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 Salvadore 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.

<table>
<thead>
<tr>
<th>Languages</th>
<th>Sources</th>
<th>Languages</th>
<th>Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tiriyó</td>
<td>Meira 1999, 2000; fn ++</td>
<td>Tamanaku</td>
<td>Gilij 1965 [1782]</td>
</tr>
<tr>
<td>Akuriyó</td>
<td>fn ++</td>
<td>Cumanagoto</td>
<td>Yangues 1683, Ruiz Blanco 1690</td>
</tr>
<tr>
<td>Karihona</td>
<td>Robayo 1987, 2000a; fn ++</td>
<td>Chayma</td>
<td>Tauste 1680</td>
</tr>
<tr>
<td>Hixkaryana</td>
<td>Derbyshire 1979, 1985 ++</td>
<td>Pemón</td>
<td>Armellada &amp; Olza 1994</td>
</tr>
<tr>
<td>Waiwai</td>
<td>Hawkins 1998; fn ++</td>
<td>Taurepán</td>
<td>Koch-Grünberg 1916</td>
</tr>
<tr>
<td>Katchuyana</td>
<td>fn; Gildea’s fn +</td>
<td>Makushi</td>
<td>Abbott 1991, Amodio &amp; Pira 1996 ++</td>
</tr>
<tr>
<td>Karinya</td>
<td>Hoff 1968, Mosonyi 1978 ++</td>
<td>Ingarikó</td>
<td>Koch-Grünberg 1916</td>
</tr>
<tr>
<td>Apalaí</td>
<td>Koehn &amp; Koehn 1986; fn ++</td>
<td>Arekuna</td>
<td>Edwards 1977; Koch-Grünberg 1916</td>
</tr>
<tr>
<td>Wayana</td>
<td>Jackson 1972; fn ++</td>
<td>Akawayo</td>
<td>Edwards 1977; fn +</td>
</tr>
<tr>
<td>De’kwana</td>
<td>Hall 1988 +</td>
<td>Panare</td>
<td>Muller 1994 ++</td>
</tr>
<tr>
<td>Yawarana</td>
<td>Méndez-Arocha 1959 +</td>
<td>Ikpeng</td>
<td>Pacheco 1997; Gildea’s fn +</td>
</tr>
<tr>
<td>Atroari</td>
<td></td>
<td>Kuhikuru</td>
<td>Franchetto’s fn ++</td>
</tr>
</tbody>
</table>

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.

The different transcription systems of the various sources have been standardized to facilitate comparisons. Certain details have been ignored (e.g. Koch-Grünberg’s
attempt at marking non-phonemic distinctions between \[e\], \[e\] and \[o\], \[ɔ\], here transcribed simply as \(e, o\); his stress marks — \(â, é\), etc. — were also left out). The symbols in need of explanation are: \(\tilde{i} = \text{IPA} \{i\}, \tilde{e} = \{ə\}, \tilde{j} = \{j\}\) (a palatal glide, English \(y\)), \(\chi = \{x\}, \gamma = \{ɣ\}\) (velar fricatives), \(\dot{\imath} = \{ʔ\}, \ddot{n} = \{n\}, x = \{ʃ\}, tx = \{tʃ\}\). Sequences of identical vowels \((aa, ee, \text{etc.})\) 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. Makushi, 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), 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 \(rê\), 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 \(wê\) found in several languages cannot be compared to \(ro \sim rê\),

---

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 \( rë \), synchronically attested in most Cariban languages (e.g. Tiriyó \( wë'I', \( wëë '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 \( rë \) result from the reanalysis of the emphatic particle as part of the stem (much like \( otros \) in Spanish

\[
\begin{array}{|c|c|c|c|c|c|}
\hline
\text{Language} & \text{1} & \text{2} & \text{1+2} & \text{1+3} & \text{2Col all of you} \\
\hline
\text{Tiriyó} & wë & émé & kînë & anja & kîmënjamo \\
\text{Akuriyó} & wi & émé & kînë & anja & kîmënjamo \\
\text{Karhona} & ëwë & êmëërë & kîmëërë & ëça & kihamoro \\
\text{Hixkaryana} & uro & omoro & kîvero & amna & kîwëjamo \\
\text{Waiwai} & ëwë & amoro & kîwëri & amna & kîwejam \\
\text{Katsuyana} & ëwë & omoro & kîmoro & amna & kîmëjari \\
\text{Karinya-Hf} & au & amoro & kîyëko & an'ëna & kîyëkaoro \\
\text{Karinya-Ms} & aau & amoro & kîm"oro & na'ëna & kîmënhëoro \\
\text{Apalai} & ëwë & omoro & kîmoro & ina & kîmaroko \\
\text{Wayana} & ëwë & émé & kunnë & enna & kunmëramkom \\
\text{De'kwa} & ëwë & êmëdë & kivi & ënëa & kînëwëno \\
\text{Yavarana} & ëwë & mëërë & kënë & ënëa & kînëwëno \\
\text{Yukpa} & ëwë, ëwë & amo, amor & nana & amora, amoru \\
\text{Waimiri} & ëwë, au & amirë & kikë & anëa & (amirië) \\
\text{Tamanaku} & ëure & amare & kiwe & jumna & kikëmo \\
\text{Cumanagoto} & ëure & amure & kute & amna & amikërokon \\
\text{Chayma} & ëure, uxe & amure & kute & amna & kutëkon \\
\text{Pemón} & ëjure & amare & jjurete & intëna & jurenkon \\
\text{Taurepán} & ëjurë & amarë & ëjurë'kon & ënë & jureñikinok \\
\text{Makushi} & ëjurë & amirë & ëjurë'kon & ënë & jureñikinok \\
\text{Akawayo} & ërë & amërë & ërë'non & ënë & amërë'non \\
\text{Ingariñiño} & ëure & amëëte & ëurenkëno & ënë & ëurenkëno \\
\text{Arekuna} & ëjurë & amëëte & ëurenkëno & ënë & ëurenkëno \\
\text{Panare} & ëjurë & amëëte & ëurenkëno & ënë & ëurenkëno \\
\text{Ikpeng} & ëjurë & amëëte & ëurenkëno & ënë & ëurenkëno \\
\text{Ara} & ëjurë & amëëte & ëurenkëno & ënë & ëurenkëno \\
\text{Bakairi} & ëjurë & amëëte & ëurenkëno & ënë & ëurenkëno \\
\text{Kuhikuru} & ëjurë & amëëte & ëurenkëno & ënë & ëurenkëno \\
\hline
\end{array}
\]

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-jëmë 'my father'), which has replaced an earlier Proto-Cariban *\( ur- \) (cf. Gildea 1998). Note also that the final syllable \( ro, rë \) occurs as \( ri \) in Makushi and Arekuna (cf. below for Makushi \( i > \), and as \( re \) in Kuhikuru (for which \( y : r \) and \( e : r \) are also regular correspondences: cf. Kuhikuru \( auy \) 'manioc bread', \( tehù 'stone', Tiriyó \( uuru, ëtëpo \)). The cases of \( re \) (Pemón, Ingariñiño, Tamanaku, Cumanagoto, Chayma) are certainly mistranscriptions of \( ri \).

5 Hoff (1968) and Mosonyi (1978) describe mutually intelligible dialects of the same language (‘Carib’ for Hoff, ‘Carita’ 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 rë, 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 (*Vwë > Vu > u). The first vowel, however, is a more difficult case: there are forms beginning with a, o, è 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) è and o seem to be diachronically related (cf. the second-person pronouns in Table 2), so that the è- 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 *awï.

Three problematic details remain, for which some suggestions are presented here. (1) Tiriyó ìi is probably the result of a metrical reanalysis of pronoun–clitic sequences: e.g. Pre-Tiriyó *ëwï rë ‘really me’ would go from [ɔi:ʔa] to [iti:γa] by losing the initial vowel, at which point the surface long vowel would be reanalyzed as underlying ìi (or else it would become short — [itγa] —, as in all CVCV words; cf. Meira 1998, 1999 on the stress system), thus yielding wï rë. (2) The initial j in Panare, Pemón, Makushi, and Taurepán may result from the resyllabification of an earlier *Vw > *ïw sequence (e.g. *ïwï rë > *ïu rë > juuri).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 ~ rë: ro, rë, rë, ri, re, ye, and also Waimiri ra and De’kwana dë (De’kwana d often corresponds to r in other languages: cf. jîwîdi ‘tapir’, Tiriyó ìwîrî). Panare n is also a likely reflex of an earlier rë; cf. Panarc ūnkë ‘horsefly’, akuñ ‘agouti’, Tiriyó turëkë, 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 mëérë, on the other hand, may represent a case of underlying ëë resulting from the loss of the initial vowel, like Tiriyó wï (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 jîu looks like an attempt at transcribing what could have been an intermediate stage (something like e.g. ìiç). 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, ê. As was mentioned above, ê and o may not be independent, which would reduce the choice to a vs. ê/o. Considering that the second vowel was probably ê/o, the ê/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 ê/o: the cases of ue (Cumanagoto, Chayma) and a (Tamanaku, Taurepán) are probably mistranscriptions of ê, and the cases of ë (Waimiri, Makushi, Arekuna) look like reflexes of ê (ê : ë is attested in Makushi, as in e.g. sikê ‘flea, chigger’, Tiriýó sikê; Waimiri has no phonemic ê). As was mentioned above, ê and o are clearly related; there are numerous examples of the ê : o correspondence (e.g. Tiriýó sikê ‘flea, chigger’, Apalai xiko). Gildea (pers. comm.) considers ê 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. Tiriýó okomo ‘wasp’, Apalai okomo; cf. the second-person forms in Table 2). The question of whether o : o and ê : o are independent correspondences has not yet been settled. Taking a conservative stance, *o will be reconstructed for o : o, and *a for ê : o. One thus ends up with a tentative proto-form *amo2. The last problem is the unexpected Kuhikuru form e(e)-γe (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-γe > e(e)-γe. However, all attested cases of m-loss in Kuhikuru are word-initial, not word-internal (e.g. Kuhikuru oto ‘worm’, Tiriýó moto). It seems thus better to suppose that the initial *e was lost first: *eme-γe > *me(e)-γe > e(e)-γe. (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 Tiriýó wî.)

The first-person dual inclusive (1+2) forms show more complex patterns. After eliminating the reflexes of the particle ro ~ rê, there are two major groups: (a) forms that contain the initial element ku, kî, ki, and (b) forms that contain an initial element ju, u (Panare, Pemon, Makushi; 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 ju-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. Tiriýó ton, kon, Apalai tomo, kome, 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

Note that the o : o and ê : o correspondences have distinct reflexes in Kuhikuru: e.g. tehu ‘stone’, Tiriýó tehu, and otî ‘field, grass’, Tiriýó otî; cf. also Kuhikuru okô ‘wasp’. Thus, Pre-Kuhikuru apparently had *o and *ê. Considering the number of (not obviously closely related) languages that have *ê, it is not impossible that Proto-Cariban *a2 was actually *ê. Not much, however, can be said without a detailed study of the distribution of o : o and ê : 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 Ingariñó 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 kī, the first idea is to reconstruct *kī. The cases of ku, however, give food for thought. First, ku occurs in Wayana and in the Southern languages (Kuhikuru, Bakairí, 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 ‘weaken to ũ’, and also the fact that k is not an obvious environment for labialization (*ũ > u). On the other hand, the possibility of deriving ku from an earlier *kiwi, at least for some languages (cf. below), must be borne in mind. All in all, reconstructing *kī 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 më (mo, m*o, nmë), wī (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(ĩ)-) 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: kĩko [Hf], kĩm*ooro [Ms]). The various forms can be separated in several groups, which correspond only imperfectly to proposed subgroupings (e.g. in Kaufman 1994): the më group (Tiriyó, Akuriyó, Karhona [= Meira’s Taranoan], Wayana, Apalai, Katsu-

8 Notice that m*o instead of mo in kĩm*ooro 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 mo results from the rhythmic stress system. The n in Wayana kunmë is less readily explained; it is tentatively reconstructed, despite the rather strange absence of its reflexes in the other më languages.

9 The long ũ in Waiwai kĩũ 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 *kiwi* is not implausible as a source for the *ku* forms (e.g. Wayana kunmë < *kuu nmo₂ < *kiwi nmo₂; for the *ku* group, one might have e.g. *kiwi ro₂ > *kuu ro₂ > Bakairí kurë). However, this leaves the ‘double-*ku*’ forms (Kuhikuru kuku, Arara ugo, Ikpeng ugu(u)) unexplained, and also Karinya kiiko; the lack of any reflex of the syllable wi 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 kurë < *ku-ro₂), 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 + ro₂ 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 *kiCV 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 ehnë possibly explained by weakening, and Waimiri 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-ñn, i-nn, i-ñ, ñ: Tiriyó, Akuriyó, Karihona [Meira’s Taranano], De’kwana, Yukpa,10 Pemón, Taurepán, Makushi, 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-Araa tsí-m; tentative reconstruction, *a-m-na); and an ‘or glottal group (n)a’-; e-: Karinya, Waimiri, 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 ñ < *i (weakening-to-ñ); 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, ña (the last form with 'weakening-to-T'); or (b) reduce to a glottal segment, yielding Karinya-Hf a'na with loss of i (and also n for Waimiri a'ua) and Yawarana ehné without loss (via e.g. *ai-p-na > *e-x-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, Ikapeng 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, ñ 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, -jarî, -jarō, -aro, -wanno, -komo, -kemo, -ton, -nî, -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, rê (< *ro2) 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 kîmê-njamo, êmê-njamo < *jamo
Akuriyó kî(mê)-njamo, êmê-njamo < *jamo
Karihona kî-ñamoro, a-ñamoro < *jamo ro2
Hixkaryana kîw-jamo, om-ñamo < *jamo
Waiwai kîw-jamo, om-ñam-ro < *jamo (ro2)
Katxuyana kîm-jari, om-jari < *jamo ro2 (?)
Karinya-Hf13 kî'k-aaro, amii-jaro < *jamo ro2

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 < ñu (as in amiijaro, note the short a here). Note that Karinya reduces nasal syllables to zero, even in synchronic morphophonology, so that *jamo ro2 > (j)aro is not surprising (cf. awoomi ‘to get up’, aj-aawo-ja ‘I am getting up’).
Karinya-Ms kîmʷ-o-ññaro, amo-ññaro < *jamo ro2
Apaláí kîm-aroñkom, am-aroñkom < *jamo ro2 komo
Wayana kunmë-ramkom, êmë-ramkom < *ro2 jamo komo
De`kwana kî-nwanno, ê-nwanno < *jamo ro2
Yukpamo-ña, amo-ja < *(ro₂) jamo
Waimiri amë-riñ < *ro2 tomo
Tamanaku ki-ñëmë, am-ñamoro < *komo; *jamo ro2
Cumanagoto am-i(a)morkom < *jamo ro2 komo
Chayma kute-kom, am-iamorkom < *komo; *jamo ro2 komo
Pemón (j)ure-nokon, amare-nokon < *no komo
Taurepán juùrì-nikom, amarë-nikom < *no komo
Makushi uarë-`nikom, amirë-`nikom < *no komo
Akawayo urë-`nogon, amërë-`nogon < *no komo
Ingariko14 kiule-nikom, ãmë-liñikom < *no komo
Arekuna jureñkom, amërë(k)-ñkom < *tomo komo; *no komo
Panare juta-kom, amën-ton < *komo; *tomo
Ikpeng ugro-çmo, om-rocño < *(ro₂) komo
Arara ugoro-çmo < *çmo
Bakairí(a)ma-reemo < *ro2 jamo
Kuhikuru amëy < *jamo ro2

Some suggestions for the problematic details are listed below.

(i) For the suffix -njamo in Tiriyó and Akuriyó, 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: *kîmë + jamo > *kîn-jamo > *kî-njamo, *kîmë + -njamo > kîmënjamo. 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. awoomi ‘get up’, aj-aawoi-ja ‘I am getting up’). The long aa in the 1+2Col form probably results from syllable fusion (*kîgko-jaro > kîgkoarò; cf. 2Col amëjaro, with a short a); for the Karinya-Ms forms, however, no obvious explanation was found. NV loss also occurs synchronically in Apaláí and Kuhikuru. Katxuyana -jarë is surprising, both because there usually is no NV loss in this language, and also because *ro2 should occur as ro, not ri. Nevertheless, an irregular evolution of *jamo ro2 still seems

---

14 The initial k, t in Ingariko (cf. also, from Koch-Grünberg 1916, Arekuna kuulïnikon ‘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 *ri or *jari without additional evidence.

(iii) The nw in De’kwana nwanno is hard to explain (it is not related to Karinya-Ms m”; cf. fn. 7). There are some correspondences between j and w in the Cariban family (e.g. Tiriyo ê-jomi ‘your language’, Wayana ê-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 *ni, 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 kukuγ in all contexts. Waimiri (Bruno, pers. comm.) has no 1+2Col form, and the 2Col form amïrïtï is often replaced by the simple form amï(ri). 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, amënkon, and Chayma kutsekon). 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 *kiC-jamo and *am-jamo (not *kiCV-jamo and *amo2-jamo, since there are no reflexes of the final vowel in any of the languages, except for Tiriyo 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 Waimiri; 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. Makushi, 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.

<table>
<thead>
<tr>
<th>Language</th>
<th>Anaphoric</th>
<th>Proximal</th>
<th>Medial</th>
<th>Distal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tiriyó15</td>
<td>iřë</td>
<td>senï, serë</td>
<td>mërë</td>
<td>ooni, mënï</td>
</tr>
<tr>
<td>Akuriyó</td>
<td>iřë</td>
<td>tserï, tsërë</td>
<td>mërë</td>
<td>o'ni, mënï</td>
</tr>
<tr>
<td>Karhona</td>
<td>iřë</td>
<td>enï, īřë</td>
<td>mërë</td>
<td>mënï</td>
</tr>
<tr>
<td>Hsxkaryana</td>
<td>iřë</td>
<td>onï</td>
<td>moro</td>
<td>mënï</td>
</tr>
<tr>
<td>Wawat</td>
<td>ero</td>
<td>on, tan</td>
<td>moro</td>
<td>mënï</td>
</tr>
<tr>
<td>Katxuyana</td>
<td>ijo</td>
<td>soro</td>
<td>moro</td>
<td>mënï</td>
</tr>
<tr>
<td>Karinya-H</td>
<td>iro</td>
<td>enë, eero</td>
<td>mooro</td>
<td>mënï</td>
</tr>
<tr>
<td>Karinya-Ms</td>
<td>ijo</td>
<td>eero</td>
<td>mooro</td>
<td>mënï</td>
</tr>
<tr>
<td>Apalai</td>
<td>iro</td>
<td>senï, soro</td>
<td>moro</td>
<td>mënï</td>
</tr>
<tr>
<td>Wayana</td>
<td>iřë</td>
<td>herë, sin</td>
<td>mërë</td>
<td>mënï</td>
</tr>
<tr>
<td>De’kwana</td>
<td>ijë(ë)</td>
<td>ee’dë</td>
<td>mënë</td>
<td>mënë</td>
</tr>
<tr>
<td>Yawarana</td>
<td>iijë</td>
<td>tije’ni</td>
<td>mënë</td>
<td>mënni</td>
</tr>
<tr>
<td>Yukpâ</td>
<td>iro</td>
<td>sëenï</td>
<td>mënï</td>
<td>mënï</td>
</tr>
<tr>
<td>Waimiri</td>
<td>iřë</td>
<td>(b)anjï, kanjï</td>
<td>mëri</td>
<td>mo’i, mîmo</td>
</tr>
<tr>
<td>Tamaranakua</td>
<td>tsëni</td>
<td>tserë</td>
<td>morë</td>
<td>mo’re</td>
</tr>
<tr>
<td>Cumanagota</td>
<td>(ts)ëni</td>
<td>(ts)ën</td>
<td>muren</td>
<td>muen</td>
</tr>
<tr>
<td>Chayma</td>
<td>senë(k), sërë</td>
<td>tserë, muere</td>
<td>mënë</td>
<td>mënë</td>
</tr>
<tr>
<td>Pemón</td>
<td>seenë(k), sërë</td>
<td>tserë, muere</td>
<td>mënë</td>
<td>mënë</td>
</tr>
<tr>
<td>Taurepán</td>
<td>se(e)në, sëriřë</td>
<td>sëni, mëriřë</td>
<td>mënë</td>
<td>mënë</td>
</tr>
<tr>
<td>Makushi</td>
<td>se(e)ni, sëriřë</td>
<td>sëni, mëriřë</td>
<td>mënë</td>
<td>mënë</td>
</tr>
<tr>
<td>Akawayo</td>
<td>se(e)Irë</td>
<td>sëiřë</td>
<td>mërë</td>
<td>mënë</td>
</tr>
<tr>
<td>Ingarikó</td>
<td>seenë(g)</td>
<td>mërë</td>
<td>mënë</td>
<td>mënë</td>
</tr>
<tr>
<td>Arekuna</td>
<td>seenë(g)</td>
<td>mërë</td>
<td>mënë</td>
<td>mënë</td>
</tr>
<tr>
<td>Panare</td>
<td>sëkë</td>
<td>mënë, ēmë</td>
<td>mëu–më’</td>
<td>mënë</td>
</tr>
<tr>
<td>Ikpeng</td>
<td>nen</td>
<td>mënë</td>
<td>mënë</td>
<td>mënë</td>
</tr>
<tr>
<td>Arara</td>
<td>nèn</td>
<td>mënë</td>
<td>mënë</td>
<td>mënë</td>
</tr>
<tr>
<td>Bakairi</td>
<td>ile</td>
<td>xërë</td>
<td>mërë</td>
<td>mënë</td>
</tr>
<tr>
<td>Kahlakuru</td>
<td>ile</td>
<td>yë</td>
<td>mërë</td>
<td>mënë</td>
</tr>
</tbody>
</table>

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

15 The terms senï and mënï usually occur in their reduced forms sen and mën, except in contexts that preserve the final I (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, ma ‘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 \( \ddot{e} : o \) correspondences (the long \( \dddot{e} \) in De’kwana remains unexplained). The first vowel is somewhat more difficult to determine; \( *i \) looks like the best reconstruction, since an \( *i > i \) change in Apalai, Hixkaryana and Kuhikuru is more likely (‘weakening’) than the reverse \( *i > \ddot{i} \), without any clear conditioning environment (for Apalai, note that the 1+3 pronoun \( \text{ina} \) also has an \( \ddot{i} \) where an \( i \) or \( a \) might be expected; for Hixkaryana, consider that \( *i > \ddot{i} \) is elsewhere attested — e.g. the third-person prefix is \( \ddot{i}- \), not \( i- \)). The length (\( i \)) 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 \( *r > \gamma \) (cf. fn. 5). One possible explanation for this pattern would be a different proto-segment (e.g. \( *r' \), 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 -\( r\ddot{i} \), which has an allomorph -\( ji \) used on stems that end in \( i \)); the possibility that this might also happen in Katsuyana cannot be excluded. For Kuhikuru, it may be argued that the same \( *i \) (which later became \( l \)) was the environment conditioning the \( l \) reflex instead of \( \gamma \). 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 \( *ni \) series, often with both terms co-existing in the same language (e.g. Tiriyó \( \text{ser}\ddot{e}, \text{sen}_i \)). 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 \( \text{tan} \) as a proximal), and that languages with one term sometimes have one and sometimes the other (e.g. De’kwana and Katsuyana have \( *ro_2 \) forms, while Yavarana and Hixkaryana have \( *ni \) forms), it seems best to reconstruct two proximals. The reconstruction of their form presents two problems: (1) the fricative initial element \( s, t\ddot{c}, h \) present in some languages but not in others; given that even closely related languages may disagree (e.g. Tiriyó 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 Waimiri, \( \ddot{e} \) in Karihona, and \( o \) in Hixkaryana, Waiwai, and Katsuyana (the Parukitoan languages). Waimiri is a very divergent language, so that the \( a \) might still simply be an idiosyncrasy; but Parukitoan \( o : e \) elsewhere is a correspondence attested

\(^{17}\) The semantic distinction between the two terms is still unsettled. Hoff (1968:272-273) argues that Karinya \( \text{ero} \) and \( \text{mooro} \) are the proximal and distal terms of a speaker-based subsystem, opposed to the speaker-and-addressee-based subsystem of \( \text{sen}_i \) and \( \text{mon}_i \). Meira, in a preliminary corpus study (to ap.-b), suggests that the difference is ‘newness’: \( \text{ser} \ddot{e} \) refers to ‘new’, ‘recently introduced’ objects, while \( \text{sen}_i \) 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’, Tiriyó je). This correspondence is probably related to ê : o, here represented as o₂, a problem that can only be solved with more comparative work. Here, e : o is simply represented as o₃. The reconstructed forms are thus *o₃ro₂ and *o₃nï.¹⁹

The medial and distal forms are easier to reconstruct, as *mo₃ro₂ and *mo₃nï, respectively. Further comments: (1) Tiriyó ooni, Akuriyó o’ni, the actual distal terms (mënï is used for referents which are hearable but not visible; about the noise made by a non-visible motor, for instance, a Tiriyó speaker might ask: atë mën? ‘what’s that?’), have no clear origin. They do not correspond to the other terms in this series (Tiriyó has no m : Ø correspondences word-initially; there is no source for length in the other words — notice that the stress system in Tiriyó does not automatically lengthen the first vowel in CVCV words —; and the final vowel does not correspond to the expected i); they must have some other, yet unknown, origin. (2) The same can be said for Makushi sîni, Pemón txinek, which are reminiscent of the proximal terms. (3) Panarê mën seems to be the true cognate (with n < *ro₂; cf. the discussion of second-person forms in the previous section); the origin of êmë, and how its meaning differs from the meaning of mën, remain unknown. (4) Kuhikuru eye exemplifies initial m loss, a normal feature of the language (cf. e.g. Tiriyó 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 ï), and those without it; it is not clear whether or not it should be reconstructed (note, in passing, that Apalaí again has ï 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 γ, 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 (rí or *rf) cannot be ruled out.

¹⁹ On remaining details: note that the final r in Makushi sîrî probably stems from the emphatic particle *ro₂, that the long vowels in Karinya (and probably in Yawarana and Taurepin) are due to the rhythmic stress system, and that the final ģ’s and k’s are probably mistranscriptions.

²⁰ There are some indications that an earlier i-form may have existed in Tiriyó. The particle ñëërëe ‘that’s the one!’ , which follows pronouns (as in e.g. mëëë inëërëe ‘it’s that one!’), looks related to ñëëë. Consider also the occurrence of ñëëë, synchronically equivalent to ñëëë + rê (the emphatic particle), but maybe diachronically related to the i-initial forms. Akuriyó ñëëë seems to be the same (although it is not known if it has the same ñëëë + rê meaning).
### Table 4. Cariban third-person pronouns: animate forms.

<table>
<thead>
<tr>
<th>Language</th>
<th>Anaphoric</th>
<th>Proximal</th>
<th>Medial</th>
<th>Distal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tiriyó</td>
<td>nërë</td>
<td>namo</td>
<td>mëë</td>
<td>mëxamo</td>
</tr>
<tr>
<td>Akuriyó</td>
<td>nërë</td>
<td>namoro</td>
<td>mëë</td>
<td>mëtxamo</td>
</tr>
<tr>
<td>Karihona</td>
<td>nërë</td>
<td>namoro</td>
<td>mëë</td>
<td>mëtxamo</td>
</tr>
<tr>
<td>Hâkaryana</td>
<td>noro</td>
<td>ñamoro</td>
<td>mëë</td>
<td>mëtxamo</td>
</tr>
<tr>
<td>Waiwai</td>
<td>noro</td>
<td>ñomoro</td>
<td>mëë</td>
<td></td>
</tr>
<tr>
<td>Katxayana</td>
<td>noro</td>
<td>mosoro</td>
<td>mëë</td>
<td></td>
</tr>
<tr>
<td>Karinya-Hf</td>
<td>inooro</td>
<td>inaaro</td>
<td>mëë</td>
<td></td>
</tr>
<tr>
<td>Karinya-Ms</td>
<td>hooro</td>
<td>hoorocon</td>
<td>mëë</td>
<td></td>
</tr>
<tr>
<td>Apalai</td>
<td>inoro</td>
<td>inaro</td>
<td>mëë</td>
<td>mëtxamo</td>
</tr>
<tr>
<td>Wayana</td>
<td>inërë</td>
<td>inamoro</td>
<td>mëë</td>
<td></td>
</tr>
<tr>
<td>De’kwana</td>
<td>nërë</td>
<td>nührro</td>
<td>mëë</td>
<td></td>
</tr>
<tr>
<td>Yawarana</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yukpa</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Waimiri</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tamanaku</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cumanagot</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chayma</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Taurepán</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Makushi</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Akawayo</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ingarikó</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arekuna</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panare</td>
<td>më(f)</td>
<td>mëhtsamon</td>
<td>kën</td>
<td></td>
</tr>
<tr>
<td>Ikpeng</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arara</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bakairi</td>
<td>inërë</td>
<td>asamor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kuhihurú</td>
<td>üple</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The an**imate 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 Bakairi, judging by the collective form). The second vowel is *o₂*, and the final vowel *o₃*, given their different correspondences (ë : 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 mëxamo, mëčjan, mëkijan, ohkijan (cf. fn. 15).
22 Meira (2000:60) listed më’txamo, më’jamo, mëkijano as Akuriyó collective pronouns. More recent data (presented here) shows that these forms were mistaken (probably Tiriyó influence).
23 De’kwana nërë is described as a distal form; the anaphoric pronoun is ëvë, a non-cognate.
24 Cf. fn. 16.
25 Bruno (pers. comm.) describes Waimiri as (surprisingly) lacking an animacy distinction. Irë (cf. Table 3) is also used to refer to people; and mëki ‘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 mëi and mëhe (for which no good semantic description is yet available), seems less likely: no other language has two forms, and, except for the Karipona mëhe (which cannot be cognate with Wayana mëhe, since Karipona 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 Karipona 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, rì in Katxuyana, Taurepán, and Makushi, and possibly also for the nì in Hixkaryana and Taurepán).

The animate medial and distal pronouns share suggestive similarities. Looking at cases such as Apaláí mokïro vs. mokì, Chayma muckere vs. muke, Waiwai mïkro vs. mïkì, 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. For instance, it is not so hard to imagine Apaláí as having a single distal term mokì 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 mokï) and mokaro (for mokïro) 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-Hê moqkan and mokkaaro. 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: moqko does not look like mokì + ro. One wonders if there could be a connection with the 1+2 pronoun kïkì).

The form of this animate distal pronoun presents relatively few problems: *moqkì seems to be the best hypothesis. Almost all languages have an initial syllable mo, mé, mue (e in Kuhikuru); it is easier to assume that Panare, Tamanaku and Ikpeng lost it. The final syllable kì, or clear reflexes of it (e.g. De’kwana ’, Tiriyó long ēè) are also overwhelming. The few problematic cases are: (1) Bakairí awëkë, which is not a clear cognate; (2) De’kwana më’kì ‘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 than Panare kën comes from *moqkì-ro2, with *ro2 > n (cf. the inanimate medial mën and the second-person pronoun amën 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 Tiriyó, mëêrë and mëkì are semantically very distinct; they are clearly two lexical items (cf. Meira to ap-a). Panare kën and muku (also attested as mìkì) also seem to be clearly independent, at least formally.
particle *ro2: The anaphoric collective can be reconstructed as resulting from *(i)n + *jamo (without the final *ro2, 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 *ro2 (e.g. Panare kamonton < *jamo ro2 tomo). In Taurepán mësémonan, 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) Makushi insemon, inkamoro with an initial unexplained i- (perhaps related to the (i)- in the anaphoric forms *(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:k-jamo and *mo:tx-amo. They are reduced, since almost all reflexes are reduced; Tiriyó mëkiñamo may have been analogically rebuilt, apparently a frequent phenomenon in Tiriyó collectives — cf. the 1+2Col and 2Col forms. In fact, Tiriyó mëesamo 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:txamo > *mëesamo; at this point *-txamo is reanalyzed as a suffix, while *mo:txo3 > mëe; then *mëe-txamo > mëesamo.

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. Makushi, Kuhikuru).

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

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>*awi</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>*amo</td>
<td>*am-jamo</td>
</tr>
<tr>
<td>1+2</td>
<td>*kiCV</td>
<td>*kiC-jamo</td>
</tr>
<tr>
<td>1+3</td>
<td>*(apina)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Categ.</th>
<th>Inanimate</th>
<th>Animate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N-Col</td>
<td>Animate</td>
</tr>
<tr>
<td>Anaph</td>
<td>*iro2</td>
<td>*(i)n(j)amo</td>
</tr>
<tr>
<td>Prox-1</td>
<td>*o:ro2</td>
<td>*(i)n(j)amo</td>
</tr>
<tr>
<td>Prox-2</td>
<td>*o:ni</td>
<td>*mo:tx-amo</td>
</tr>
<tr>
<td>Med</td>
<td>*mo:ro2</td>
<td>*mo:ki</td>
</tr>
<tr>
<td>Dist</td>
<td>*mo:ni</td>
<td>*mo:k-jamo</td>
</tr>
</tbody>
</table>

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 — *ro₂, *mî, *mo₂, *kî — which suggest that the third-person pronouns are actually old combinations of yet older pronouns. The anaphoric *iro₂ could be a combination of a third-person marker *î- (from Gildea’s *jî-) with the element ro₂, 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 *ro₂). The element *mo₂, also found in combination with *ro₂, might be compared to the ‘evidential’ mo or më that, in some languages, occurs with the third-person prefix to indicate certain evidential values (e.g. Wayana nî-tê-jai ‘he is going’, më-n-tê-jai ‘he is going (but I do not see him)’; Hixkaryana mo-n-ewe-h-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

Amodio, Emanuele, & Vicente Pira

Armellada, Cesáreo de, & Jesús Olza
1994 Gramática de la lengua pemón (morfosintaxis), San Cristóbal (Venezuela): Universidad Católica del Táchira.

Bruno, Ana Carla dos Santos

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

Edwards, Walter F.

1977 An introduction to the Akawayo and Arekuna peoples, Amerindian Languages Project, Georgetown: University of Guyana.

Gildea, Spike


Gilij, Felipe [Filippo] S.


Girard, Victor James


Hall, Katherine L.


Hawkins, Robert E.


Hoff, Berend J.


Jackson, Walter S.


Kaufman, Terrence K.


Koch-Grünberg, Theodor

Koehn, E., & S. Koehn.

Meira, Sérgio

to ap-a ‘Non-contrastive exophoric use of Tiriyó 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.

Muller, Marie-Claude Mattei

Pacheco, Frantomé Bezerra

Robayo, Camilo

Ruiz Blanco, Matias

Souza, Isaac de
Steinen, Karl von den.

Tauste, Francisco de

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

Wheatley, James


Yangues, Manuel de