

A FIRST COMPARISON OF PRONOMINAL AND DEMONSTRATIVE SYSTEMS IN THE CARIBAN LANGUAGE FAMILY*

Sérgio Meira

Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

1. Introduction

The Cariban language family is composed of approximately 25 languages (numbers ranging from 20 to 50, depending on different researchers' opinions about which varieties are dialects and which are independent languages), spoken by approximately 100,000 people in lowland South America, from south-eastern Colombia (where Karihona is spoken) to the Oiapoque river in Brazil (Karinya), from the coast of the Guianas (Karinya) down to the southern Xingu area in central Brazil (Bakairi).

The field of comparative Cariban studies was initiated more than two hundred years ago, when the relationship between a number of Cariban languages was first noticed by Filippo Salvatore Gilij (1782). Unfortunately, the historical-comparative method has been only very rarely applied to Cariban languages, for two main reasons: (1) most of the languages are, to this day, poorly known, which means that there is very little reliable material to compare; (2) most people who compared Cariban languages were not trained comparativists. Girard (1971) remains the only methodical attempt at reconstructing Proto-Cariban lexical items and proposing a classification (unfortunately based on a still very poor data base). In the area of morphosyntax, Gildea (1998) presents the first reconstruction of the person-marking and tense-aspect-mood (TAM) systems of Proto-Cariban and their syntactic properties.

The present work attempts to contribute to the development of historical studies in the Cariban family by presenting a first preliminary reconstruction of the pronominal system of Proto-Cariban (including non-third-person and third-person, i.e. anaphoric and demonstrative, pronouns). For this purpose, the available sources (cf. Table 1 below) were scanned in search of pronouns, which were then sorted in cognate sets (Tables 2-4), according to what is known about the sound correspondences between Cariban languages (taking Girard 1971 as a guide), so as to draw conclusions on the historical evolution of these forms.

In Section 2 below, the sources and standardized transcription are introduced. Section 3 has a summary discussion of pronouns in Cariban languages, which is the main background for the rest of the paper. The actual reconstructions are discussed in

* This paper is a revised version of a short comparison of Cariban demonstrative systems that was presented at the 50th International Congress of Americanists in Warsaw. It includes a significant amount of new material, and discusses also non-third-person pronouns. I wish to thank Hein van der Voort and Mily Crevels for comments on an earlier version, and Ana Carla Bruno and Bruna Franchetto for sharing their data. Any remaining mistakes are, of course, my own.

Sections 3.1 (non-third-person pronouns) and 3.2 (third-person pronouns). In Section 4, a summary table presents the reconstructed forms, followed by some further speculative comments on the relationships between these forms.

2. Sources and transcription

Sources on Cariban languages, as is the case with most other language families in lowland South America, are very different in their level of reliability, accuracy, and breadth of coverage. For some languages, the best available sources are word lists from the last century; for others, there are recently published high-quality grammatical descriptions and occasionally even dictionaries. In view of that, the actual availability of data was a factor of importance in the selection of the languages to compare. Table 1 contains a list of the languages and sources selected for this study. Data from the best sources (marked with ‘++’ in Table 1) is assumed to be good in all respects; missing pronouns from these sources will thus be considered as non-existent. The less good sources (marked with ‘+’ in Table 1), and especially the worst sources (unmarked), are less reliable, and need to be handled with care. Mis-transcriptions and inadequate phonological analyses are a real danger; missing pronouns may in many cases actually result from gaps in the data.

Languages	Sources	Languages	Sources
Tiriyó	Meira 1999, 2000; fn ++	Tamanaku	Gilij 1965[1782]
Akuriyó	fn ++	Cumanagoto	Yangues 1683, Ruiz Blanco 1690
Karihona	Robayo 1987, 2000a; fn ++	Chayma	Tauste 1680
Hixkaryana	Derbyshire 1979, 1985 ++	Pemón	Armellada & Olza 1994 +
Waiwai	Hawkins 1998; fn ++	Taurepán	Koch-Grünberg 1916
Katxuyana	fn; Gildea’s fn +	Makushí	Abbott 1991, Amodio & Pira 1996 ++
Karinya	Hoff 1968, Mosonyi 1978 ++	Ingarikó	Koch-Grünberg 1916
Apalaí	Koehn & Koehn 1986; fn ++	Arekuna	Edwards 1977; Koch-Grünberg 1916 +
Wayana	Jackson 1972; fn ++	Akawayo	Edwards 1977; fn +
De’kwana	Hall 1988 +	Panare	Muller 1994 ++
Yawarana	Méndez-Arocha 1959 +	Ikpeng	Pacheco 1997; Gildea’s fn +
Yukpa	Vengamián 1978 ¹	Arara	Souza 1992
Waimirí-	Bruno 1996; Bruno’s fn	Bakairí	Steinen 1892, Wheatley 1973, 1978 +
Atroarí		Kuhikuru	Franchetto’s fn ++

Table 1. Cariban languages and sources considered in this work. The ‘+’ signs mark the best sources; fn = field notes (Meira’s if unidentified). The dotted lines identify probable lower-level genetic subgroups.²

The different transcription systems of the various sources have been standardized to facilitate comparisons. Certain details have been ignored (e.g. Koch-Grünberg’s

¹ A couple of forms also from Robayo (2000b).

² Tiriyó, Akuriyó, and Karihona were classified together already in Girard (1971) and in Kaufman (1994); Meira (2000) proposed the name *Taranoan* for this subgroup and reconstructed part of the grammar and vocabulary. Gildea (pers. comm.; cf. also 1998:91-96) considers Hixkaryana, Waiwai, and Katxuyana a subgroup, which he named *Parukotoan*. The other groupings, present in Girard (1971) and Kaufman (1994), contain very closely related languages or dialects.

attempt at marking non-phonemic distinctions between [e], [ɛ] and [o], [ɔ], here transcribed simply as *e*, *o*; his stress marks — *á*, *é*, etc. — were also left out). The symbols in need of explanation are: *ĩ* = IPA [i], *ẽ* = [ɛ], *j* = [j] (a palatal glide, English *y*), *ɣ* = [x], *ɣ* = [ɣ] (velar fricatives), ' = [ʔ], *ñ* = [ɲ], *x* = [ʃ], *tx* = [tʃ]. Sequences of identical vowels (*aa*, *ee*, etc.) are phonetically long.

The language names, which often vary from source to source, have been respelled here for clarity, avoiding language-specific uses ('Karihona' instead of the Spanish 'Carijona'). Names with final stress have a diacritic mark (e.g. Makushí, Apalaí); the others have penultimate stress (e.g. Arekuna, Akawayo = Arekúna, Akawáyo).

3. *Cariban pronominal systems*

A typical Cariban system has pronouns for: first person (1), second person (2), first person dual inclusive (1+2), first person exclusive (1+3),³ and third person (3). In terms of number, Cariban languages oppose 'collective' (i.e. focus on a group) vs. 'non-collective' (i.e. focus on less than a group, but not necessarily a single individual). Usually, there is a second person collective pronoun (2Col), based on the non-collective form plus a collective ending, a first person inclusive collective (1+2Col), based on the first person dual inclusive form, and third-person collectives; the first person exclusive form (1+3) is unmarked for number. Note that the first-person pronoun does not have a collective form; semantically, the 1+3 and 1+2Col forms play this role.

The third-person pronouns form a relatively complicated system, including anaphoric and demonstrative (proximal, medial, distal) forms which, along with number (collective vs. non-collective), distinguish also animacy.

Derbyshire (1999:53-54) gives a first comparative overview of Cariban pronominal systems. For the sake of convenience, we shall follow his system of separating non-third-person from third-person pronouns as two subsystems, discussed in Sections 3.1 and 3.2, respectively.

3.1. *Non-third-person forms*

The pronominal forms to be compared, from the sources in Table 1, are listed in Table 2 below. As can be seen, they seem to form good cognate sets.

The *first-person* forms can be first divided in those that end in *ro* or *rë*, and those that do not. The same syllable is present at the end of other pronouns in many other languages. Considering its frequency (14 occurrences), one might feel tempted to reconstruct it, at least to some intermediate level. However, for the following reasons, this is not a good idea: (i) this syllable has no clear cognates in the other languages (the final *wĩ* found in several languages cannot be compared to *ro* ~ *rë*,

³ Syntactically, the 1+3 form is treated as a third person (e.g. verbs agree with it as if it were a third person pronoun); one wonders if it could have been an old non-possessible noun (cf. e.g. Brazilian Portuguese *a gente* 'we', literally 'the people').

since there is no regular *w* : *r* correspondence in the Cariban family); (ii) it has an obvious source in the ‘emphatic’ particle *ro* or *rě*, synchronically attested in most Cariban languages (e.g. Tiriyó *wĩ* ‘I’, *wĩ rě* ‘really me’, ‘yes, that’s me’; cf. Hoff 1990:508 for *ro* in Karinya [Carib of Surinam], Derbyshire 1985:250 for *ro* in Hixkaryana). It seems best to assume that the endings *ro* and *rě* result from the reanalysis of the emphatic particle as part of the stem (much like *otros* in Spanish *nosotros*, *vosotros*).⁴

Language	1 I	2 you	1+2 you and I	1+3 we excl.	1+2Col all of us	2Col all of you
Tiriyó	<i>wĩ</i>	<i>ěmě</i>	<i>kĩmě</i>	<i>anja</i>	<i>kĩměnjamo</i>	<i>ěměnjamo</i>
Akuriyó	<i>wĩ</i>	<i>ěmě</i>	<i>kĩmě</i>	<i>anja</i>	<i>kĩ(mě)njamo</i>	<i>ěměnjamo</i>
Karihona	<i>ěwĩ</i>	<i>ěměřě</i>	<i>kĩměřě</i>	<i>aña</i>	<i>kĩñamoro</i>	<i>añamoro</i>
Hixkaryana	<i>uro</i>	<i>omoro</i>	<i>kĩwro</i>	<i>amna</i>	<i>kĩwjamo</i>	<i>omñamo</i>
Waiwai	<i>owĩ</i>	<i>amoro</i>	<i>kĩwĩ</i>	<i>amna</i>	<i>kĩwjam</i>	<i>amjamro</i>
Katsuyana	<i>owĩ</i>	<i>omoro</i>	<i>kĩmoro</i>	<i>amna</i>	<i>kĩmjarĩ</i>	<i>omjarĩ</i>
Karinya-Hf ⁵	<i>au</i>	<i>amooro</i>	<i>kĩxko</i>	<i>a'na</i>	<i>kĩxkaaro</i>	<i>amĩjaro</i>
Karinya-Ms	<i>aau</i>	<i>amooro</i>	<i>kĩm'ooro</i>	<i>na'na</i>	<i>kĩm'oññaaro</i>	<i>amoññaaro</i>
Apalaí	<i>ĩwĩ</i>	<i>omoro</i>	<i>kĩmoro</i>	<i>ĩna</i>	<i>kĩmarokomo</i>	<i>amarokomo</i>
Wayana	<i>ĩu</i>	<i>ěmě</i>	<i>kunmě</i>	<i>emna</i>	<i>kunměramkom</i>	<i>ěměramkom</i>
De'kwana	<i>ĩwĩ</i>	<i>ěmědě</i>	<i>kĩwĩ</i>	<i>ñña</i>	<i>kĩnwanno</i>	<i>ěnwanno</i>
Yawarana	<i>wĩrě</i>	<i>měřě</i>		<i>ehně</i>		
Yukpa	<i>awĩ, awě</i>	<i>amo, amor</i>		<i>nana</i>		<i>amora, amoja</i>
Waimirí	<i>awĩ, aa</i>	<i>amĩ, amĩrĩ</i>	<i>kĩkĩ</i>	<i>a'a</i>		<i>(amĩrĩrĩ)</i>
Tamanaku	<i>ure</i>	<i>amare</i>	<i>kĩwe</i>	<i>jumna</i>	<i>kikemo</i>	<i>amñamoro</i>
Cumanagoto	<i>ure</i>	<i>amuere</i>		<i>amna</i>		<i>amia(mo)rkom</i>
Chayma	<i>ure, utxe</i>	<i>amuere</i>	<i>kutxe</i>	<i>amna</i>	<i>kutsekon</i>	<i>amiamorkon</i>
Pemón	<i>(j)ure</i>	<i>amare</i>	<i>(j)uureto</i>	<i>in(n)a</i>	<i>juurenokon</i>	<i>amarenokon</i>
Taurepán	<i>ĩjurě</i>	<i>amaarĩ</i>		<i>iná</i>	<i>ĩjurĩnĩkon</i>	<i>amaarĩnĩkon</i>
Makushĩ	<i>uurĩ</i>	<i>amĩrĩ</i>	<i>uurĩ'kon</i>	<i>anna</i>	<i>uurĩ(')nĩkon</i>	<i>amĩrĩ(')nĩkon</i>
Akawayo	<i>urě</i>	<i>amě(rě)</i>		<i>ina, nja</i>	<i>urě'nogon</i>	<i>aměřě'nogon</i>
Ingarikó	<i>ĩure</i>	<i>aměěre</i>			<i>kiulenĩkon</i>	<i>tĩmĩlĩnĩkon</i>
Arekuna	<i>jurě</i>	<i>aměřě</i>		<i>inna</i>	<i>jurětokon</i>	<i>aměřě(k)nokon</i>
Panare	<i>ju</i>	<i>aměn</i>	<i>juto, juta</i>	<i>ana</i>	<i>jutakon</i>	<i>aměnton</i>
Ikpeng	<i>uro</i>	<i>omro, omo</i>	<i>ug(u)ro</i>	<i>tximna</i>	<i>ugroymmo</i>	<i>omroymmo</i>
Arara	<i>uro</i>		<i>ugoro</i>	<i>tximna</i>	<i>ugoroymmo</i>	
Bakairĩ	<i>urě</i>	<i>ěmě</i>	<i>kurě</i>	<i>(x)ina</i>		<i>(a)mareemo</i>
Kuhikuru	<i>uře</i>	<i>e(e)ře</i>	<i>kukurye</i>	<i>tisurye</i>		<i>amayyo</i>

Table 2. Cariban non-third-person pronouns. Elements in parentheses did not occur consistently.

⁴ Interestingly, in Hixkaryana, the new first-person pronoun *uro* has given rise to a new first-person prefix *ro-*, *r-* (e.g. *ro-jĩmĩ* ‘my father’), which has replaced an earlier Proto-Cariban **u-* (cf. Gildea 1998). Note also that the final syllable *ro*, *rě* occurs as *rĩ* in Makushĩ and Arekuna (cf. below for Makushĩ *ě* > *ĩ*), and as *ře* in Kuhikuru (for which *ř* : *r* and *e* : *ě* are also regular correspondences: cf. Kuhikuru *uře* ‘manioc bread’, *tehu* ‘stone’, Tiriyó *uru*, *těpu*). The cases of *re* (Pemón, Ingarikó, Tamanaku, Cumanagoto, Chayma) are certainly mistranscriptions of *rě*.

⁵ Hoff (1968) and Mosonyi (1978) describe mutually intelligible dialects of the same language (‘Carib’ for Hoff, ‘Cariña’ for Mosonyi), here labeled ‘Karinya’. They are here treated independently (Hoff’s as Karinya-Hf, Mosonyi’s as Karinya-Ms) because their pronouns differ in form.

If we ignore the final *ro* or *rě*, all first-person forms seem to contain a *w*, or a reflex of it in the form of the vowel *u*; the longer forms contain a preceding and a following vowel (**VwV*). The second vowel can be reconstructed as *i*, and the cases of *u* can be seen as the result of vowel loss and syllable reduction (**Vwĩ > Vu > u*). The first vowel, however, is a more difficult case: there are forms beginning with *a*, *o*, *ě* and *i*. Such problematic vowel correspondences are not infrequent in Cariban languages, due to (often irregular) vowel assimilation (cf. Girard 1971:79). A final reconstruction must wait for better lower-level comparisons. For a tentative reconstruction, consider that: (a) *i* is often the result of the weakening of an earlier vowel in Cariban languages, possibly as a first step in the process of syllable reduction and loss (Gildea, pers. comm.; cf. Gildea 1995 on Cariban syllable reduction); (b) *ě* and *o* seem to be diachronically related (cf. the second-person pronouns in Table 2), so that the *ě*- and *o*-initial forms are probably not independent. Taking (a) and (b) into account, **a* is the best tentative reconstruction: with a following *w*, an **a > o* assimilation would be much more natural than **o > a* (cf. Gildea 1998:83-84 for a similar argument concerning the reconstruction of the second-person prefix **a(j)-*). Tentatively, one could suggest a protoform **awĩ*.

Three problematic details remain, for which some suggestions are presented here. (1) Tiriyo *ĩ* is probably the result of a metrical reanalysis of pronoun-clitic sequences: e.g. Pre-Tiriyo **ěwĩ rě* ‘really me’ would go from [əví:ɾə] to [ví:ɾə] by losing the initial vowel, at which point the surface long vowel would be reanalyzed as underlying *ĩ* (or else it would become short — [vĩɾə] —, as in all CVCV words; cf. Meira 1998, 1999 on the stress system), thus yielding *wĩ rě*. (2) The initial *j* in Panare, Pemón, Makushí, and Taurepán may result from the resyllabification of an earlier **Vw > *iw* sequence (e.g. **iwĩ rě > *iu rě > juurĩ*).⁶ (3) Chayma *txe* is rather puzzling; one might suggest that an element *txe* was added to an earlier **u* (still attested in *u-re*), maybe by analogy with the 1+2 form *kutxe* (but note that the *txe* in *kutxe* is also of mysterious origin).

In the *second-person* forms, one can again exclude the final syllables that reflect the particle *ro ~ rě*: *ro*, *rě*, *rĩ*, *re*, *ye*, and also Waimirí *ra* and De’kwana *dě* (De’kwana *d* often corresponds to *r* in other languages: cf. *jwũdĩ* ‘tapir’, Tiriyo *ĩwĩrĩ*). Panare *n* is also a likely reflex of an earlier *rě*; cf. Panare *tunkě* ‘horsefly’, *akuñ* ‘agouti’, Tiriyo *turěkě*, *akuri*. The cases of long vowels in the second syllable (Karinya, Taurepán, Ingarikó, Arekuna) are probably phonetic effects of the rhythmic stress system (cf. Meira 1998); Yawarana *měērě*, on the other hand, may represent a case of underlying *ěě* resulting from the loss of the initial vowel, like Tiriyo *wĩ* (cf. above). Looking at what remains, the second consonant *m* is almost always present (except in Kuhikuru; cf. below) and can safely be reconstructed, together with two

⁶ Taurepán *jũ* looks like an attempt at transcribing what could have been an intermediate stage (something like e.g. *ĩũ:*). Cf. the case of Portuguese *eu* [ew] and Spanish *yo* [jo], which have stressed different parts of an earlier **eo < Lat. ego*.

adjacent vowels: **VmV*. The exact nature of the vowels is less clear; one can only make tentative suggestions.

For the first vowel, one has the possibilities *a*, *o*, *ě*. As was mentioned above, *ě* and *o* may not be independent, which would reduce the choice to *a* vs. *ě/o*. Considering that the second vowel was probably *ě/o*, the *ě/o* cases in the first vowel could be the consequence of assimilation (**a > o*, possibly made easier by the intervening labial **m*), whereas the *a* cases are harder to derive from **o*. The best hypothesis is thus **a*.

For the second vowel, one basically has *ě/o*: the cases of *ue* (Cumanagoto, Chayma) and *a* (Tamanaku, Taurepán) are probably mistranscriptions of *ě*, and the cases of *ĩ* (Waimirí, Makushí, Arekuna) look like reflexes of *ě* (*ě : ĩ* is attested in Makushí, as in e.g. *sikĩ* ‘flea, chigger’, Tiriyo *sikě*; Waimirí has no phonemic *ě*). As was mentioned above, *ě* and *o* are clearly related; there are numerous examples of the *ě : o* correspondence (e.g. Tiriyo *sikě* ‘flea, chigger’, Apalaí *xiko*). Gildea (pers. comm.) considers *ě* to be always a reflex of Proto-Cariban **o*, which is quite plausible phonetically. There are, however, *o : o* correspondences without apparent conditioning factors (e.g. Tiriyo *okomo* ‘wasp’, Apalaí *okomo*; cf. the second-person forms in Table 2). The question of whether *o : o* and *ě : o* are independent correspondences has not yet been settled. Taking a conservative stance, **o* will be reconstructed for *o : o*, and **o₂* for *ě : o*.⁷ One thus ends up with a tentative proto-form **amo₂*. The last problem is the unexpected Kuhikuru form *e(e)-ye* (long *ee* attested in the author’s [Meira’s] field notes; short *e* attested in Franchetto’s field notes). One idea could be intervocalic *m*-loss: Pre-Kuhikuru **eme-ye > e(e)-ye*. However, all attested cases of *m*-loss in Kuhikuru are word-initial, not word-internal (e.g. Kuhikuru *oto* ‘worm’, Tiriyo *moto*). It seems thus better to suppose that the initial **e* was lost first: **eme-ye > *me(e)-ye > *e(e)-ye*. (The long *ee*, in case it is not a transcription mistake, might result again from the influence of an earlier rhythmic stress system, as in the case of Tiriyo *wĩĩ*.)

The *first-person dual inclusive* (1+2) forms show more complex patterns. After eliminating the reflexes of the particle *ro ~ rě*, there are two major groups: (a) forms that contain the initial element *ku*, *kĩ*, *ki*, and (b) forms that contain an initial element *ju*, *u* (Panare, Pemón, Makushí; presumably, the other languages of the same group also have similar forms, unfortunately unattested). The best idea seems to be, since there is no initial *k* loss rule for the (b) languages, that these two groups of forms are not cognate. In fact, the *(j)u*-initial forms all seem to be based on the first-person plus a final element *to*, *ta*, *kon*, all reminiscent of number (collective) markers (e.g. Tiriyo *ton*, *kon*, Apalaí *tomo*, *komo*, etc.; cf. below the discussion of collective forms). This would imply a path of evolution whereby an original 1+2 form was lost and replaced

⁷ Note that the *o : o* and *ě : o* correspondences have distinct reflexes in Kuhikuru: e.g. *tehu* ‘stone’, Tiriyo *těpu*, and *oti* ‘field, grass’, Tiriyo *oi*; cf. also Kuhikuru *okō* ‘wasp’. Thus, Pre-Kuhikuru apparently had **o* and **o₂*. Considering the number of (not obviously closely related) languages that have **ě*, it is not impossible that Proto-Cariban **o₂* was actually **ě*. Not much, however, can be said without a detailed study of the distribution of *o : o* and *ě : o* in the family.

with an analytical 1 + Col form. One may further suggest that this form had originally collective, not simply dual, meaning, and that the collective forms (which have additional collective suffixes) may have originally been more emphatic synonyms. The Ingarikó and Arekuna forms would thus represent — in case they are not simple mistranscriptions — a retention of original *k* forms.

The *k*-initial languages all share an initial syllable reconstructible as **kV*. Given the overwhelming majority of cases of *kĩ*, the first idea is to reconstruct **kĩ*. The cases of *ku*, however, give food for thought. First, *ku* occurs in Wayana and in the Southern languages (Kuhikuru, Bakairi, Arara, Ikpeng), which are as far away from each other as is possible within the family. One may consider also the earlier mentioned tendency for vowels to ‘weaken to *ĩ*’, and also the fact that *k* is not an obvious environment for labialization (**ĩ > u*). On the other hand, the possibility of deriving *ku* from an earlier **kĩwĩ*, at least for some languages (cf. below), must be borne in mind. All in all, reconstructing **kĩ* seems to be still the best tentative hypothesis.

The second syllable of the *k*-initial forms, however, varies quite wildly; it does not seem possible to view *mě* (*mo*, *m^wo*, *nmě*), *wĩ* (*wi*, *we*), *txe*, *χko*, *ku* (*gu*, *go*) as all cognate. Rather, it would seem that an initial element **ku* (probably the same as the 1+2 prefix that Gildea (1998:92, 114) reconstructs as Proto-Cariban **k(i)-*) was added to several independent elements (maybe old possessible nouns) to make 1+2 pronouns; even dialects may end up with different forms (e.g. Karinya: *kĩχko* [Hf], *kĩm^wooro* [Ms]). The various forms can be separated in several groups, which correspond only imperfectly to proposed subgroupings (e.g. in Kaufman 1994): the *mě* group (Tiriyó, Akuriyó, Karihona [= Meira’s Taranoan], Wayana, Apalaí, Katxuyana, Karinya-Ms; most of Kaufman’s Guianan branch plus two Central branch languages; tentative reconstruction **kĩ-nmo₂*),⁸ the *wĩ*-group (Hixkaryana, Waiwai [= Gildea’s Parukotoan without Katxuyana], De’kwana, Tamanaku; two Guianan and two Central branch languages; tentative reconstruction **kĩwĩ*),⁹ and the *ku*-group (Kuhikuru, Arara, Ikpeng, Waimirí; the Southern branch without Bakairi, plus one North Amazonian language; tentative reconstruction **kuku*). Bakairi might be added to the *ku*-group (so that it includes all of Kaufman’s Southern branch) by assuming that *ku-rě* actually results from **kuku ro₂* (which would also yield Arara-Ikpeng *ug(u)ro*, *ugoro* if one assumes the loss of the initial *k*). It is not unthinkable that Karinya-Hf *kĩχko* is related to the *ku*-group: **kuku + *ko* could yield present-day *kĩχko*, but not **kĩnmo₂ + *ko* or **kĩwĩ + *ko* (cf. Gildea 1995 on syllable reduction). Chayma *kitxe* remains isolated.

⁸ Notice that *m^wo* instead of *mo* in *kĩm^wooro* represents no problem, since Mosonyi’s (Venezuelan) Karinya has rules of palatalization and labialization of consonants depending on the quality of the adjacent vowels; the long *oo* results from the rhythmic stress system. The *n* in Wayana *kunmě* is less readily explained; it is tentatively reconstructed, despite the rather strange absence of its reflexes in the other *mě* languages.

⁹ The long *ĩ* in Waiwai *kĩwĩ* results from an idiosyncratic change (probably related to the stress system) that lengthened the first vowel of all CVCV words.

Going farther than this means going into the realm of speculation, which, all in all, is not a bad source of ideas. One first notices that **kīwĩ* is not implausible as a source for the **ku* forms (e.g. Wayana *kunmẽ* < **kīu nmo*₂ < **kīwĩ nmo*₂; for the **ku* group, one might have e.g. **kīwĩ ro*₂ > **kīu ro*₂ > Bakairí *kurẽ*). However, this leaves the ‘double-*ku*’ forms (Kuhikuru *kuku*, Arara *ugo*, Ikpeng *ug(u)*) unexplained, and also Karinya *kĩχko*; the lack of any reflex of the syllable *wĩ* in Tiriyó, Akuriyó, Karihona, Apalaí, and Katxuyana (one would expect at least a long vowel) is a further difficulty. One might also suspect that a simple **ku* could have been the original source of both the 1+2 pronouns and the 1+2 person-marking prefix; it may even have been an independent element at some point (maybe still preserved in Bakairí *kurẽ* < **ku-ro*₂), and would later on have blended with other elements (erstwhile independent nouns). However, the evidence for this element as an independent word in Proto-Cariban is very scant (Bakairí, the only apparent case of retention, could also result from **kuku* + **ro*₂ with syllable reduction).

Thus, in view of the variety of forms, it does not seem possible to reconstruct the form of a 1+2 pronoun to Proto-Cariban. Notice that it must have existed, since there are 1+2 pronouns in all languages (even those who lost the **ku*-forms innovated new 1+2 pronouns) and the 1+2-marking prefix can be reconstructed; its form, however, must remain unreconstructed. This fact will be represented with the formula **kīCV* for the presumed Proto-Cariban 1+2 pronoun.

The *first person exclusive* (1+3) forms, like the first person dual inclusive forms, are also all apparently partially, but not completely, cognate. All forms end in *na* (Yawarana *ehnẽ* possibly explained by weakening, and Waimirí *a’a* maybe from an earlier **a’na*), so that a final syllable **na* can be reconstructed.

However, the initial syllables, like the final syllables in 1+2 forms, clearly do not form a single cognate set. One can separate the attested forms into: an *n-ñ* or *palato-alveolar* group (*a-nj*, *a-ñ*, *a-nn*, *i-nn*, *i-n*, *ñ*: Tiriyó, Akuriyó, Karihona [Meira’s Taranoan], De’kwana, Yukpa,¹⁰ Pemón, Taurepán, Makushí, Akawayo, Panare, and probably also Apalaí and Bakairí;¹¹ there are members of Kaumfan’s Guiana, North Amazonian, and Central branches; tentative reconstruction, **a-in(n)a*); an *m* or *labial* group (*a-m*: Hixkaryana, Waiwai, Katxuyana [Gildea’s Parukotoan], Cumanagoto, Chayma, and probably also Wayana *e-m*, Tamanaku *ju-m*, and Ikpeng-Arara *txi-m*; tentative reconstruction, **a-m-na*); and an *’* or *glottal* group (*(n)a-’*, *e-*: Karinya, Waimirí, Yawarana; two Central branch languages, one isolate; tentative reconstruction, **a/e-h/’-na*).

At this point, one may speculate further. It would seem that the three groups could be unified if one presupposes an initial element **ap* which, in contact with an original **ina*, could then: (a) nasalize to **am* and yield *amna* with the loss of the vowel *i*, or *emna* without this loss (e.g. via **aimna* < **am-ina*), and further

¹⁰ It may be that the Yukpa form is missing a glottal stop (*na’na*), in which case it would be transferred to the glottal group.

¹¹ For Apalaí, one may suggest *ĩ* < **i* (weakening-to-*ĩ*); for Bakairí, the initial *x-* may be a later addition: notice that *xina* is found in only one of the two dialects, the other having *ina*.

assimilate to the *n*, creating *anna*, *inna*, **ainna* > *anja*, *aña*, *ana*, *ina* (the last form with ‘weakening-to-*i*’); or (b) reduce to a glottal segment, yielding Karinya-Hf *a’na* with loss of *i* (and also *n* for Waimiri *a’a*) and Yawarana *ehně* without loss (via e.g. **ai-p-na* > **e-χ-na*). This would suggest reconstructing **ap-ina* or **apina* to Proto-Cariban, as depicted in Fig. 1 below. However, the *in-/ñ*-initial form could as easily be reflexes of a simpler **ina*, without **ap*; and Tamanaku *jumna*, Ikpeng and Arara *tximna* suggest that initial elements other than **ap* could also occur (though their final *m* does suggest some relation to **ap*). The formula **(ap)ina* will be adopted here to stress the tentative status of the reconstruction of the initial element **ap*.

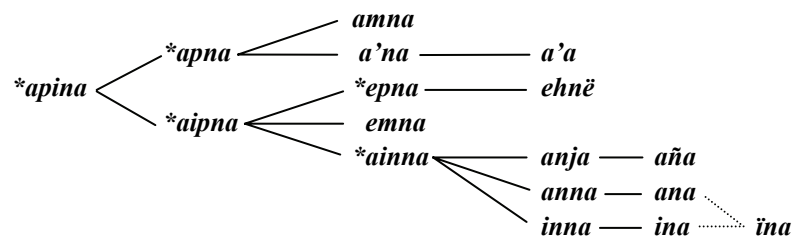


Figure 1. A speculation on the evolution of **(ap)ina* ‘1+3’. The remaining initial elements *n*, *ju*, *x*, *tx* are not included.¹²

The *collective* forms (1+2Col, 2Col) all seem to be derived from the respective non-collective forms with the help of the collective suffixes *-njamo*, *-jamo*, *-jarĩ*, *-jaro*, *-aro*, *-wanno*, *-komo*, *-kemo*, *-ton*, *-nĩ*, *-no* often more than one and not in the same order as other languages (though, with a few exceptions — Tamanaku, Panare, Arekuna —, every language uses the same suffixes in the same order for its 1+2Col and 2Col forms); the ‘emphatic’ particle *ro*, *rě* (< **ro*₂) often occurs, sometimes between suffixes. The best hypothesis seems to be the reconstruction of three collective markers, **jamo*, **komo*, **tomo* (all still attested synchronically as such in several languages), and maybe also **no*. The various collective forms would then be derived as follows:

Tiriyó *kĩmě-njamo*, *ěmě-njamo* < **jamo*
 Akuriyó *kĩ(mě)-njamo*, *ěmě-njamo* < **jamo*
 Karihona *kĩ-ñamoro*, *a-ñamoro* < **jamo ro*₂
 Hixkaryana *kĩw-jamo*, *om-ñamo* < **jamo*
 Waiwai *kĩw-jamo*, *om-ñam-ro* < **jamo (ro)*₂
 Katxuyana *kĩm-jarĩ*, *om-jarĩ* < **jamo ro*₂ (?)
 Karinya-Hf¹³ *kĩχk-aaro*, *amü-jaro* < **jamo ro*₂

¹² It is also possible to derive **apna* from **aipna*, rather than directly from **apina*; in this case, **aipna* would be Proto-Carib, and **apina* either unnecessary, or maybe pre-Proto-Cariban.

¹³ Long *aa* < **ja* (as in *amijaro*; note the short *a* here). Note that Karinya reduces nasal syllables to zero, even in synchronic morphophonology, so that **-jamo ro*₂ > *(j)aro* is not surprising (cf. *awoomĩ* ‘to get up’, *aj-aawo-ja* ‘I am getting up’).

Karinya-Ms *kīm^wo-ññaro, amo-ññaaro* < **jamo ro₂*
 Apalaí *kīm-arokomo, am-arokomo* < **jamo ro₂ komo*
 Wayana *kunmë-ramkom, ëmë-ramkom* < **ro₂ jamo komo*
 De'kwana *kĩ-nwanno, ë-nwanno* < **jamo ro₂*
 Yukpaamo-ra, *amo-ja* < *(*ro₂*) *jamo*
 Waimirí *amĩ-rĩĩ* < **ro₂ tomo*
 Tamanaku *kĩ-kemo, am-ñamoro* < **komo; *jamo ro₂*
 Cumanagoto *am-ia(mo)rkom* < **jamo ro₂ komo*
 Chayma *kutxe-kon, am-iamorkom* < **komo; *jamo ro₂ komo*
 Pemón (j) *ure-nokon, amare-nokon* < **no komo*
 Taupépán *jĩurĩ-nĩkon, amaarĩ-nĩkon* < **no komo*
 Makushí *uurĩ-(')nĩkon, amĩrĩ-(')nĩkon* < **no komo*
 Akawayo *urë-'nogon, amërë-'nogon* < **no komo*
 Ingarikó¹⁴ *kiule-nĩkon, tĩmũ-lĩnĩkon* < **no komo*
 Arekuna *jurëtokon, amërë(k)-nokon* < **tomo komo; *no komo*
 Panare *juta-kon, amën-ton* < **komo; *tomo*
 Ikpeng *ugro-ŋmo, om-roŋmo* < *(*ro₂*) *komo*
 Arara *ugoro-ŋmo* < **komo*
 Bakairí(a) *ma-reemo* < **ro₂ jamo*
 Kuhikuru *am-ayo* < **jamo ro₂*

Some suggestions for the problematic details are listed below.

- (i) For the suffix *-njamo* in Tiriyo and Akuriyo, Meira 2000:59 suggests that it results from the reinterpretation, in an earlier collective pronoun, of the **n-jamo* sequence as **-njamo*, followed by the forming of new collective pronouns with **-njamo*. He suggests the following steps for the 1+2Col form: **kĩmë + jamo* > **kĩn-jamo* > **kĩ-njamo*, **kĩmë + -njamo* > *kĩmënjamo*. For Hixkaryana *-ñamo*, the obvious answer is nasalization by the preceding *m* (**om-jamo* > *omñamo*). A similar explanation for the *ñ* in Karinya-Ms was not found thus far, but it probably exists.
- (ii) Karinya-Hf *-jaro* < **jamo ro₂*, without nasal reflex, is not surprising: Karinya loses NV syllables, even in synchronic morphophonology (cf. *awoomĩ* 'get up', *aj-aawoi-ja* 'I am getting up'). The long *aa* in the 1+2Col form probably results from syllable fusion (**kĩŋko-jaro* > *kĩŋkaaro*; cf. 2Col *amũjaro*, with a short *a*); for the Karinya-Ms forms, however, no obvious explanation was found. NV loss also occurs synchronically in Apalaí and Kuhikuru. Katxuyana *-jarĩ* is surprising, both because there usually is no NV loss in this language, and also because **ro₂* should occur as *ro*, not *rĩ*. Nevertheless, an irregular evolution of **jamo ro₂* still seems

¹⁴ The initial *k*, *t* in Ingarikó (cf. also, from Koch-Grünberg 1916, Arekuna *kuulĩnĩkon* 'we') are probably mistranscriptions (but the *k*'s might also be remnants of earlier *k*-initial forms).

less strange than a whole new collective element **rĩ* or **jarĩ* without additional evidence.

- (iii) The *nw* in De'kwana *nwanno* is hard to explain (it is not related to Karinya-Ms *m^w*; cf. fn. 7). There are some correspondences between *j* and *w* in the Cariban family (e.g. Tiriyó *ẽ-jomi* 'your language', Wayana *ẽ-womi*), so that it may still be derivable from **jamo ro₂* (> **jan-no*, with nasalization of **r*). The preceding *n*, however, remains unexplained (though it may indicate a connection with **kinmo₂* languages).
- (iv) The **no* found in the Pemón group languages (reconstructed as **no*, rather than **nĩ*, because of its frequency, including in the best documented languages: Aka-wayo, Arekuna, Pemón) is a surprising element, without obvious equivalent in the other languages. Its origin remains unknown (though one may compare it to the 'postposition collective' *-:ne*, *-'ne*, which sometimes occurs on nouns; cf. Tamanaku *jeje* 'tree', *jeje-'ne* 'trees'; notice that a similar marker occurs in Apalai on inanimate demonstratives: *moro* 'that (medial)', *moro-'ne* 'those (medial)').

Some languages seem to lack collective forms. Kuikuru (Franchetto, pers. comm.) has no 1+2Col pronoun and uses the simple 1+2 *kukuye* in all contexts. Waimirí (Bruno, pers. comm.) has no 1+2Col form, and the 2Col form *amĩrĩĩ* is often replaced by the simple form *amĩ(rĩ)*. Some of the gaps in Table 2 may also indicate actual non-existent forms, and even some of the attested forms may be *ad hoc*, non-lexicalized formations (maybe Panare *jutakon*, *amėnkon*, and Chayma *kutxekon*). One therefore wonders if collective forms should be reconstructed to Proto-Cariban at all.

If one looks only at **komo* and **tomo*, the answer is probably 'no'; but **jamo*, which is apparently older than **komo* (it is always closer to the stem when the two co-occur) is so frequent that it seems at least equally possible that Proto-Cariban **jamo* forms were lost in the languages that lack them (the Pemón group, Panare, and Ikpeng-Arara). Collective **jamo* forms are thus tentatively reconstructed here as **kĩC-jamo* and **am-jamo* (not **kĩCV-jamo* and **amo₂-jamo*, since there are no reflexes of the final vowel in any of the languages, except for Tiriyó and Akuriyó, in which it results from analogy — cf. (i) above — and Wayana, in which the final vowel was protected by the following **ro₂*).

3.2. Third-person forms

Cariban languages usually distinguish *animate* from *inanimate* forms (the only exception being apparently Waimirí; cf. below). For the sake of convenience, these two sets will be examined separately, in Tables 3 (inanimate forms) and 4 (animate forms). Note that only animate pronouns have lexicalized collective forms. A general classification (cf. Derbyshire 1999:54) recognizes *anaphoric* (or *referential*) and *demonstrative* (*proximal*, *medial*, and *distal*) forms; though not all languages fit exactly

into these categories, they are still frequent enough to be useful for comparative purposes.

The *anaphoric* pronoun is not attested in most of the Venezuelan languages (Chayma, Cumanagoto, the Pemón group, Yawarana, Yukpa, Panare). In some cases, this may be due to gaps in the data; however, even the languages with the best sources (e.g. Makushí, Panare) do not mention special anaphoric terms. It is also absent in the Southern languages Arara and Ikpeng, but this is possibly a spurious gap, given the very poor available sources on these languages. If they are not taken into account, the languages without anaphoric pronouns form a geographically contiguous area, and may be more closely related to each other, while those with anaphoric pronouns occupy a larger area and do not seem to form any subgroup within the family. Based on this pattern, an anaphoric term may be reconstructed for Proto-Cariban.

Language	Anaphoric	Proximal	Medial	Distal
Tiriyó ¹⁵	<i>irě</i>	<i>senĩ, serě</i>	<i>měřě</i>	<i>oonĩ, mēnĩ</i>
Akuriyó	<i>irě</i>	<i>txenĩ, txerě</i>	<i>měřě</i>	<i>o'ni</i>
Karihona	<i>irě</i>	<i>enĩ, ẽrě</i>	<i>měřě</i>	<i>mēnĩ</i>
Hixkaryana	<i>ĩro</i>	<i>onĩ</i>	<i>moro</i>	<i>monĩ</i>
Waiwai	<i>ero</i>	<i>on, tan</i>	<i>moro</i>	<i>mĩnĩ</i>
Katxuyana	<i>ijo</i>	<i>soro</i>	<i>moro</i>	<i>monĩ</i>
Karinya-Hf	<i>ĩro</i>	<i>eenĩ, eero</i>	<i>mooro</i>	<i>moonĩ</i>
Karinya-Ms	<i>ijjo</i>	<i>eero</i>	<i>mooro</i>	<i>moonĩ</i>
Apalaĩ	<i>ĩro</i>	<i>senĩ, sero</i>	<i>moro</i>	<i>monĩ</i>
Wayana	<i>irě</i>	<i>herě, sin</i>	<i>měřě</i>	<i>mĩn</i>
De'kwana	<i>ijě(ẽ)</i>	<i>ee'dě</i>	<i>měñě</i>	
Yawarana		<i>(s)eenĩ</i>		<i>měnnĩ</i>
Yukpa ¹⁶				
Waimirĩ	<i>irĩ</i>	<i>(h)anji, kanji</i>	<i>mĩrĩ</i>	<i>mo'o, mĩmo</i>
Tamanaku		<i>txene</i>	<i>more</i>	
Cumanagota		<i>(tx)en</i>	<i>muere</i>	<i>muen</i>
Chayma		<i>(tx)en</i>	<i>muere</i>	<i>muen</i>
Pemón		<i>sene(k), sere</i>	<i>txine(k), muere</i>	
Taurepán		<i>seene(k), silě</i>	<i>měřĩ</i>	<i>mēnĩg</i>
Makushĩ		<i>se(e)ni, sĩrĩrĩ</i>	<i>sĩni, mĩrĩrĩ</i>	
Akawayo		<i>se(e)rě</i>	<i>měřě</i>	
Ingarikó				
Arekuna		<i>seenĩ(g)</i>	<i>měřě</i>	
Panare		<i>sĩ(h)</i>	<i>měñ, ẽmẽ</i>	<i>mu-mĩ'</i>
Ikpeng		<i>nen</i>		<i>mun</i>
Arara				
Bakairĩ	<i>ilě</i>	<i>xĩrě</i>	<i>měřě</i>	<i>(awěřě)</i>
Kuhikuru	<i>ile</i>	<i>iyẽ</i>	<i>eyẽ</i>	

Table 3. Cariban third-person pronouns: inanimate forms.

¹⁵ The terms *senĩ* and *mēnĩ* usually occur in their reduced forms *sen* and *měñ*, except in contexts that preserve the final *ĩ* (a following C(CV)-initial clitic or suffix).

¹⁶ The Yukpa sources contain a wealth of terms, all very poorly analyzed (e.g. Spanish 'ese': *obsek, opse, okano, otka, maa, orko*). Although some of them may be cognate with terms in Tables 3 and 4 (e.g. *mari, maari, mas* 'this'), it seems wiser not to take them into account and wait for better data to become available.

The final vowel of this pronoun was clearly *o*₂, given the *ē* : *o* correspondences (the long *ēē* in De'kwana remains unexplained). The first vowel is somewhat more difficult to determine; **i* looks like the best reconstruction, since an **i* > *ī* change in Apalaí, Hixkaryana and Kuhikuru is more likely ('weakening') than the reverse **ī* > *i*, without any clear conditioning environment (for Apalaí, note that the 1+3 pronoun *īna* also has an *ī* where an *i* or *a* might be expected; for Hixkaryana, consider that **i* > *ī* is elsewhere attested — e.g. the third-person prefix is *ī-*, not *i-*). The length (*īī*) in Karinya-Hf and De'kwana is probably the result of the stress system and should not be reconstructed. The intermediate consonant is usually an *r*, but (a) there are reflexes as *j*, and (b) also as *l* in Kuhikuru, a language in which **r* > *ɣ* (cf. fn. 5). One possible explanation for this pattern would be a different proto-segment (e.g. **rʲ*, or maybe a cluster **rj*). However, two of the languages with *j* reflexes, De'kwana and Karinya, have synchronic morphophonological rules that change *r* into *j* in the vicinity of *i* in at least some cases (cf. e.g. the De'kwana possessive suffix *-rī*, which has an allomorph *-jī* used on stems that end in *i*); the possibility that this might also happen in Katxuyana cannot be excluded. For Kuhikuru, it may be argued that the same **i* (which later became *ī*) was the environment conditioning the *l* reflex instead of *ɣ*. In the absence of more detailed comparisons, it seems better not to postulate a new segment for Proto-Cariban. The anaphoric pronoun is thus reconstructed as **iro*₂.

The *proximal* terms seem to belong either to a **ro*₂ or to a **nī* series, often with both terms co-existing in the same language (e.g. Tiriyo *serē*, *senī*). The fact that many languages do not have both terms raises the question of whether they should be both reconstructed to Proto-Cariban. More work on the actual distribution of those terms, their semantic value,¹⁷ and their diachronic relations to each other is clearly necessary. For the time being, considering that many absences may actually be gaps in the data, that there are some indications of occasional loss of a term (e.g. the Waiwai **ro*₂ term seems to have taken up the anaphoric role, being replaced by the non-cognate *tan* as a proximal), and that languages with one term sometimes have one and sometimes the other (e.g. De'kwana and Katxuyana have **ro*₂ forms, while Yawarana and Hixkaryana have **nī* forms), it seems best to reconstruct two proximals. The reconstruction of their form presents two problems: (1) the fricative initial element *s*, *tx*, *h* present in some languages but not in others; given that even closely related languages may disagree (e.g. Tiriyo and Karihona), it seems best not to reconstruct it;¹⁸ and (2) their initial vowel, which occurs almost always as *e*, but as *a* in Waimirí, *ē* in Karihona, and *o* in Hixkaryana, Waiwai, and Katxuyana (the Parukotoan languages). Waimirí is a very divergent language, so that the *a* might still simply be an idiosyncrasy; but Parukotoan *o* : *e* elsewhere is a correspondence attested

¹⁷ The semantic distinction between the two terms is still unsettled. Hoff (1968:272-273) argues that Karinya *eero* and *mooro* are the proximal and distal terms of a speaker-based subsystem, opposed to the speaker-and-addressee-based subsystem of *eenī* and *monī*. Meira, in a preliminary corpus study (to ap.-b), suggests that the difference is 'newness': *serē* refers to 'new', 'recently introduced' objects, while *senī* refers to previously known objects.

¹⁸ Ikpeng initial *n* is probably not cognate with this element; its origin remains unknown.

also in other words (e.g. Hixkaryana *jo* ‘tooth’, Tiriyo *je*). This correspondence is probably related to *ě* : *o*, here represented as *o*₂, a problem that can only be solved with more comparative work. Here, *e* : *o* is simply represented as *o*₃. The reconstructed forms are thus **o₃ro₂* and **o₃nĩ*.¹⁹

The *medial* and *distal* forms are easier to reconstruct, as **mo₂ro₂* and **mo₂nĩ*, respectively. Further comments: (1) Tiriyo *ooni*, Akuriyo *o’ni*, the actual distal terms (*měni* is used for referents which are hearable but not visible; about the noise made by a non-visible motor, for instance, a Tiriyo speaker might ask: *atĩ mēn?* ‘what’s that?’), have no clear origin. They do not correspond to the other terms in this series (Tiriyo has no *m* : *Ø* correspondences word-initially; there is no source for length in the other words — notice that the stress system in Tiriyo does *not* automatically lengthen the first vowel in CVCV words —; and the final vowel does not correspond to the expected *ĩ*); they must have some other, yet unknown, origin. (2) The same can be said for Makushí *siini*, Pemón *txinek*, which are reminiscent of the proximal terms. (3) Panare *mēn* seems to be the true cognate (with *n* < **ro₂*; cf. the discussion of second-person forms in the previous section); the origin of *ēmē*, and how its meaning differs from the meaning of *mēn*, remain unknown. (4) Kuhikuru *eye* exemplifies initial *m* loss, a normal feature of the language (cf. e.g. Tiriyo *moto* ‘worm’, Kuhikuru *oto*); one wonders whether it has become homophonous with the second-person pronoun or not; they might provide a minimal pair for length (in case Meira’s *ee* is not a mistake). (5) Ikpeng *u* is surprising; it is not known if this is a normal reflex. (For additional details, cf., *mutatis mutandis*, fn. 19).

The *animate anaphoric* pronoun, as was the case with its inanimate counterpart, is mostly not attested in Venezuela (but notice Tamanaku *nare*). Again, since it exists in most other branches, it should be reconstructed to Proto-Cariban. The languages are more or less evenly divided into those with an initial vowel (*i* or *ĩ*), and those without it; it is not clear whether or not it should be reconstructed (note, in passing, that Apalaí again has *ĩ* where other languages have *i*, as was the case for the 1+3 inanimate anaphoric pronouns).²⁰ It can be tentatively added to the final reconstructed form: **(i)no₂ro₂*. (The *a*’s in Bakairí and in Tamanaku are probably mistranscriptions; the glottal stop ‘ in De’kwana remains unexplained.) Note Kuhikuru *l* instead of *r*: the idea that the original **i* ‘palatalizes’ the **r* and keeps it from becoming *ɣ*, though still possible, becomes less plausible, since the **r* is separated from the *i* by one syllable. The possibility that the intermediate consonant should be reconstructed as having a palatal element (**r^j*, or **rj*) cannot be ruled out.

¹⁹ On remaining details: note that the final *rĩ* in Makushí *sĩrĩrĩ* probably stems from the emphatic particle **ro₂*, that the long vowels in Karinya (and probably in Yawarana and Taurepán) are due to the rhythmic stress system, and that the final *g*’s and *k*’s are probably mistranscriptions.

²⁰ There are some indications that an earlier *i*-form may have existed in Tiriyo. The particle *iněērě* ‘that’s the one!’, which follows pronouns (as in e.g. *měrě iněērě* ‘it’s that one!’), looks related to *něrě*. Consider also the occurrence of *něērě*, synchronically equivalent to *něrě* + *rě* (the emphatic particle), but maybe diachronically related to the *i*-initial forms. Akuriyo *něērě* seems to be the same (although it is not known if it has the same *něrě* + *rě* meaning).

Language	Anaphoric		Proximal		Medial		Distal	
	N-Col.	Col.	N-Col.	Col.	N-Col.	Col.	N-Col.	Col.
Tiriyó ²¹	<i>něřě</i>	<i>namo</i>	<i>měe</i>	<i>měesamo</i>	<i>měērě</i>	<i>měějamo</i>	<i>měķī</i>	<i>měķijamo</i>
Akuriyó ²²	<i>něřě</i>	<i>namoro</i>	<i>mě(')e</i>	<i>mětxamo</i>	<i>měķīřě</i>	<i>mě(ě)txamo</i>	<i>ohķī</i>	<i>ohķijan</i>
Karihona	<i>něřě</i>	<i>namoro</i>	<i>měhe</i>				<i>měķē</i>	<i>měkamoro</i>
Hixkaryana	<i>noro</i>	<i>ñamoro</i>	<i>mosonĩ</i>	<i>moxamo</i>	<i>mokro</i>	<i>mokjamo</i>	<i>mokĩ</i>	<i>mokjamo</i>
Waiwai	<i>noro</i>	<i>ñexamro</i>	<i>moso</i>	<i>moxam</i>	<i>mikro</i>	<i>mikjam</i>	<i>mikĩ</i>	<i>mikjam</i>
Katxuyana	<i>noro</i>		<i>mosoro</i>		<i>mokoro</i>		<i>mokĩ</i>	
Karinya-Hf	<i>inooro</i>	<i>inaaro</i>	<i>moose</i>	<i>moojan</i>	<i>mořko</i>	<i>mořkaaro</i>	<i>mookĩ</i>	<i>mořkan</i>
Karinya-Ms	<i>ñooro</i>	<i>ñoorokon</i>	<i>mohse</i>	<i>mohseekon</i>	<i>mohko</i>	<i>mohkaaro</i>	<i>mookĩ</i>	<i>mookĩkon</i>
Apalaĩ	<i>ĩnoro</i>	<i>ĩnaro</i>	<i>mose</i>	<i>moxiamo</i>	<i>mokĩro</i>	<i>mokaro</i>	<i>mokĩ</i>	<i>mokamo</i>
Wayana	<i>iněřě</i>	<i>inamoro</i>	<i>měi, měhe</i>	<i>měham</i>	<i>měkrě</i>		<i>měk</i>	<i>měķjam</i>
De'kwana ²³	<i>ně'dě</i>	<i>nñanno</i>			<i>mě'dě</i>	<i>kanno</i>	<i>mě'kĩ</i>	<i>ma'kamo</i>
Yawarana								
Yukpa ²⁴								
Waimiri ²⁵							<i>mikĩ</i>	
Tamanaku	<i>nare</i>		<i>motxe</i>	<i>txamoro</i>	<i>krere</i>	<i>kiamoro</i>	<i>make</i>	<i>mukiamo</i>
Cumanagot			<i>metxe</i>	<i>metxamo</i>	<i>muekrere</i>		<i>muek</i>	<i>mukiamo</i>
Chayma			<i>metxe</i>	<i>metxam(o)</i>	<i>muekere</i>		<i>muek</i>	<i>mukiam(o)</i>
Pemón			<i>mesere</i>	<i>itxamo(re)</i>	<i>muere</i>			
Taurepán			<i>měserĩ, měsenĩ</i>	<i>měšemonan</i>	<i>maarĩ</i>			
Makushĩ			<i>mĩserĩ</i>	<i>insemoro</i>	<i>mĩķĩrĩ</i>	<i>inkamoro</i>		
Akawayo					<i>ķĩrě</i>			
Ingarikó					<i>měķ(ĩ)re</i>	<i>mĩkamoro</i>		
Arekuna			<i>mĩserě</i>	<i>měĩtxamorĩ</i>	<i>mĩķrěřě</i>			
Panare			<i>mě(ĩ)'</i>	<i>měhtxanton</i>	<i>ķěn</i>	<i>kamonton</i>	<i>muku</i>	<i>mukukon</i>
Ikpeng					<i>oren</i>	<i>wam</i>	<i>ugun</i>	<i>ugjam</i>
Arara								
Bakairĩ	<i>iněra</i>		<i>asaemo</i>		<i>(awěķě)</i>		<i>měķē</i>	<i>akaemo</i>
Kuhikuru	<i>ĩgele</i>		<i>ese</i>		<i>ekise</i>			

Table 4. Cariban third-person pronouns: animate forms.

The *animate proximal* terms all seem to form a good cognate set. The initial consonant is clearly **m* (which is, as expected, lost in Kuhikuru, and maybe also in Bakairĩ, judging by the collective form). The second vowel is **o*₂, and the final vowel **o*₃, given their different correspondences (*ě* : *o* and *e* : *o*). The intermediate consonant is a fricative, probably **tx* (cf. **c* in Girard 1971); notice, however, that Karihona *h* is an unexpected reflex (*h* in this language is supposed to come from **p*; cf. Meira 2000). The Karinya-Ms form suggests the reconstruction of a **hs* (or **htx*)

²¹ The Tiriyó collective forms usually occur as *měesan*, *měėjan*, *měķijan*, *ohķijan* (cf. fn. 15).

²² Meira (2000:60) listed *mě'etxamo*, *mě'jamo*, *měķijamo* as Akuriyó collective pronouns. More recent data (presented here) shows that these forms were mistaken (probably Tiriyó influence).

²³ De'kwana *ně'dě* is described as a distal form; the anaphoric pronoun is *ñwĩ*, a non-cognate.

²⁴ Cf. fn. 16.

²⁵ Bruno (pers. comm.) describes Waimiri as (surprisingly) lacking an animacy distinction. *Irĩ* (cf. Table 3) is also used to refer to people; and *mikĩ* 'that' to inanimate objects.

cluster, which is not a bad hypothesis; however, there is no evidence yet of a Proto-Cariban **h*. The possibility of reconstructing two proximals, suggested by the two Wayana forms *měi* and *měhe* (for which no good semantic description is yet available), seems less likely: no other language has two forms, and, except for the Karihona *měhe* (which cannot be cognate with Wayana *měhe*, since Karihona *h* < **p* and Wayana *h* < **tx*; cf. Girard 1971), all forms look cognate (i.e. there do not seem to be two sets, but only one). For these reasons, the proximal form is here reconstructed as **mo₂txo₃*, and the Karihona *h* is left unexplained. (The two Wayana forms might come from combinations with non-deictic elements, e.g. particles; this is certainly the explanation for the *ro*, *rĩ* in Katxuyana, Taurepán, and Makushí, and possibly also for the *nĩ* in Hixkaryana and Taurepán).

The *animate medial* and *distal* pronouns share suggestive similarities. Looking at cases such as Apalaí *mokĩro* vs. *mokĩ*, Chayma *muekere* vs. *muek*, Waiwai *mĩkro* vs. *mĩkĩ*, etc., one has the impression that the distal terms are simply combinations of the medial term with a reflex of the emphatic particle **ro₂*. This is probably true diachronically, but it even may be true synchronically for some languages.²⁶ For instance, it is not so hard to imagine Apalaí as having a single distal term *mokĩ* that, when co-occurring with the emphatic particle *ro*, is used for closer referents: the ‘closer range’ may be an effect of the semantics of the particle. The two plural forms *mokamo* (for *mokĩ*) and *mokaro* (for *mokĩro*) are also as expected: with the total reduction of the final syllable *mo*, one would expect *mokamo* + *ro* > *mokaro* (though the failure of the vowel *a* to nasalize is unexpected); cf. also Karinya-Hf *moʒkan* and *moʒkaaro*. In Waiwai, there even is only one collective form *mokjam* corresponding to both the medial and the distal pronouns. All of this strongly suggests that Proto-Cariban did not have two non-proximal pronouns, but only one: all forms in the medial and distal columns of Table 4 would then belong to one cognate set. (The only problematic case is Karinya: *moʒko* does not look like *mokĩ* + *ro*. One wonders if there could be a connection with the 1+2 pronoun *kĩʒko*).

The form of this animate distal pronoun presents relatively few problems: **mo₂kĩ* seems to be the best hypothesis. Almost all languages have an initial syllable *mo*, *mě*, *mue* (*e* in Kuhikuru); it is easier to assume that Panare, Tamanaku and Ikpeng lost it. The final syllable *kĩ*, or clear reflexes of it (e.g. De'kwana ', Tiriyo long *ěě*) are also overwhelming. The few problematic cases are: (1) Bakairí *awěkě*, which is not a clear cognate; (2) De'kwana *mě'kĩ* ‘distal’, with an unexpected ' (glottal stop); (3) Ikpeng *oren*, which may not be cognate; *ugun*, with loss of initial *m*, looks like a better candidate. Note that Panare *kěn* comes from **mo₂kĩ-ro₂*, with **ro₂* > *n* (cf. the inanimate medial *měn* and the second-person pronoun *aměn* above).

The *animate collective* forms are also, as was the case with the non-third-person pronouns, formed with reflexes of the collective elements **jamo*, **komo*, and the

²⁶ This is not, of course, true for all languages. In Tiriyo, *měěrě* and *měkĩ* are semantically very distinct; they are clearly two lexical items (cf. Meira to ap-a). Panare *kěn* and *muku* (also attested as *mĩkĩ*) also seem to be clearly independent, at least formally.

particle **ro*₂. The *anaphoric collective* can be reconstructed as resulting from **(i)n* + **jamo* (without the final **ro*₂, since it does not occur in Tiriyó). It is not clear whether the palatalization in Hixkaryana *ñamoro*, Waiwai *ñexamrom* comes from the preceding **i* (in which case one could reconstruct **(i)namo*), or from the following **j* (in which case one could reconstruct **(i)njamo*). To keep both possibilities in mind, the formula **(i)n(j)amo* will be used. The other collectives are again derivable from the non-collective stem plus a combination of collective markers and **ro*₂ (e.g. Panare *kamonton* < **jamo ro*₂ *tomo*). In Taurepán *mēsēmonan*, there seems to be a new collective element *nan* (< **jamo nan*). The problematic cases are: (1) Ikpeng *wam*, which might simply be the element **jamo*, without any original stem, or else non-cognate; (2) Makushí *insemoro*, *inkamoro* with an initial unexplained *i-* (perhaps related to the *(i)-* in the anaphoric forms *(i)no*₂*ro*₂, *(i)n(j)amo*). As was the case for the non-third-person pronouns, collective forms with **jamo* possibly existed; they can be reconstructed as **mo*₂*k-jamo* and **mo*₂*tx-amo*. They are reduced, since almost all reflexes are reduced; Tiriyó *mēkijamo* may have been analogically rebuilt, apparently a frequent phenomenon in Tiriyó collectives — cf. the 1+2Col and 2Col forms. In fact, Tiriyó *mēsamo* also looks like an analogically rebuilt word, given the fact that it conserves an intervocalic reflex of **tx* (cf. Meira 2000:31, 54 for the loss of intervocalic **tx* in Tiriyó). A hypothesis would be: **mo*₂*txamo* > **mētxamo*; at this point **-txamo* is reanalyzed as a suffix, while **mo*₂*txo*₃ > *mēe*; then **mēe-txamo* > *mēsamo*.

As a final observation, it is interesting to note that, apparently, the most complicated Cariban demonstrative systems are found in the Guiana area (from Tiriyó to De'-kwana in the tables). As one moves away from this area, the systems become simpler: there may be no anaphoric term, and often only two distance terms (distal vs. proximal, without medial; e.g. Makushí, Kuhikuru).

4. Conclusion

The Proto-Cariban pronouns reconstructed in the preceding two sections are summarized in Table 5 below.

Pers.	N-Col.	Col.	Categ.	Inanimate	Animate	
					N-Col	Col
1	<i>*awī</i>		Anaph	<i>*iro</i> ₂	<i>*(i)no</i> ₂ <i>ro</i> ₂	<i>*(i)n(j)amo</i>
2	<i>*amo</i> ₂	<i>*am-jamo</i>	Prox-1	<i>*o</i> ₃ <i>ro</i> ₂	<i>*mo</i> ₂ <i>txo</i> ₃	<i>*mo</i> ₂ <i>tx-jamo</i>
1+2	<i>*kīCV</i>	<i>*kīC-jamo</i>	Prox-2	<i>*o</i> ₂ <i>nī</i>		
1+3		<i>*(ap)ina</i>	Med	<i>*mo</i> ₂ <i>ro</i> ₂	<i>*mo</i> ₂ <i>kī</i>	<i>*mo</i> ₂ <i>k-jamo</i>
			Dist	<i>*mo</i> ₂ <i>nī</i>		

Table 5. Proto-Cariban pronominal and demonstrative system.

The non-third-person pronouns form a typical Cariban system, with all categories duly represented. They correspond to the set of person-marking prefixes reconstructed by Gildea (1998:114) as **u-* '1', **a-* '2', and **k-* '1+2'.

The set of third-person pronouns is also typical, despite the absence of a medial-distal distinction (which may be less frequent than the available descriptions suggest). Since most semantic analyses of demonstratives in Cariban languages are not very sophisticated, the meanings of the reconstructed terms are very approximative. In fact, the cognate sets were determined by how well their members fit the known correspondences in the family, rather than by putting together terms with the same gloss; especially for the older sources, glosses such as ‘this’, ‘that’, ‘este’, ‘ese’, ‘aquel’ etc. are not very trustworthy.

The elements in Table 5 display certain recurrent similarities that lead to some speculative ideas. Certain elements can be identified — **ro*₂, **nĩ*, **mo*₂, **kĩ* — which suggest that the third-person pronouns are actually old combinations of yet older pronouns. The anaphoric **iro*₂ could be a combination of a third-person marker **i-* (from Gildea’s **jĩ-*) with the element *ro*₂, which could be the emphatic particle — i.e. ‘really third-person’. (This presupposes that the third-person prefix would have been an independent element in the past, so that it could be followed by the particle **ro*₂). The element **mo*₂, also found in combination with **ro*₂, might be compared to the ‘evidential’ *mo* or *mẽ* that, in some languages, occurs with the third-person prefix to indicate certain evidential values (e.g. Wayana *nĩ-tẽ-jai* ‘he is going’, *mẽ-n-tẽ-jai* ‘he is going (but I do not see him)’; Hixkaryana *mo-n-eweh-no* ‘he took a bath (out of sight)’). More comparative research should help decide how much truth there is in such speculations.

References

Abbott, Miriam

1991 ‘Macushi’, in Desmond C. Derbyshire & Geoffrey K. Pullum (eds.), *Handbook of Amazonian languages*, vol. 3, Berlin: Mouton de Gruyter, pp. 23-160.

Amodio, Emanuele, & Vicente Pira

1996 *Língua Makuxi — Makusi Maimu. Guia para a aprendizagem e dicionário da língua makuxi*, Boa Vista (Brazil): Diocese de Roraima & Missionários de Scarboro (Canada).

Armellada, Cesáreo de, & Jesús Olza

1994 *Gramática de la lengua pemón (morfosintaxis)*, San Cristóbal (Venezuela): Universidad Católica del Táchira.

Bruno, Ana Carla dos Santos

1996 *Guia de aprendizagem da língua Waimirí-atroari*, Manaus: Programa Waimirí-Atroari, Convênio FUNAI / Eletronorte.

Derbyshire, Desmond C.

1979 *Hixkaryana*, *Lingua Descriptive Studies*, vol. 1. Amsterdam: North-Holland.

1985 *Hixkaryana and Linguistic Typology*, Dallas: Summer Institute of Linguistics and the University of Texas at Arlington.

- 1999 'Carib', in R. M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *The Amazonian languages*, Cambridge Language Surveys, Cambridge: Cambridge University Press, pp. 23-64.
- Edwards, Walter F.
- 1977 *An introduction to the Akawayo and Arekuna peoples*, Amerindian Languages Project, Georgetown: University of Guyana.
- Gildea, Spike
- 1995 'A comparative description of syllable reduction in the Cariban language family', *International Journal of American Linguistics* 61:62-102.
- 1998 *On reconstructing grammar: comparative Cariban morphosyntax*. Oxford Studies in Anthropological Linguistics, vol. 18. Oxford: Oxford University Press.
- Gilij, Felipe [Filippo] S.
- 1665 *Ensayo de historia americana*, 3 vols, transl. by Antonio Tovar, Caracas: Academia Nacional de la Historia (original 1782, *Saggio di Storia Americana*, 4 vols., Rome).
- Girard, Victor James
- 1971 *Proto-Carib phonology*, Berkeley (CA): University of California Ph.D. dissertation.
- Hall, Katherine L.
- 1988 *The morphosyntax of discourse in De'kwana Carib*, Saint Louis: Washington University Ph.D. dissertation.
- Hawkins, Robert E.
- 1998 'Waiwai', in Desmond C. Derbyshire & Geoffrey K. Pullum (eds.), *Handbook of Amazonian languages*, vol. 4, Berlin: Mouton de Gruyter, pp. 25-224.
- Hoff, Berend J.
- 1968 *The Carib language*, The Hague: Martinus Nijhoff.
- 1990 'The non-modal particles of the Carib language of Surinam and their influence on constituent order', in Doris L. Payne (ed.), *Studies in Lowland South American Languages*, Austin: University of Texas Press, pp. 495-541.
- Jackson, Walter S.
- 1972 'A Wayana grammar', in Joseph E. Grimes (ed.), *Languages of the Guianas*, Norman: Summer Institute of Linguistics and University of Oklahoma Press, pp. 47-77.
- Kaufman, Terrence K.
- 1994 'The native languages of South America', in Christopher Moseley and R. E. Asher (eds.), *Atlas of the world's languages*, London, New York: Routledge, pp. 46-76.
- Koch-Grünberg, Theodor
- 1916 *Vom Roroima zum Orinoco. Ergebnisse einer Reise in Nordbrasilien und Venezuela in den Jahren 1911-1913*, vol. 4: *Sprachen*, Stuttgart: Strecker u. Schröder.

- Koehn, E., & S. Koehn.
 1986 'Apalai', in Desmond C. Derbyshire & Geoffrey K. Pullum (eds.), *Handbook of Amazonian languages*, vol. 1, Berlin: Mouton de Gruyter, pp. 33-127.
- Meira, Sérgio
 1998 'Rhythmic stress in Tiriyo (Cariban)', *International Journal of American Linguistics* 64:352-378.
 1999 *A grammar of Tiriyo*, Houston (USA): Rice University Ph.D. dissertation (to appear in the Mouton grammar series).
 2000 *A reconstruction of Proto-Taranoan: phonology and morphology*, München: LINCOM Europa.
- to ap-a 'Non-contrastive exophoric use of Tiriyo demonstratives', in Michael Dunn & Sérgio Meira (eds.), *Demonstratives in Cross-Linguistic Perspective*.
 to ap-b 'Les démonstratifs proximaux non-animés de la langue tiriyo (caribe): une étude de corpus', *Amérindia*.
- Méndez-Arocha, Alberto
 1959 'Un vocabulario yabarana con apuntes fonémicos', *Antropológica* (Venezuela) no. 7:65-84.
- Mosonyi, Jorge C.
 1978 *Diccionario básico del idioma caríña*, Thesis, Caracas: Universidad Central de Venezuela.
- Muller, Marie-Claude Mattei
 1994 *Diccionario ilustrado panare-español, español-panare: un aporte al estudio de los Panares-E'ñepa*, Caracas: Comisión Quinto Centenario, Gráficas Armitano.
- Pacheco, Frantomé Bezerra
 1997 *Aspectos da gramática ikpeng (karib)*. M.A. Thesis, Campinas (Brazil): Universidade de Campinas (UNICAMP), Instituto de Estudos da Linguagem.
- Robayo, Camilo
 1987 *Le système des personnes de la langue carijona*, DEA Thesis, Paris: Université de Paris VII - Jussieu.
 2000a 'Avance sobre morfología carijona', in María Pérez & María Montes (eds.), *Lenguas indígenas de Colombia: una visión descriptiva*, Bogotá: Instituto Caro y Cuervo, pp. 171-180.
 2000b 'Introducción al estudio de la lengua yupo o yukpa', in María Pérez & María Montes (eds.), *Lenguas indígenas de Colombia: una visión descriptiva*, Bogotá: Instituto Caro y Cuervo, pp. 709-717.
- Ruiz Blanco, Matias
 1690 *Arte y tesoro de la lengua cumanagota*, Madrid (facsimilar edition by Julio Platzmann, Leipzig: B. G. Teubner, 1888).
- Souza, Isaac de
 1983 *Arara - Karib*. Relatório para a FUNAI. (Includes unpublished material — hand-outs — from presentations at two conferences, dated 1992).

Steinen, Karl von den.

- 1892 *Die Bakairi-Sprache: Wörterverzeichnis, Sätze, Sagen, Grammatik (mit Beiträgen zu einer Lautlehre der Karaibischen Grundsprache)*, Leipzig: K. F. Koehler's Antiquarium.

Tauste, Francisco de

- 1680 *Arte y vocabulario de la lengua de los indios chaymas, cumanagotos, cores, parias, y otros diversos de la provincia de Cumana, o Nueva Andalucia*, Madrid (facsimilar edition by Julio Platzmann, Leipzig: B. G. Teubner, 1888).

Vegamián, Felix M.

- 1978 *Diccionario ilustrado yupa-español, español-yupa (con onomástica y apuntes gramaticales)*, Caracas: Formateca C.A.

Wheatley, James

- 1973 'Pronouns and nominal elements in Bacairi discourse', *Linguistics* 104:105-115.

- 1978 *Vocabulário Bakairi-Português, Português-Bakairi (edição experimental)*, Brasília: Summer Institute of Linguistics.

Yangués, Manuel de

- 1683 *Principios y reglas de la lengua cummanagota*, Burgos (facsimilar edition by Julio Platzmann, Leipzig: B. G. Teubner, 1888).