-referenced on the verb as a 3PL direct object (*ma-*). In (22b), the questioned element is a benefactive object, and as such it is cross-referenced on the verb (*ja-*).<sup>16</sup>

#### 4.6. Other semantic and pragmatic distinctions

Siewierska (1999) also notes tendencies of person markers in polysynthetic languages to show co-occurrence restrictions with their antecedents in terms of person, part of speech, and animacy. These can be stated in terms of hierarchies for which the generalisation is if the

<sup>16</sup>The marker *ja-* is a phonologically conditioned allomorph of *ka-*; [k] generally changes to [h], spelled *j*, after a vowel.

antecedent displays characteristics that are higher up the hierarchy (further to the left), there will be a stronger preference for person agreement.

- (23) a. Nominal hierarchy  
           pronoun > noun  
       b. Person hierarchy  
           1st > 2nd > 3rd  
       c. Animacy hierarchy  
           human > animate > inanimate > abstract

These hierarchies can also be read to reflect environments where pronouns would typically or exclusively be used towards the left, whereas towards the right pronouns are not or less often used. We might hypothesise that they represent aspects of pathways of change from bound pronouns towards agreement markers.

For Yurakaré, these hierarchies have no influence on whether or not pronominal affixes appear. The nominal hierarchy (23a) has no bearing on Yurakaré agreement:

- (24) a. *l-ati = w*      *ma-wita = ya*      **ma-che-jti-ø**  
           REF-DEM = PL    3PL-arrive = IRR.SS    3PL-eat-HAB-3  
           ‘Those he brought and ate.’  
       b. *ayma ma-che*    *ta-ñuma-shama = w*  
           fire    3PL-eat-3    1PL-mother-FMR = PL  
           ‘The fire burned (ate) our ancestors.’

The direct-object marker on the verb *che* ‘eat’ is identical in both (24a) and (24b); in (24a) the marker is co-indexed with a pronoun (*latiw* ‘them’); in (24b) with a noun (*tañumashamaw* ‘our late mothers’).

With regard to the person hierarchy (23b), it might be said that Yurakaré has no person agreement with the third person, only number agreement:

- (25) a. *bobo*  
           *hit*  
           ‘He hits him.’  
       b. *ma-bobo-w*  
           PL-*hit-PL*  
           ‘They hit them.’

However, the absence of a phonological form should be interpreted as grammatical person marking, because they do refer, as can be seen in (25a). Moreover, the fact that there is a plural marker with the 3PL which is identical to the nominal plural marker, and which also coincides with plural marking pattern in the free pronouns (suppletive for first and second person, transparent for third persons), suggests a zero pronominal element referring to third persons.

As for object marking, the opposition zero vs. *ma-* can also be said to exhibit only number marking. However, in other paradigms, and also in some circumstances in the direct object paradigm, the third person has an overt marker *ka-* or a variant, which still stands in opposition to *ma-* for 3PL.

As for the animacy hierarchy (23c), agreement occurs with humans, animates and inanimates:

- (26) a. ta-ñu        **ma**-bějta = ya    awēwē-tu    chajtiya  
*1PL-child:PL 3PL-see = IRR.SS cry-1PL.S always*  
 ‘Always when we see our children, we cry.’
- b. petche = w    **ma**-bějta = ya    amala = ya    ma-bobo-ta-∅ = w = laba  
*fish = PL 3PL-see = IRR.SS come = IRR.SS 3PL-kill-HYP-3 = PL = PRS*  
 ‘Possibly, when they see these fishes, they will come to kill them.’
- c. lëshie    chishta = w    **ma**-bějti  
*two sieve = PL 3PL-see:1SG.S*  
 ‘I see two sieves.’

Yurakaré generally avoids the use of abstract nouns, and if they are used, they are usually in subject or oblique role, so there is no visible evidence on the verb of cross referencing.

##### 5. BOUND PERSON MARKERS AND ECONOMY

As appears from the above discussion, the Yurakaré bound pronominal markers have all the characteristics of agreement marking, except for two: they do not require their antecedent to be present within the same domain, and they are capable of distinguishing many case roles. I will argue that this ‘polysynthetic’ system is advantageous when compared to the major alternatives.

Andrews (2007), discussing the functions of the noun phrase, makes a distinction between semiotic (semantic and pragmatic) functions, grammatical functions, and coding features. Semiotic functions have to do with the meaning of a sentence (e.g. semantic roles and pragmatic functions). Grammatical functions encode relationships between units in the sentence which serve to facilitate the interpretation of the semiotic functions. The grammatical functions depend on the grammatical structure of a language, i.e. the coding features a language has available in its structural make-up.

Because of the fact that Yurakaré has object prefixes and an elaborate applicative system, the case-role encoding function, facilitating the interpretation of the semantic roles participants play can be fulfilled to a large degree by the pronominal affixes.<sup>17</sup> The obligatoriness of the pronominal affixes, however, makes them unsuitable for distinguishing between different referential or pragmatic statuses. Languages with obligatory pronominal affixes may deal with this problem in a number of ways.

Navajo, apart from having referentially used pronominal prefixes, has three special sets of pronominal prefixes, used for generic reference, reference to unspecified participants, and what is called ‘areal’ or ‘spatial’, which is, among other things, used ‘for ambient conditions, characteristics of an area, and much more’ (Mithun 2003: 265). Indefinite participants can be cross-referenced with the referential prefixes, provided they have been established in the previous context, which also may be within the same clause (p. 266).

Hungarian (Kenesei, Vago, & Fenyesei 1998) marks definiteness of object participants on the verb, fused with subject inflection (pp. 257–8):

- (27) a. lát-om        a    kutyá-m-at  
*see-DEF.1SG the dog-1SG.P-ACC*  
 ‘I see my dog.’

<sup>17</sup>Limitations are that maximally three participants can be marked on a verb, and certain relations must be indicated by postpositions, as discussed above.

- b. lát-ok            egy kutyá-t  
     *see-INDEF.1SG a dog-ACC*  
     ‘I see a dog.’

The system of definiteness marking in Hungarian is much more complex than this, and raises some interesting questions on the relativity of the notion of definiteness (see Kenesei et al. 1998: 321–7). For the current paper, it suffices to show that definiteness can have a place in the verbal conjugation paradigm.

The solution to the referentiality/pragmatic status problem that Yurakaré chooses<sup>18</sup> is to leave the expression of definiteness, referentiality, and pragmatic status to the free forms, whereas case-role assignment is largely accomplished by the pronominal affixes, in combination with the applicative system. This division of labour allows for pro-drop, which can be restated as an economy rule: do not express an argument as a free NP unless required. The reasons for expressing arguments as free NPs in polysynthetic languages can differ from language to language, and arguably relates to a large degree to the structure and informativeness of the paradigm (but see the discussion in Section 4.2). General reasons for expressing arguments as free NPs, regardless of the structure of the person-marking paradigm, are disambiguation and adding lexical content. For Yurakaré, if a referent is focused, or indefinite, expression of a free NP is normally required. Non-expression of an indefinite referent as a free form yields the interpretation of generic referents like ‘people’, ‘someone’.

This type of system, which leaves open the possibility of applying the economy rule discussed above, can be said to combine the advantages of systems that do not mark participants on the verb with those of agreement systems. The advantage of non-verb-marking languages over agreement languages is that they do not express the same information twice, which is more economical. The functionality of agreement markers is discussed in Corbett (2006: 274–5), on the basis of Levin (2001: 21–7). Three suggestions are put forward, all with their problems and exceptions, leading Corbett to conclude that there is no single function for agreement markers. The three suggestions are that (i) agreement markers facilitate the hearer’s task by repeating information, (ii) they contribute to reference tracking, (iii) they mark constituency. The last suggested function is arguably more valid for within-NP agreement, and not so much for verb agreement markers. The first two functions of agreement markers are shared by syntactic agreement markers and pronominal affixes that allow for their antecedent to be absent. In fact, it can be argued that pronominal affixes, which are normally assumed to be phonetically more salient and semantically more informative than syntactic agreement markers, are better at performing these functions than syntactic agreement markers. For reference tracking, for instance, it helps to distinguish different numbers and genders in your paradigm (Corbett 2006: 275), and repeating more information arguably facilitates the hearer’s task even further.

Pronominal affixes that allow for the absence of their antecedent within their domain are therefore functionally optimal: they combine (part of) the economy of non-verb-marking systems in situations of pro-drop with the functionality of agreement markers in non-pro-drop situations. Extending this system to include more case-roles by means of applicative constructions only extends the situations to other types of participants.

<sup>18</sup>I do not mean to sharply contrast Yurakaré to languages like Central Alaskan Yup’ik or Navajo, which also allow for the non-expression of arguments as free NPs.

## 6. REMARKS ON DIACHRONY

The diachronic path from free pronoun to agreement marker is a widely discussed topic (see e.g. Corbett 2006; Siewierska 1999; 2004; Givón 1976). It involves changes on several dimensions: a morphosyntactic change from free forms to bound forms, a syntactic change from an optional marker to an obligatory marker (with different binding conditions), and a functional-semantic change from informative referring expressions to (in principle) redundant markers. The process broadly involves three stages:

- (i) morphological attachment of pronouns;<sup>19</sup>
- (ii) required presence of bound pronominal forms;
- (iii) phonetic attrition of bound pronominal forms and obligatory expression of antecedents.

The general idea of the process is that the morphosyntactic change comes first<sup>20</sup> (either as a result of clitic doubling or morphological reduction and bounding as a result of discourse principles or frequency; cf. Siewierska 2004). The second change results in polysynthetic systems like the one in Yurakaré. The obligatory presence of the bound pronominal elements leads to the loss of pronominal referential qualities (see Evans 2002),<sup>21</sup> and to the co-occurrence of the bound pronouns with their antecedent in the same clause.<sup>22</sup> Frequency of use of the obligatory bound pronouns leads to phonetic attrition, and as a result loss of informativeness of the person markers. The consequence of this is that antecedents have to be expressed more and more often until they eventually are required.

The development of applicative markers is in principle independent of the development of pronominal affixes, since there are languages that do not have (object) pronominal affixes but which do have applicative constructions. However, in his typological study of applicative constructions, Petersen (2007: 209) found that languages with applicatives tend to mark more arguments on the verb by means of verbal indexing (an average of 1.72 participants in languages with applicatives versus an average of 1.32 for non-applicative languages, on the basis of a 100-language sample, 50 with and 50 without applicative constructions). One possible motivation for the rise of bound pronominal markers is that they represent highly accessible (topical) information, and the more accessible a referent, the less coding required (see also Figure 2 above). The main motivation for the direct rise of applicative markers is that they allow a peripheral and oblique-marked participant to be expressed as a topical (unstressed pronominal) core argument (Petersen 2007: 141–2). A tentative reason for the correlation found by Petersen is that pronouns, together with their case-role markers, grammaticalise and become bound markers.

In any case, even if applicative markers in a language grammaticalise at a different point in time than the pronominal affixes (in either order), the result is a situation in which bound object pronominal affixes can be marked for more than one case role. In this situation, object affixes may become obligatory, and may lose their ability to indicate specific referential and pragmatic statuses, but they are still informative with respect to case roles.

In such a situation, phonological erosion of the person-marking paradigm leads to a loss of informativeness with respect to distinctions between the different person and/or number

<sup>19</sup>I gloss over the complexities that may be involved in this process of morphological attachment (see e.g. Cysouw 2003).

<sup>20</sup>This is an empirical observation made in Evans (2002), who emphasises that there is no a priori reason to exclude the possibility of obligatory pronominal forms that are not morphologically bound.

<sup>21</sup>The data of Navajo and Hungarian, however, show that obligatoriness does not entail the loss of certain referentiality distinctions; there can be separate markers within the paradigm.

<sup>22</sup>This scenario is slightly different when topic-shift constructions are at the basis of the diachronic development (see Givón 1976). The result, however, is the same.

distinctions, leading in turn to a decrease of the number of situations where the economy rule of pro-drop can apply, and eventually to a situation where the person markers have become redundant agreement markers that necessarily have their antecedent within the same domain (although this seems to be an exceptional, if not unattested, situation for object affixes; see Siewierska 1999). However, other development scenarios are also imaginable. For instance, the applicative system may arise or expand after the object markers have become obligatory, or (as seems to be the case in Yurakaré) the applicative case roles may develop semantically in such a way that they do no longer have clear dependent-marked counterparts. In these alternative scenarios, the division of labour between verbal pronominal affixes and free NP arguments with respect to NP functions, which was argued above to be functionally optimal, is maintained and further developed and strengthened.

## 7. CONCLUSION

Pronominal affixes in polysynthetic languages often have characteristics of both pronouns and agreement markers. This is because the continuum between free pronouns on the one hand and agreement markers on the other encompasses several different levels (morphology, semantics, pragmatics), which do not necessarily align functionally. One of the aspects that connects many polysynthetic languages is that they allow for pro-drop, i.e. the expression of core arguments as free NPs is constrained by a rule of economy. It was argued that such a system in fact combines economy of non-verb-marking languages with the functional advantages of agreement markers.

From a diachronic perspective, polysynthetic pronominal affixes are often seen as an intermediate phase of a diachronic process in which free pronouns develop into bound pronouns, which in turn develop into agreement markers. The change from free to bound pronoun precedes the obligatory presence of the pronominal elements, and with the obligatory presence comes the loss of referentiality of the pronominal markers. Syntactic agreement markers are then claimed to arise when the person-marking paradigm undergoes phonetic attrition and loses its informativeness to such a degree that the presence of the antecedent is required (a rare situation cross-linguistically).

However, given the functional optimality of pro-drop systems, languages may develop elaborate systems of (obligatory) verbal person marking which fulfil several functions performed by NPs in non-pro-drop languages, involving special affixes indicating definiteness or referential distinctions (Navajo, Hungarian), or, as is the case in Yurakaré, by an elaborate applicative system distinguishing several different case roles.

*Department of Linguistics  
Radboud University Nijmegen  
PO Box 9103  
6500 HD Nijmegen, Netherlands  
Email: e.vangijn@let.ru.nl*

## REFERENCES

- ABBOTT, MIRIAM, 1991. 'Macushi', in C. Derbyshire & Geoffrey K. Pullum (eds), *Handbook of Amazonian Languages*, vol. 3, Berlin: Mouton de Gruyter, 23–160.
- ANDREWS, AVERY D., 2007. 'The major functions of the noun phrase', in Timothy Shopen (ed.), *Language Typology and Syntactic Description*, vol. 1: *Clause Structure*, Cambridge: Cambridge University Press, 132–223.
- BHAT, D. N. S., 2004. *Pronouns*, Oxford: Oxford University Press.
- BRESNAN, JOAN & MCHOMBO, SAM, 1987. 'Topic, pronoun and agreement in Chicheŵa', *Language* 63, 741–82.
- CHAFE, WALLACE, 1994. *Discourse, Consciousness, and Time*, Chicago: University of Chicago Press.
- CHEN, PING, 2009. 'Aspects of referentiality', *Journal of Pragmatics* 41(8), 1657–74.

- CORBETT, GREVILLE G., 2003. 'Agreement: the range of the phenomenon and the principles of the Surrey Database of Agreement', *Transactions of the Philological Society* 101(2), 155–202.
- CORBETT, GREVILLE G., 2006. *Agreement*, Cambridge: Cambridge University Press.
- CYSOUW, MICHAEL, 2003. 'Towards a typology of pronominal cliticization', paper presented at the 5th conference of the Association for Linguistic Typology (ALT-V), Cagliari, Italy: [http://web.me.com/cysouw/presentations/index\\_files/cysouwCLITICS\\_handout.pdf](http://web.me.com/cysouw/presentations/index_files/cysouwCLITICS_handout.pdf)
- EVANS, NICHOLAS, 2002. 'The true status of grammatical object affixes: evidence from Bininj Gun-wok', in Nicholas Evans & Hans-Jürgen Sasse (eds), *Problems of Polysynthesis*, Berlin: Akademie Verlag, 15–50.
- GIVÓN, TALMY, 1976. 'Topic, pronoun, and grammatical agreement', in Charles N. Li (ed.), *Subject and Topic*, New York: Academic Press, 151–88.
- KENESEI, ISTVÁN, VAGO, ROBERT M. & FENYESI, ANNA, 1998. *Hungarian*, London: Routledge.
- LEVIN, MAGNUS, 2001. *Agreement with Collective Nouns in English*, Stockholm: Almqvist & Wiksell.
- MALCHUKOV, ANDREI, 2005. 'Case pattern splits, verb types and construction competition', in M. Amberber & H. de Hoop (eds), *Competition and Variation in Natural Languages: The Case for Case*, Amsterdam: Elsevier, 73–118.
- MITHUN, MARIANNE, 2003. 'Pronouns and agreement: the information status of pronominal affixes', *Transactions of the Philological Society* 101(2), 235–78.
- PETERSEN, DAVID A., 2007. *Applicative Constructions*, Cambridge: Cambridge University Press.
- SAXENA, ANJU, 2006. 'Pronouns', in Keith Brown (ed.), *Encyclopedia of Language and Linguistics*, 2nd edn, Amsterdam: Elsevier, 131–3.
- SIEWIERSKA, ANNA, 1999. 'From anaphoric pronoun to grammatical agreement marker: why objects don't make it', *Folia Linguistica* 33(2), 225–51.
- SIEWIERSKA, ANNA, 2004. *Person*, Cambridge: Cambridge University Press.
- STEELE, SUSAN, 1978. 'Word-order variation: a typological survey', in Joseph H. Greenberg, Charles A. Ferguson & Edith A. Moravcsik (eds), *Universals of Human Language IV: Syntax*, Stanford, CA: Stanford University Press, 585–623.
- VAN GIJN, RIK, 2004. 'Number in the Yurakaré noun phrase', in Leonie Cornips & Jenny Doetjes (eds), *Linguistics in the Netherlands 2004*, Amsterdam: Benjamins, 69–79.
- VAN GIJN, RIK, 2005. 'Head marking and dependent marking of case relations in Yurakaré', in Mengistu Amberber & Helen de Hoop (eds), *Competition and Variation in Natural Languages: The Case for Case*, New York: Elsevier, 41–72.
- VAN GIJN, RIK, 2006. 'A grammar of Yurakaré', Ph.D. dissertation, Radboud University, Nijmegen.
- VAN GIJN, RIK, in press. 'Subjects and objects in Yurakaré: a semantic account', in A. Guillaume & F. Rose (eds), *Argument Coding Systems in Bolivian Lowland Languages*, special issue of the *International Journal of American Linguistics*.
- VAN VALIN, ROBERT D., JR., 2005. *Exploring the Syntax–Semantics Interface*, Cambridge: Cambridge University Press.