A comparative account of intransitive verbs with conservative first person forms in Cariban

Nine Cariban languages have a group of intransitive verbs irregularly inflected for first person. The irregular first person markers are conservative, contrasting with innovative regular markers. They are a result of person marker extensions not affecting some verbs, which happened independently in six (proto-)languages. These six incomplete extensions left between one and seven conservatively inflected verbs, which show a great etymological overlap across (proto-)languages. Bybee's network model of morphology was used to generate hypotheses about factors causing the distribution of conservative and innovative markers in each language. Predictions of different possible factors were then tested against the data. Because of patterns reconstructible to Proto-Cariban, the hypothesized factors largely overlap in their predictions, though phonology (combined with frequency) shows a strong overall performance.

Keywords: Cariban, person marking, irregular, language change, network morphology

Contents

List of Figures

List of Tables

1 Introduction

The Cariban language family is one of the largest of South America, with between 60'000 and 100'000 speakers unevenly distributed between 22 to 25 extant languages (Gildea 2012: 441). The family is concentrated in Venezuela, the Guianas and Northern Brazil, with three Western and four Southern outliers. Figure 1 shows the geographical distribution and genealogical affiliation of the extant Cariban languages. For overviews of and comparative work on the family, readers are referred to Gildea (1998), Derbyshire (1999), Meira (2002), Meira & Franchetto (2005), Meira (2006a), Gildea & D. Payne (2007), Meira et al. (2010), Gildea et al. (2010), Gildea (2012), Matter (2021a), and Gildea & Cáceres (in preparation).

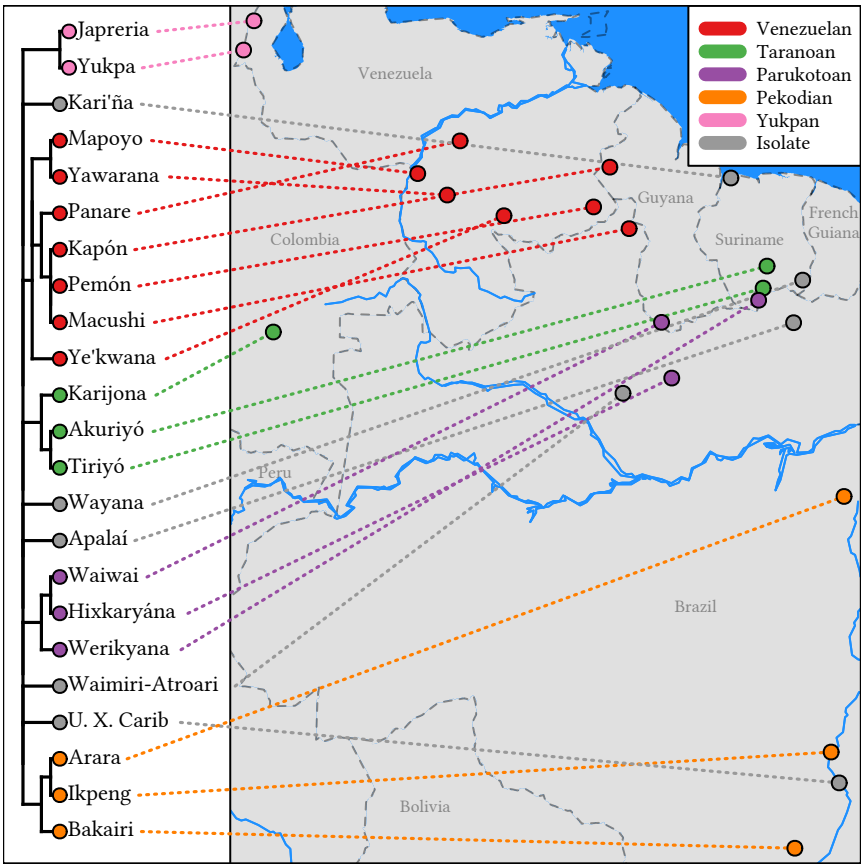

Figure 1: The Cariban language family. Classification from Matter (2021a), positions from Hammarström et al. (2020). Made with a modified version of `phytools` (Revell 2012)

In some Cariban languages, a small group of verbs show a divergent first person inflection pattern, a topic which has not received much attention in the literature. This is illustrated for Hixkaryána in Table 1,[1] with paradigms of four verbs, all members of the S_A inflectional class. In this language, the

Table 1: Some Hixkaryána verbs (Howard 2001: 150, 510-511, 513, 520; Derbyshire 1985: 197-198)

	'to fall'	'to be afraid'	'to walk'	'to cut self'	'to be'
1	*k-ehurka-*	*k-oser^{j}ehɨ-*	*k-atar^{j}eknohɨ-*	*k-atama-*	*w-eʃe-*
2	*m-ehurka-*	*m-oser^{j}ehɨ-*	*m-atar^{j}eknohɨ-*	*m-atama-*	*m-eʃe-*
1+2	*t-ehurka-*	*t-oser^{j}ehɨ-*	*t-atar^{j}eknohɨ-*	*t-atama-*	*t-eʃe-*
3	*ɲ-ehurka-*	*n-oser^{j}ehɨ-*	*n-atar^{j}eknohɨ-*	*n-atama-*	*n-eʃe-*

[1]The presence of a 1+2 person value implies that of a 1+3 value. This is usually expressed with a free pronoun combined with third person morphology in Cariban languages, so it is not represented as a distinct value in Table 1 and other paradigm

Table 2: Some Tiriyó verbs (Meira 1999: 292, 294; Carlin 2004: 274)

	'to sleep'	'to see self'	'to bathe (INTR)'	'to yawn'	'to go'
1	*t-əənɨkɨ-*	*t-əene-*	*s-epɨ-*	*s-entapo-*	*wɨ-tən-*
2	*m-əənɨkɨ-*	*m-əene-*	*m-epɨ-*	*m-entapo-*	*mɨ-tən-*
1+2	*kɨt-əənɨkɨ-*	*k-əene-*	*ke-epɨ-*	*k-entapo-*	*kɨ-tən-*
3	*n-əənɨkɨ-*	*n-əene-*	*n-epɨ-*	*n-entapo-*	*nɨ-tən-*

verb 'to be' diverges from other S_A verbs like 'to fall' by having a first person marker *w-*, rather than *k-*. A similar pattern exists in Tiriyó (Table 2), where the verb 'to go' has a first-person prefix *wɨ-* while other S_A verbs have a prefix with phonologically conditioned allomorphs *t-* / *_ə* and *s-* / *_e*. In both languages, the first person prefix of the verbs on the left is representative for the vast majority of S_A verbs.

Such divergent verbs have been identified for Hixkaryána (Derbyshire 1985: 188), Waiwai (Gildea 1998: 90), the three Taranoan languages (Meira 1998: 112–115), Bakairi (Meira 2003a), and Arara (Alves 2017: 153), but have only been subject to comparative scrutiny in Meira's (1998) reconstruction of Proto-Taranoan. In synchronic analyses, these verbs and their first person prefixes may be called IRREGULAR, contrasting with regular prefixes (like Hixkaryána *kɨ-* and Tiriyó *t-*/*s-*) on regular verbs. Note that there is no widely accepted definition of irregularity (Stolz et al. 2012), and many stricter definitions (e.g., Haspelmath & Sims 2010) require the pattern to occur at a single place in the grammar. In such approaches, the verbs simply belong to a small inflectional (sub-)class, an analysis applied to the Pekodian languages Bakairi and Arara (Meira 2003a: 4; Alves 2017: 149).

Regardless of the details of synchronic analysis, the cause for the divergent inflectional patterns is found in the diachrony of the languages in question. The goal of this study is to approach these patterns from a comparative perspective and to provide a diachronic and functional account, proceeding as follows: In Section 2, relevant aspects of the Proto-Cariban verbal system are introduced, and it is shown that the mechanism of person marker extensions is responsible for patterns like in Tables 1 and 2 . In Section 3, six incomplete person marker extensions and the verbs unaffected by them are described. Since languages show considerable etymological overlap in their conservative verbs, these are further discussed and reconstructed. Section 4 uses Bybee's (1985) network model of morphology to search explanations for the verbs (un-)affected by each extension. Section 5 summarizes and discusses the results of the study.

2 The origins of irregular first person inflections

The irregular first person prefixes from Section 1 are relics, inherited from the ancestral Proto-Cariban system (Section 2.1). That system underwent much innovation; the mechanism responsible for the irregular forms are person marker extensions not spreading through the entire S_A lexicon (Section 2.2). A specific aspect of the system, the S_A vs S_P distinction, plays a role in incomplete extensions and is discussed in Section 2.3.

2.1 Proto-Cariban person marking and inflectional relics

Proto-Cariban is reconstructed by Gildea (1998) as using a person paradigm called Set I in its independent verb forms, shown in Table 3. Person indexation in transitive verbs was conditioned by a basic hierarchy $\boxed{1/2 > 3}$. The locuphoric markers had two forms, an A-oriented one for direct (SAP>3) scenarios and a

tables. Further, TAM suffixes in the attested forms are omitted, since a) the focus lies on the prefixes and stems, and b) full paradigms with the same TAM suffix are rare in the available sources. Standard IPA symbols are used in the transcription of Cariban languages, except for coronal rhotics, which are simply represented with ⟨r⟩, rather than ⟨ɽ⟩ for Wayana or ⟨ɾ⟩ for Ye'kwana etc. In languages with strong morphophonological processes and/or non-phonemic orthography the original form is shown in an additional line when presented in interlinearized examples. Gildea (2018) is followed in using ⟨ə⟩ for the Proto-Cariban reconstructed by Meira & Franchetto (2005), although it was likely more back (Gildea et al. 2010).

Table 3: Proto-Cariban Set I (main clause) person markers (Meira et al. 2010: 495; Gildea & Zúñiga 2016: 497)

(a) Transitive

A/P	1	2	1+2	3
1		*k-		*t(i)-
2	*k-			*m(i)-
1+2				*kɨt(i)-
3	*u(j)-	*ə(j)-	*k-	*n(i)-

(b) Intransitive

	S_A	S_P
1	*w-	*u(j)-
2	*m-	*ə(j)-
1+2	*kɨt-	*k-
3	*n-	*n(i)-

P-oriented one for inverse (3>SAP) scenarios. There was a single aliophoric marker *n(i)-, which only surfaced in nonlocal (3>3) scenarios, without morphologically expressed distinctions between different third person referents. Local scenarios were marked identically and non-transparently with the 1+2 prefix *k-.

Formally identical or etymologically related markers occured in intransitive verbs, which showed a split-S system (Table 3b). S_A verbs took similar markers as the A-oriented ones in transitive verbs, with the exception of first person ($1S_A$ *w- vs 1>3 *t(i)-) and the absence of *i after all S_A prefixes. On the other hand, S_P verbs took markers fully identical to the P-oriented ones, with $3S_P$ *n(i)- aligning with 3>3 scenarios.

Knowledge about the ancestral system makes clear that the divergent Hixkaryána and Tiriyó forms in Tables 1 and 2 behave irregularly because they preserve the original Proto-Cariban $1S_A$ prefix *w-; they are therefore CONSERVATIVE. They contrast with regular S_A verbs, which are innovative in both languages. The reflexes of *w- are RELICS, old and restricted to a few lexemes, contrasting with the innovative prefixes found elsewhere. These verbs and their prefixes are comparable with the few English nouns like *ɒks*, which preserve the old plural suffix *-ən*. It was once more widespread as the normal plural suffix of the weak inflection, compare German *ɔks-ən* 'ox-en', *na:mə-n* 'name-s', *ha:zə-n* 'hare-s', *bɛ:ʁ-ən* 'bear-s'. Since the regular Hixkaryána and Tiriyó prefixes are innovative, the question arises how they developed.

2.2 Person marker extensions and lexical diffusion

In his discussion of the Proto-Cariban split-S system (Section 2.3) and reconstruction of the intransitive person prefixes, Gildea (1998: 88–96) shows that the system has undergone many different changes in daughter languages. The main mechanism of these changes are PERSON MARKER EXTENSIONS, person prefixes being extended to verbal paradigm cells previously occupied by other prefixes. There have been many person marker extensions in Cariban languages, some still ongoing. Gildea (1998) illustrates this with the three Parukotoan languages Werikyana, Hixkaryána, and Waiwai. Apart from segmental changes to individual morphemes, the following innovations happened in the Set I paradigm in Parukotoan:

(1)

Proto-Parukotoan	$1S_A$ *w- to 1>3
	1+2 *k- to $1S_P$[a]
	1+2 *kɨt- to $1+2S_P$[a]
Proto-Waiwaian	$1S_P$ *k- to $1S_A$
	*owɨ(ro) j- '1PRO LK' for 1P
Waiwai	$2S_A$ m- to $2S_P$

[a] Completed in Proto-Waiwaian, ongoing in Werikyana.

All innovations are person marker extensions except 2b, which combined a pronoun with the linker *j-. Figure 2 shows them in bold and reproduces Gildea's (1998) tables as a tree diagram, with adapted transcription and an additional Werikyana $1S_P$ marker ∅/*j*- (Spike Gildea, p.c.). Hixkaryána has preserved

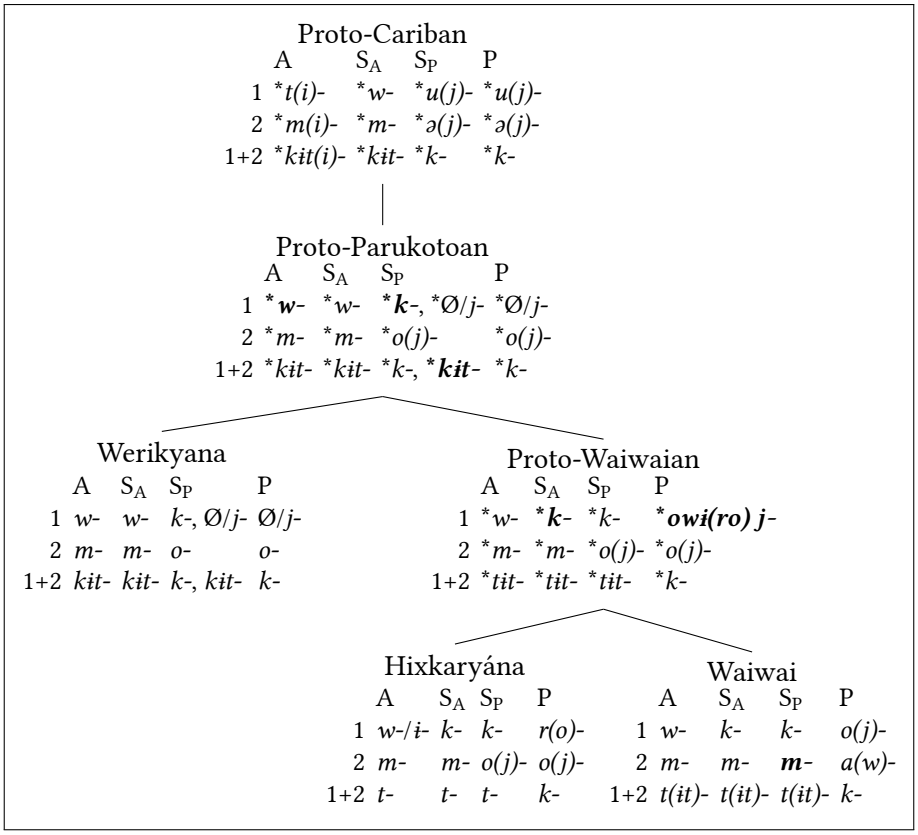

Figure 2: Person marker extensions in Parukotoan, after Gildea (1998: 94)

split-S only in the second person prefixes, while Werikyana still shows the variation reconstructible to Proto-Parukotoan in its first person and 1+2 prefixes. Waiwai has lost the system entirely, which notably happened via three diachronically distinct innovations.

For Gildea (1998), person marker extensions are relevant for the loss of split-S and the accompanying changes to indexing alignment, whereas this study focuses on a different aspect. Namely, they most likely took place via lexical diffusion, characterized as a type of extension by Harris & Campbell (1995: 106–115), a hypothesis supported by three facts. First, the variation in first person and 1+2 prefixes described above for Werikyana is not completely free. Some verbs only allow e.g. first person *k-*, but not *j-*, while others can occur with both, a pattern expected in a lexical diffusion scenario. In addition, this is speaker-dependent (Spike Gildea, p.c.), which points to an ongoing change. Second, while there is no detailed diachronic account of the switch of 1>3 **t-* and $1S_A$ **w-* in the Tiriyoan languages (Section 3.1.3), Meira (1998: 111–112) argues that it must have happened gradually rather than instantaneously, and entailed both markers spreading simultaneously. Whether or not this gradual switch followed ordered lines, lexical diffusion must have played a role.

The third argument in favor of the lexical diffusion scenario goes back to conservative forms like in Hixkaryána and Tiriyó. Both innovative $1S_A$ prefixes were introduced by a person marker extension spreading via lexical diffusion. The continued presence of the old $1S_A$ prefix in a few verbs is the result of the extension stopping short of these verbs, rather than spreading through the entire S_A lexicon. In a family-wide search, 18 distinct extensions affecting intransitive verbs were identified, 6 of them incomplete. The latter have left between 1 and 7 conservatively inflected verbs in 9 Cariban languages (Section 3).

Interestingly, all six featured innovative first person markers on S_A verbs. All other (complete) extensions[2] either occurred with other person values and/or targeted S_P verbs. Illustrative examples for

[2] As an honorable mention, when Ikpeng replaced third person Set I with Set II prefixes, *a* 'to be' and *ke* 'to say' retained *n-* (Matter 2021b: 12). However, innovative markers were not spreading within a paradigm, but rather from former subordinate

Table 4: Some examples for completed extensions (Gildea 1998: 90–92)

(a) Apalaí

	S_A	S_P
1	*ɨ-*/∅	*ɨ-*/*j-*
2	*m(ɨ)-*	*o-*
1+2	***s(ɨ)-***	
3	*n(ɨ)-*	

(b) Panare

	S_A	S_P
1	*w(ɨ)-*	∅/*j-*
2	***m(ɨ)-***	
1+2	*n(ɨ)-*[a]	
3	*n(ɨ)-*	

[a] 1+2 was lost as a person value.

(c) Waimiri-Atroari

	S
1	***w(ɨ)-/i-***
2	***m(ɨ)-***
1+2	***h(ɨ)-***
3	*n-*/∅

complete extensions are shown in Table 4: the extension of 1+2S_A *s(ɨ)-* (< **kɨt-*) to S_P verbs in Apalaí (Table 4a), of 2S_A *m(ɨ)-* in to S_P verbs in Panare (Table 4b), and the extension of the entire S_A set to S_P verbs in Waimiri-Atroari (Table 4c). The starkly different behavior of S_A and S_P verbs regarding extensions points to the split-S system playing a role, so its main properties will be discussed in Section 2.3. This will also clarify how the S_A/S_P distinction can be lost for a single person, or how S_P verbs can take on S_A markers with apparent semantic impunity.

2.3 The Cariban split-S system

As seen in Section 2.1, the split between S_A and S_P verbs was instantiated by inflection classes within the Proto-Cariban Set I person paradigm, but this was not the only difference: In deverbalized forms, S_A verbs took a class marker **w-*, while S_P verbs lacked that prefix (Gildea 1998: 89, 141–142; Meira 2000: 208). Also, S_P verbs took the 2S_P prefix **ə(j)-* in imperatives, while S_A verbs were unprefixed (Gildea 1998: 89; Meira 2000: 208).

In modern instantiations of Cariban split-S, mismatches between the semantics of verbs and their S_A or S_P status are common, exemplified with Kari'ña data in (2).

(2) Kari'ña

a. *sipi tɨnka-rɨ m-ekema-non hen*
net pull-NMLZ 2S_A-be.afraid-PRS.UNCERT eh?
'You're afraid to pull up the net, aren't you? ' (Courtz 2008: 253)

b. *aya:woiya*
aj-awomɨ-ja
2S_P-get.up-PRS
'You are getting up.' (Hoff 1968: 167)

The S_A verb *ekema* 'to be afraid' takes an A-oriented marker (2a), while the S_P verb *awomɨ* 'to get up' takes a P-oriented marker (2b). In both cases, the prefix does not appear to contribute to the semantics of the predicate, since there are clear mismatches: 'to be afraid' with an "agentive" marker can hardly be considered a volitional act, while 'to get up' with a "patientive" marker is clearly volitional. Meira (2000) investigates a corpus of intransitive verbs from Tiriyó, Kari'ña, Apalaí, and Wayana, and categorizes them by applying different criteria commonly encountered in split-S systems. He shows that neither (non)activities, (non-)agency, (in-)animacy, nor Aktionsart satisfactorily predict the class membership of intransitive verbs in any of the languages.

Rather, the reason for a verb to take A- or P-oriented prefixes is (at least diachronically) a morphological one. Meira (2000: 217–221) demonstrates that those intransitive verbs which (etymologically) have a derivational detransitivizing prefix are treated as S_A verbs, while essentially all others are S_P verbs:

to main clauses.

> Almost all verbs in the S_A class are detransitivized forms of transitive verbs, either synchronically (with still exisiting transitive sources) or diachronically (with reconstructible but no longer existing transitive sources) (Meira 2000: 201)

He notes that this leads to an inflectional split not based in meaning, but rather morphology:

> Apparently, the morphological behavior of the S_A verb class is an accidental consequence of the fact that detransitivization, as far back as we can reconstruct, entails all the morphology described [...] as typical of S_A verbs. The alignment of person-marking prefixes appears not to be driven by any semantic forces in the language; it is as though they were being dragged by the evolution of the reflexive marker. (Meira 2000: 226)

Regarding the form of this marker, Meira et al. (2010: 505–512) reconstruct two distinct prefixes for Proto-Cariban: reciprocal **ətɨ-* and reflexive **e-*, although their reflexes on verbs have been merged into a single morpheme in modern languages.[3] Reflexes of **ətɨ/e-*show a range of meanings summarizable as "detransitive", illustrated with Tiriyó S_A verbs in (3).

(3) Tiriyó (Meira 2000: 218–219, 1999: 128, 256)

nonta 'abandon'	→	*e-nonta*, *ətɨ-nonta*	'abandon each other' (reciprocal)
suka 'wash'	→	*e-suka*, *ətɨ-suka*	'wash self' (reflexive)
pahka 'break (TR)'	→	*e-pahka*	'break (INTR)' (anticausative)
puunəpɨ 'think about'	→	*əh-puunəpɨ*, *ətɨ-puunəpɨ*	'think, meditate' (antipassive)

The morphological variation in 'abandon each other' and 'wash self' is due to the collapse of the two Proto-Cariban prefixes: *e-* comes from the reflexive prefix **e-*, while the form *ətɨ-* originates in reciprocal **ətɨ-*. However, both can occur with either meaning – at least for these two verbs. In the next section, it will be seen that many of the verbs not affected by person marker extensions belong to the small group of S_A verbs without a reflex of **ətɨ/e-*.

3 Prefixes and verbs: innovation and resistance

As shown in Section 2.2, irregularly inflected first person forms are leftovers of incomplete person marker extensions. Section 3.1 presents the six identified incomplete extensions, the prefixes they introduced and the verbs they spared. Since languages show considerable etymological overlap in their conservative verbs, those are compared and reconstructed in Section 3.2. Further, their reflexes which did get affected by an incomplete extension are identified. Section 3.3 summarizes the results.

3.1 Incomplete extensions: the innovative 1S_A markers

As stated in Section 2.2, the six person marker extensions which left a group of verbs untouched all introduced innovative first person markers on S_A verbs. Of these extensions, half can be reconstructed to intermediate proto-languages, and half happened in pre-modern stages of single languages. The sources

[3] In the Proto-Cariban transitive verb template, derivational prefixes were in a paradigmatic relationship with the earlier third person marker **i-*: **m-i-*V 'you V it', **m-e-*V 'you V yourself', **m-ətɨ-*V 'you V each other'. This analysis is applied to Tiriyó by Carlin (2004: 268–269), who interprets *i-* as marking transitive diathesis.

Table 5: Regular Pekodian S_A verbs (Meira 2003a: 4; Alves 2017: 150; Pachêco 2001: 52)

	Bakairi 'to go up'	Arara 'to dance'	Ikpeng 'to run'
1	*k-əku-*	*k-origu-*	*k-aranme-*
2	*m-əku-*	*m-origu-*	*m-aranme-*
1+2	*kɨd-əku-*	*kud-origu-*	*kw-aranme-*
3	*n-əku-*	*∅-origu*	*∅-aranme-*

of innovative markers vary, but not much: the innovative $1S_A$ prefix is formally identical to the $1+2P/S_P$ marker (Proto-Cariban **k*-) in three cases, to the $1P/S_P$ marker (Proto-Cariban **u(j)*-) in two cases, and to the 1>3 marker (Proto-Cariban **t*-) in one case. For each extension, regular (innovative) verbs are contrasted with irregular (conservative) ones, and verb forms are reconstructed where necessary. Section 3.1.1 demonstrates the extension of **k*- in Proto-Pekodian, reflected in the three daughter languages Arara, Ikpeng, and Bakairi. Section 3.1.2 details the extension of **k*- in Proto-Waiwaian, briefly shown in Section 2.2. Section 3.1.3 focuses on innovative **t*- in Proto-Tiriyoan, reflected in modern Tiriyó and Akuriyó. The topic of Sections 3.1.4 to 3.1.6 are innovative $1S_A$ markers found in single languages: *k*- in Akuriyó, and *j*- in Carijona and Yukpa.

3.1.1 Proto-Pekodian **k*-

The Pekodian branch consists of closely related Arara and Ikpeng, with Bakairi as a more distant member. The contribution establishing the branch (Meira & Franchetto 2005) focused on phonology and lexicon, so no reconstructions of Proto-Pekodian morphosyntax are found in the literature. However, all three Pekodian languages have a regular $1S_A$ marker *k*- (Table 5), allowing the reconstruction of a Proto-Pekodian $1S_A$ marker **k*-.

The most detailed description of a Pekodian language (Alves 2017) names six[4] Arara S_A verbs forming a subclass defined by a first person marker *w(ɨ)*- rather than *k*-, shown in (4). There is also a reflex of the copula **a[p]* (Section 3.2.3), serving syntactically as a postposition introducing adverbial clauses meaning 'if' or 'when' (Alves 2017: 199–201). However, it inflects with verbal Set I prefixes, including first person *w*- (5).

(4) Arara (Alves 2017: 153)

wɨ-genɨ	'I said'
w-ɨ͡ʃ͡ʃinɨ	'I was, lied down'
w-ebɨnɨ	'I came'
w-ibɨnɨ	'I bathed'
w-iptoŋrɨ	'I went down'
w-ɨdolɨ	'I went'

(5) Arara (Alves 2017: 200)

1	*w-aptam* 'when/if I was'
2	*m-od-aptam*
1+2	*kud-aptam*
3	*∅-aptam*

In his brief but precise discussion of Bakairi verbal person marking, Meira (2003a) reports the existence of two S_A subclasses, defined by $1S_A$ *w*- and *k*-, respectively.[5] The first group is illustrated with *i* 'to bathe' (6).

[4] Seven in her analysis, treating the two meanings of *ɨ͡ʃ͡ʃi* 'to be, to lie down' as different verbs.

[5] Meira (2003a) indicates that the same verbs which take first person *w*- in Bakairi also take a 1+2 marker *k*-. However, this marker is only illustrated for 'to bathe', both by Meira (2003a) and von den Steinen (1892). Given the lack of data for other verbs, this potential additional pattern will not be further discussed. If the characterization by Meira is accurate, then verbs with innovative first person prefixes have conservative 1+2 prefixes, and vice versa.

(6) Bakairi (Meira 2003a: 4)
ɯ-i-də
$1S_A$-bathe-IMM
'I bathed'

While Meira (2003a: 4) lists some Bakairi cognates of the Arara verbs in (4) as S_A verbs, he does not indicate whether they belong to the S_A-1 class with *k-*, or the S_A-2 class with *w-*. However, inflected forms can be found in von den Steinen (1892), represented in (7) based on the analyses of Bakairi phonology and verbal morphology by Wheatley (1969), Meira (2003a, 2005), and Franchetto & Meira (2016).

(7) Bakairi (von den Steinen 1892: 131, 397, 76, 137, 374, 130)

a. ⟨u-ɣépa⟩
u-ge-pa
$1S_A$-say-NEG
'I don't say.'

b. ⟨wi-táki⟩ / ⟨wi-tági⟩
ɯ-i-taki
$1S_A$-be-INT
'I was.'

c. ⟨kχaewí-le⟩
k-əewɨ-lɨ
$1S_A$-come-IMM
'I came.'

d. ⟨kχ-itaké-he⟩
k-ɨtəgɨ-se
$1S_A$-go.down-NPST?
'I go down.'

e. ⟨űta⟩ / ⟨uúta⟩
u-tə
$1S_A$-go
'I go.'

f. ⟨töre-w-akine⟩
tərə ɯ-a-kɨne
there $1S_A$-be-PST.CONT
'I was there.'

All descriptions of Ikpeng list *k-* as the only $1S_A$ marker (Pachêco 1997: 55; Campetela 1997: 105; Pachêco 2001: 64; Alves Chagas 2013: 205). However, most Ikpeng cognates of the verbs in question do not take *k-*, but rather *ɨ-* or Ø (8), with the exception of *k*-prefixed 'to go' (9). There is a formally identical Ikpeng cognate of Arara *iptoŋ* 'to go down', but no first person forms are attested (Angela Chagas, p.c.). While reflexes of **a[p]* 'be-1' do exist in Ikpeng, apparently only reflexes of **eti* 'be-2' occur with first person prefixes (Gildea 2018: 401).

(8) Ikpeng

a. *ɨ-ge-lɨ*
1-say-REC
'I said.' (Pachêco 2001: 209)

b. *Ø-eʧi-lɨ*
1-be-REC
'I was.' (Pachêco 2001: 139)

c. *aʧagotpop Ø-ip-ʧi ik-gwa-kʧi*
always 1-bathe-NPST river-LOC.aquatic-ALL
'I always bathe in this river.' (Pachêco 1997: 68)

(9) Ikpeng (Pachêco 2001: 80)
k-aran-ʧi
1-go-NPST
'I'm going.'

Table 6: Verbs preserving 1S$_A$ **w-* in Proto-Pekodian (Alves 2017: 153, 200; Pachêco 2001: 42, 80, 139, 209; von den Steinen 1892: 76, 130-131, 374, 397; Pachêco 1997: 68; Meira 2003a: 4)

	Proto-Pekodian	Arara	Ikpeng	Bakairi
'say'	**wɨ-ge-*	*wɨ-ge-*	*ɨ-ge-*	*u-ge-*
'go'	**w-ɨtən-*	*w-ɨdo-*	*k-aran-*	*u-tə-*
'be-1'	**w-ap-*	*w-ap-*	–	*w-a-*
'be-2'	**w-e͡ʧi-*	*w-iʧi-*	*∅-e͡ʧi-*	*w-i-*
'come'	**w-epɨ-*	*w-ebɨ-*	*k-arep-*	*k-əewɨ-*
'go down'	**w-ɨptə-*	*w-ɨptoŋ-*	*?-ɨptoŋ-*	*k-ɨtəgɨ-*
'bathe'	**w-ipɨ-*	*w-ibɨ-*	*∅-ip-*	*w-i-*

Table 7: Loss of **w* in Ikpeng (de Souza 1993: 44, 70; Alves Chagas 2013: 118; Alves 2017: 143; Pachêco 2001: 21, 164; de Souza 2010: 9; Campetela 1997: 40)

Meaning	Arara	Ikpeng
'to defecate'	*watke*	*atke*
'DAT'	*wɨna*	*ɨna*
'dog'	*wokori*	*akari*
'capuchin monkey'	*tawe*	*tae*
'to sleep'	*wɨnkɨ*	*ɨnkɨ*

Reconstructed Proto-Pekodian forms of conservatively inflected verbs are given in Table 6. Newly identified Ikpeng *ɨ-*/Ø is demonstrably a reflex of Proto-Xinguan **w(ɨ)-*, based on other (albeit irregular) cases of loss of **w* (Table 7). Similarly, the change of **wɨ* to Bakairi *u* is found in correspondences like *udo* (Meira & Franchetto 2005) from Proto-Cariban **wɨtoto* 'person' (Gildea & D. Payne 2007: 4). Thus, a 1S$_A$ prefix **w(ɨ)-* can securely be reconstructed to Proto-Pekodian, identical to its Arara reflex in form and distribution. In later, individual developments, Bakairi extended *k-* to 'to go down', and Ikpeng to 'to go'.

Reconstructions of verb stems are deferred to Section 3.2, but a comment on 'to come' is in order: The stems are not fully cognate, as Ikpeng and Bakairi both show a reflex of the Proto-Pekodian detransitivizer **əd-* in combination with a root reconstructible as **epɨ* (Section 3.2.4). In contrast, the Arara first person form has a bare reflex of **epɨ*. While reflexes of **əd-epɨ* can be found elsewhere in the Arara paradigm (10), Ikpeng and Bakairi uniformly reflect **əd-ebɨ*.

(10) Arara (Alves 2017: 150)
m-odebɨ-nɨ
2S$_A$-come-REC
'You came.'

Following Meira's (1998: 114) line of reasoning for a similar pattern in Taranoan (see also Section 3.1.3), the idiosyncratic pattern in Arara can be reconstructed to Proto-Pekodian, with Bakairi and Ikpeng independently leveling the paradigm in favor of **əd-epɨ*.

3.1.2 Proto-Waiwaian **k-*

This extension, one of the Parukotoan innovations shown in Section 2.2, resulted in the Hixkaryána patterns from Section 1. Proto-Waiwaian further extended the 1S$_P$ prefix **k-* (innovated in Proto-Parukotoan) to 1S$_A$. For regularly inflected verbs, this created a unified 1S category (Table 8).

Table 8: Regular 'to fall' (S_A) and 'to sleep' (S_P) in Proto-Waiwaian (Derbyshire 1985: 189-190, 196; W. N. Hawkins & R. E. Hawkins 1953: 209, 211; Howard 2001: 150; R. E. Hawkins 1998: 30)

	'to fall'			'to sleep'		
	Proto-Waiwaian	Hixkaryána	Waiwai	Proto-Waiwaian	Hixkaryána	Waiwai
1	*k-eɸurka-*	*k-ehurka-*	*k-eɸirka-*	*kɨ-wɨnikɨ-*	*kɨ-nikɨ-*	*kɨ-wɨnikɨ-*
2	*m-eɸurka-*	*m-ehurka-*	*m-eɸirka-*	*o-wɨnikɨ-*	*o-wnikɨ-*	*mɨ-wɨnikɨ-*
1+2	*t-eɸurka-*	*t-ehurka-*	*ʧ-eɸirka-*	*tɨt-wɨnikɨ-*	*tɨ-nikɨ-*	*tɨt-wɨnikɨ-*
3	*ɲ-eɸurka-*	*ɲ-ehurka-*	*ɲ-eɸirka-*	*nɨ-wɨnikɨ-*	*nɨ-nikɨ-*	*nɨ-wɨnikɨ-*

Not all S_A verbs were affected: Waiwai *ka* 'to say' does not take *kɨ-*, but rather conservative *wɨ-* (11a). Its Hixkaryána counterpart has a prefix *ɨ-* (11b), which also occurs in 1>3 scenarios in Hixkaryána (11c), corresponding to Waiwai *wɨ-* (11d).

(11) a. Waiwai (R. E. Hawkins 1998: 71)
wɨɨkekɲe
wɨ-ka-jakɲe
1-say-PST
'I said.'

b. Hixkaryána (Derbyshire 1985: 124)
roxehra nay hamɨ Kaywerye ɨkekonɨ
ro-ʃe-hɨra n-a-je hamɨ kajwer^{j}e ɨ-ka-jakonɨ
1-DES-NEG 3-be-NPST.UNCERT EVID K. 1S_A-say-REM.CONT
'I said (to myself), "Kaywerye evidently doesn't like me".'

c. Hixkaryána (Derbyshire 1985: 191)
ɨ-koroka-no
1>3-wash-IMM
'I washed him.'

d. Waiwai (R. E. Hawkins 1998: 192)
wîyesî
wɨ-jo-jasɨ
1>3-boil-NPST
'I will boil it.'

The regular correspondence in transitive verbs points to Hixkaryána *ɨ-* on intransitive verbs as another reflex of **wɨ-*, with a similar phonological reduction as in Ikpeng (Section 3.1.1). Notably, Derbyshire (1985) analyzes this *ɨ-* as the regular 1>3 prefix, because he considers Hixkaryána *ka* 'to say' to be transitive (Section 3.2.1).

There are three more verbs which did not take innovative **k-* in Proto-Waiwaian (Table 9). The two forms for 'to be' are unproblematic, whereas 'to go' is a special case. While Hixkaryána has the expected *ɨ-*, Waiwai seems to have combined innovative *k-* with old **w-*, an analysis also considered by Gildea (1998: 90). Alternatively, this form may have been influenced by deverbalized forms of 'to go' showing a fossilized of the S_A class marker **w-* (e.g., *o-**w**to-topo-nho* 'my trip' [R. E. Hawkins 1998: 92]). Either way, the first person form Hixkaryána 'to go' clearly points to Proto-Waiwaian **wɨ-tom-*.

3.1.3 Proto-Tiriyoan **t-*

The moniker Tiriyoan (Hammarström et al. 2020) subsumes Tiriyó and Akuriyó, the more closely related of the three Taranoan languages identified by Girard (1971), with Carijona as a sister. Meira (1998) provides an extensive phonological, morphological, and lexical reconstruction of Proto-Taranoan, facing

Table 9: Verbs preserving $1S_A$ *w- in Proto-Waiwaian (Derbyshire 1979: 4; R. E. Hawkins 1998: 71, 85; Derbyshire 1985: 70, 197-198; p.c., Spike Gildea)

	Proto-Waiwaian	Hixkaryána	Waiwai
'say'	*wɨ-ka-	ɨ-ka-	wɨ-ka-
'go'	*wɨ-tom-	ɨ-to-	kɨw-tom-
'be-1'	*w-ah-	w-ah-	w-a-
'be-2'	*w-eʃi-	w-eʃe-	w-eeʃi-

Table 10: Regular Proto-Tiriyoan S_A verbs (Gildea 1994: 87; Meira 1999: 292, 294)

	'to bathe (INTR)'			'to sleep'		
	Proto-Tiriyoan	Tiriyó	Akuriyó	Proto-Tiriyoan	Tiriyó	Akuriyó
1	*t͡ʃ-epɨ-	s-epɨ-	t͡ʃ-epɨ-	*t-əənikɨ-	t-əənikɨ-	k-əənikɨ-
2	*m-epɨ-	m-epɨ-	m-epɨ-	*m-əənikɨ-	m-əənikɨ-	m-əənikɨ-
1+2	*ke-epɨ-	ke-epɨ-	ke-epɨ-	*kɨt-əənikɨ-	kɨt-əənikɨ-	kəʔ-əənikɨ-
3	*n-epɨ-	n-epɨ-	n-epɨ-	*n-əənikɨ-	n-əənikɨ-	n-əənikɨ-

an interesting puzzle in the Set I paradigms of Tiriyó and Akuriyó: Proto-Cariban 1>3 *t- and $1S_A$ *w- seem to have switched places, creating a regular $1S_A$ marker of the form *t͡ʃ- / _e, *t- / _ə (Table 10).[6] The question of how this switch happened in detail (Meira 1998: 107–112) is still open, although a scenario seems necessary in which both *t- and *w- for a time occurred on both transitive and intransitive verbs (Meira 1998: 112).[7]

As for verbs unaffected by the spread of *t-, Meira (1998) reconstructs four of the items in Table 11 as taking *w- in Proto-Taranoan, for which reconstructed Proto-Tiriyoan forms are substituted here.[8] As a fifth verb *eʔi 'be-1' can be added, whose Tiriyó reflex retains w-. The idiosyncratic Akuriyó first person prefix ə- on 'to go' is identified as a reflex of *wɨ- by Meira (1998: 113), which is supported the fact that both components of the idiosyncratic change *wɨ- > ə- (*w → ∅ and *ɨ → ə) are found in other person prefixes (12a–b).

(12) Akuriyó

Table 11: Verbs preserving $1S_A$ *w- in Proto-Tiriyoan (Meira 1999: 292, 294, 339, 1998: 112-115, 165)

	Proto-Tiriyoan	Tiriyó	Akuriyó
'say'	*wɨ-ka-	wɨ-ka-	wɨ-ka-
'go'	*w-ɨtəmɨ-	wɨ-tən-	ə-təmɨ-
'be-1'	*w-a-	w-a-	∅-a-
'be-2'	*w-eʔi-	w-ei-	?-eʔi-
'come'	*w-əʔepɨ-	w-əepɨ-	∅-eepɨ-

[6] The latter allomorph was subsequently replaced by k- in Akuriyó (Section 3.1.4).

[7] In fact, even the issue of *when* this happened is open. It could have happened at the Proto-Taranoan stage, but the subsequent introduction of j- in Carijona (Section 3.1.5) would have erased any traces of such an innovation. Accordingly, Meira (1998) hesitates to assign this extension to a specific proto-language. The decision taken here to assign the innovation to Proto-Tiriyoan does not affect the results of this study.

[8] The present reconstruction of 'to come' diverges from Meira's (1998: 114–115), who reconstructs Proto-Taranoan *əepɨ for first, but *eepɨ for other persons, based on the paradigmatic pattern in Tiriyó and the vowel length in Akuriyó. Akuriyó and Carijona would then have levelled that pattern, similar to what was suggested for the Pekodian languages (Section 3.1.1). Here, both Tiriyó əepɨ and Akuriyó eepɨ are identified as reflecting *ətepɨ (Section 3.2.4), via Proto-Taranoan *(ət͡ʃ-)epɨ and Proto-Tiriyoan *(əʔ)epɨ.

Table 12: Regular Akuriyó $1S_A$ markers (Gildea 1994: 77, 79, 82, 84-87)

first person *k-*	first person *t͡ʃ-*
əempa- 'to learn'	*epɨ-* 'to bathe (INTR)'
əət͡ʃena- 'to cry'	*ekɨrɨka-* 'to stay back'
əiwa- 'to tremble'	*entapo-* 'to yawn'
əməmɨ- 'to enter'	*etonema-* 'to lie down'
ətajiŋka- 'to run'	*ewai-* 'to sit down'
əturu- 'to talk'	*ehpa-* 'to bathe (INTR)'
əənɨkɨ- 'to sleep'	

257 a. *wi-toka*
1>3-hit
'I hit him/her.' (Gildea 1994: 86)

258 b. *kəʔ-eepɨ*
1+2-come
'We came.' (Meira 1998: 114)

259 3.1.4 Akuriyó *k-*

260 After the split of Proto-Tiriyoan, when **t-*/**t͡ʃ-* had largely replaced **w-*, Akuriyó innovated a third $1S_A$
261 marker *k-*. *k-* and *t͡ʃ-* show a clear phonologically conditioned distribution in Gildea's (1994) Akuriyó
262 data (Table 12). Meira (1998: 107) largely confirms that distribution, but mentions "several cases of
263 first person *t-* in Akuriyó" (on *ə*-initial verbs), albeit without any examples. He also suggests that *k-*
264 may be more recent, which is plausible: since the distribution **t-* / _*ə* / **t͡ʃ-* / _*e* is reconstructible to
265 Proto-Tiriyoan, the most straightforward scenario is *k-* replacing **t-* but not **t͡ʃ-* in Akuriyó. The few
266 *t-* mentioned by Meira (1998) were then perhaps reintroduced under Tiriyó influence. However, since
267 there are no examples of – or more information about – *ə*-initial verbs with *t-*, these cases cannot be
268 discussed further.
269 The verbs listed for Proto-Tiriyoan in Table 11 in Section 3.1.3 of course also resisted the extension
270 of *k-* in Akuriyó, although the first-person form of the copular verb *eʔi* is unknown. In addition, there is
271 an S_A verb *ɨ(h)tə* 'to go down', which has an irregular first person marker *p-*, apparently reconstructible
272 to Proto-Tiriyoan (13). It was not affected by the extension of Akuriyó *k-*, but whether it was an S_A verb
273 when Proto-Tiriyoan **t-* was introduced is unclear (see Section 3.2.5).

274 (13) First person forms of 'to go down'
275 Tiriyó *p-ɨhtə-* (Meira 1999: 294)
Akuriyó *p-ɨtə-* (Gildea 1994: 84)

276 3.1.5 Carijona *j-*

277 Carijona, the third Taranoan language, extended the $1S_P$ marker *j(i)-*[9] to S_A verbs (Meira 1998: 105–107).
278 Combined with the extension of $2S_A$ *m-* and $1+2S_A$ *kɨt-*/*kɨs-* to S_P verbs, this created a single unified S
279 category for regular verbs (Table 13). Although the split-S system was lost entirely, former S_A verbs can
280 be identified by the presence of a detransitivizer, like *ehɨnəhɨ* 'to fight' (14), derived from *hɨnəhɨ* 'to kill'
281 (Robayo Moreno 2000: 179).

282 (14) Carijona (Koch-Grünberg 1908: 79)
283 *hẹnẹ(x)tónoko-mãɻẹ y-ẹ-hẹnẹ́(x)yaɻ̣*
hɨnəhtono-ko=marə j-e-hɨnəhɨ-jai
enemy-PL=with 1-DETRZ-kill-NPST.CERT
'I fight with the enemies.'

[9] Since all affected S_A verbs are V-initial, only the / _V allomorph *j-* occurs in that context.

Table 13: Regular Carijona verbs (Robayo Moreno 2000: 173; Meira 1998: 106)

	'to dance'	'to arrive'
1	*j-eharaga-*	*ji-tuda-*
2	*m-eharaga-*	*mɨ-tuda-*
1+2	*kɨs-eharaga-*	*kɨsi-tuda-*
3	*n-eharaga-*	*ni-tuda-*

Table 14: Regular Yukpa verbs (Largo 2011: 72, 76; Meira 2006b: 139)

	'to wash self'	'to sleep'	'to fall'
1	*j-otum-*	*jɨ-nɨ-*	*j-ata-*
2	*m-otum-*	*mɨ-nɨ-*	*m-ata-*
3	*n-otum-*	*nɨ-nɨ-*	*n-ata-*

As noted in Section 3.1.3, this extension also erased any traces of putative Proto-Taranoan $1S_A$ **t-*. However, it did not fully eclipse the old $1S_A$ marker **w-*, which is attested as being preserved in the verbs *tə* 'to go' (15a) and *a* 'to be' (15b). In addition, the verb *ka* 'to say' has a zero-marked first-person form (15c).

(15) Carijona (Guerrero-Beltrán 2016: 5, 42, personal communication)

a. *wɨ-tə-e=rehe*
1-go-NPST=FRUST
'I almost go (but I am not going to go).'

b. *əji-marə-ne w-a-e*
2-COM-PL 1-be-NPST
'I am with you all.'

c. *dëmëmara kae ëwï iya*
n-təmə=mara ∅-ka-e əwɨ i-ja
3-go=DUB 1-say-NPST.CERT 1PRO 3-OBL
'"Did s/he leave?", I say to him.'

Based on other C-initial verbs like *tə* 'to go' or *tuda* 'to arrive', *ka* 'to say' should either have conservative *wɨ-* or innovative *ji-*, so the zero is unexpected. It is analysed here as a reflex of **wɨ-*, primarily due to the loss of **w* in Ikpeng and Hixkaryána. While those developments were more regular, an already irregular marker undergoing idiosyncratic phonological erosion is not that surprising, see Akuriyó **wɨ- > ə-* in the preceding section. Alternatively, the divergent development of **w-* on *ka* 'to say' and *tə* 'to go' may be due to the latter's originally V-initial nature (Section 3.2.2).

3.1.6 Yukpa *j-*

The divergent nature of Yukpa[10] is *inter alia* visible in the loss of many Set I forms and the formation of non-cognate innovative constructions (Meira 2006b). However, it does preserve the Set I prefixes in the immediate past, which shows a unified intransitive paradigm (Table 14). The wholesale loss of 1+2 as an inflectional value was combined with the extension of $2S_A$ *m(ɨ)-* to (now former) S_P verbs like *nɨ* 'to sleep'. These verbs share their first person marker *j(ɨ)-* with former S_A verbs like *otum* 'to wash self', identifiable by their semantics and the reflex of **əte/e-*. Since *j(ɨ)-* is the reflex of the Proto-Cariban $1S_{(P)}$

[10]Very little is known about its sister Japreria, whose status as a dialect of Yukpa is contested by Oquendo (2004).

marker **u(j)-* (Gildea 1998: 92), it also occurs on transitive verbs in 3>1 scenarios (16a). In contrast, 1>3 scenarios are zero-marked (16b).

(16) Yukpa (Meira 2006b: 139)

a. *aw j-esare*
1PRO 3>1-see
'S/he saw me.'

b. *aw ∅-esare*
1PRO 1>3-see
'I saw it.'

Since Proto-Cariban $1S_A$ **w(ɨ)-* was extended to 1>3 scenarios in most languages (Gildea 1998: 81–82), and given its inclination for phonological erosion (Sections 3.1.1 and 3.1.2), the zero marking in 1>3 scenarios can be identified as the Yukpa reflex of $1S_A$ **w-*.

In intransitive verbs, this first-person zero marking is attested with a single verb, *to* 'to go' (17). It diverges from regular C-initial verbs with *jɨ-*, like 'to sleep' (Table 14). It can thus be identified as having resisted the extension of *j-* in Yukpa.

(17) Yukpa (Meira 2006b: 139)
aw ∅-to
1PRO $1S_A$-go
'I went.'

3.2 Conservative verbs in comparison

In Section 3.1, six incomplete extensions of personal prefixes into $1S_A$ territory and the verbs unaffected by them were identified. This set is rather small in most cases, and many of the verbs remain conservative in different (proto-)languages. Here, these verbs are investigated from a comparative perspective: **ka[ti]* 'to say' (Section 3.2.1), **ɨtə[mə]* 'to go' (Section 3.2.2), both roots of the copula **eti* and **a[p]* (Section 3.2.3), **(ət)jəpɨ* 'to come' (Section 3.2.4), **ɨpɨtə* 'to go down' (Section 3.2.5), and **e-pɨ* 'to bathe' (Section 3.2.6). The large and phonologically coherent group of *e*-initial verbs not affected by the extension of Akuriyó *k-* (Section 3.1.4) will not be discussed.

3.2.1 **ka[ti]* 'to say'

This verb was not affected by any of the extensions in Proto-Pekodian, Proto-Waiwaian, Proto-Tiriyoan, Akuriyó, or Carijona (Sections 3.1.1 to 3.1.5), while no first person form of its Yukpa reflex *ka* is attested. Most reflexes are simply *ka*, but a fleeting syllable **ti* is reconstructed by Gildea & D. Payne (2007), visible in the imperative forms of some languages. Table 15 shows a comparison of the longest attested forms for each language.[11]

As mentioned in Section 3.1.2, Derbyshire (1985) analyzes Hixkaryána *ka[s]* as transitive, a choice not only motivated by a desire to avoid an idiosyncratic intransitive first person prefix *ɨ-* instead of regular *kɨ-*. The verb also shows the complementary distribution of third person *n-* and preceding objects typical of transitive verbs in Cariban (Gildea 1998: 60–81). Due to its semantics, these objects are either ideophones or direct speech (18).

[11] Cognate segments in Tables 15 to 21 were aligned automatically with LingPy (List et al. 2021), for improved exposition of correspondences. Brackets indicate segments not present in all forms.

Table 15: Reflexes of **ka[ti]* 'to say' (Meira 2005: 267, 2003a: 4; Franchetto 2008: 48; Pachêco 2001: 209, 279; Alves 2017: 80, 153; Derbyshire 1985: 182; Meira 1998: 113; E. Koehn & S. Koehn 1986: 107; R. E. Hawkins 1998: 26; Camargo & Tapinkili 2010: 66; Abbott 1991: 59; Swiggers 2010: 123; Courtz 2008: 430; Caesar-Fox 2003: 125; M.-C. Mattéi-Müller 1994: 102; Largo 2011: 63; p.c., Spike Gildea)

Language	Form				
Werikyana	*ka[s]*	k	a	s	
Proto-Waiwaian	**ka[s]*	k	a	s	
Hixkaryána	*ka[s]*	k	a	s	
Waiwai	*ka[s]*	k	a	s	
Proto-Pekodian	**ge*	g	e		
Proto-Pekodian	**ke*	k	e		
Arara	*ge*	g	e		
Arara	*ke*	k	e		
Ikpeng	*ge*	g	e		
Ikpeng	*k[e]*	k	e		
Bakairi	*ge*	g	e		
Bakairi	*ke*	k	e		
Proto-Tiriyoan	**ka*	k	a		
Tiriyó	*ka*	k	a		
Akuriyó	*ka*	k	a		
Carijona	*ka*	k	a		
Wayana	*ka[i]*	k	a		i
Apalaí	*ka[ʃi]*	k	a	ʃ	i
Kari'ña	*ka*	k	a		
Kapón	*ka*	k	a		
Pemón	*ka*	k	a		
Macushi	*ka*	k	a		
Panare	*ka[h]*	k	a	h	
Upper Xingu Carib	*ki*	k	i		
Yukpa	*ka*	k	a		

(18) Hixkaryána

a. *oni wyaro nkekoni biryekomo, tiyoni wya*
oni wjaro n-ka-jakoni bir^{j}ekomo ti-joni wja
this like 3-say-REM.CONT boy COR-mother OBL
'This is what the boy said to his mother.' (Derbyshire 1985: 36)

b. *moro ha, ketxkoná hatá.*
moro ha ka-jaʃʃkoni hati
MED.DEM.INAN INTS say-REM.CONT.PL HSY
'"That one there" they said.' (Derbyshire 1965: 14)

In (18a), the prefix *n-* occurs because there is no preceding object ('he said it like this'), while does not occur in (18b) where the verb is preceded by direct speech. Outside of Hixkaryána, at least the Tiriyó reflex shows the same pattern, albeit inconsistently so (Carlin 2004: 267).

Reflexes of **ka[ti]* 'to say' also show transitive patterns in their derivational suffixes: In Tiriyó, it is the only intransitive verb (inconsistently) taking transitive *-po* (CAUS) and *-ne* (AGT.NMLZ) (Meira 1999: 263, 169). It also exceptionally takes the former suffix in Kari'ña (Courtz 2008: 82) and Wayana (Tavares 2005: 258). The agent nominalizer **-ne* became the Panare inflectional suffix *-ɲe* on transitive verbs (Gildea 1998: 184–185). The combination of *ka* and *-ɲe* likely led T. E. Payne & D. L. Payne (2013: 214) to categorize it as transitive, disagreeing with M.-C. Mattéi-Müller (1994: 102). Finally, reflexes of the transitive causativizer **-metipo* (Gildea 2015) are found with *ka* in Apalaí (E. Koehn & S. Koehn 1986: 51) and Waiwai (R. E. Hawkins 1998: 52).

Arguments in favor of intransitive 'to say' primarily come from its inflectional prefixes. Kari'ña has a minimal pair between transitive *ka* 'to remove' and intransitive *ka* 'to say', *sikai* 'I took it away' vs *wikai* 'I said' (Courtz 2008: 288, 45). Similarly, Proto-Pekodian **ke* 'to say' had 1S_A **w-* (Section 3.1.1), rather than 1>3 *s-* (Bakairi) or **ini-* (Proto-Xinguan). Additionally, languages differentiating transitive from S_A prefixes by the presence of *i* (Meira et al. 2010: 495) have *i*-final prefixes, see Akuriyó in (19a), as well as Meira (1999: 294), Tavares (2005: 195), Pachêco (2001: 288), Alves (2017: 150), and Hoff (1968: 168) for cognate forms in other such languages. Finally, the S_A class marker *w-* occurs on nominalizations in Kari'ña (19b), and is probably reflected as vowel length in the Tiriyó (Meira 1999: 333) and Wayana (Tavares 2005: 196) participles.

(19) a. Akuriyó (Meira 1998: 113)
mi-ka
2-say
'You said.'

b. Kari'ña (Courtz 2008: 202)
Òmakon 'wa oti ywykàpo kaiko.
o-ʔma-kon ʔwa oti i-wi-ka-ʔpo kai-ko
2-child-PL OBL greeting 1-S_A-say-PST.NMLZ say-IMP
'Pass my greetings to your children.'

In summary, this verb can be reconstructed as intransitive based on its (inflectional) prefixes, but transitive based on some (derivational) suffixes. Hixkaryána has lost the main intransitive criteria, making its reflex look more like a transitive verb.

3.2.2 **iti[mə]* 'to go'

This verb was not affected by the any of the extensions in Section 3.1. Gildea & D. Payne (2007) reconstruct it as C-initial **tə[mə]*, like **ka[ti]* 'to say' with a fleeting second syllable. While many reflexes are unambiguously *t*-initial (e.g. Hixkaryána *ntoje* 'he went' [Derbyshire 1985: 27], Tiriyó *təkə* 'go!' [Meira 1999: 246], or Wayana *kuptəm* 'we went' [Tavares 2005: 195]), the contrast with C-initial **ka[ti]* becomes

Table 16: Reflexes of **ɨtə[mə]* 'to go' (Cruz 2005: 291; Meira 1999: 292; Tavares 2005: 195; Gildea 1994: 87; Alves 2017: 80, 153; Derbyshire 1985: 27, 248; R. E. Hawkins 1998: 45, 62; Pachêco 2001: 54, 80; von den Steinen 1892: 112, 374; Cáceres 2011: 181, 216; Meira 1998: 112; Hoff 1968: 168; Meira 2006b: 139; Cáceres & Gildea 2018: 4; M. C. Mattéi-Müller 1975: 74; M.-C. Mattéi-Müller 1994: 198; Abbott 1991: 48, 50; García Ferrer 2006: 172; Franchetto 2002: 6; Camargo 2002: 99; Lacerda 1991: 14; p.c., Spike Gildea)

Language	Form					
Werikyana	*to[mo]*		t	o	m	o
Proto-Waiwaian	**[ɨ]to[m]*	ɨ	t	o	m	
Hixkaryána	*[ɨ]to*	ɨ	t	o		
Waiwai	*[e]to[m]*	e	t	o	m	
Proto-Pekodian	**ɨtə[n]*	ɨ	t	ə	n	
Arara	*to*		t	o		
Arara	*ɨdo*	ɨ	d	o		
Ikpeng	*aran*	a	r	a	n	
Ikpeng	*ero*	e	r	o		
Bakairi	*[ɨ]tə*	ɨ	t	ə		
Proto-Tiriyoan	**[ə]tə[mɨ]*	ə	t	ə	m	ɨ
Tiriyó	*tə[n]*		t	ə	n	
Akuriyó	*[ə]tə[mɨ]*	ə	t	ə	m	ɨ
Carijona	*tə[mə]*		t	ə	m	ə
Wayana	*[ɨ]tə[m]*	ɨ	t	ə	m	
Apalaí	*ɨto*	ɨ	t	o		
Kari'ña	*[ɨ]ʔ*	ɨ	ʔ			
Kari'ña	*to*		t	o		
Ye'kwana	*ɨtə[mə]*	ɨ	t	ə	m	ə
Ingarikó	*ətə*	ə	t	ə		
Pemón	*[e]tə*	e	t	ə		
Macushi	*[a]tɨ*	a	t	ɨ		
Panare	*tə[n]*		t	ə	n	
Yawarana	*tə*		t	ə		
Mapoyo	*tə*		t	ə		
Upper Xingu Carib	*[e]te*	e	t	e		
Yukpa	*to*		t	o		
Waimiri-Atroari	*ɨʔ*	ɨ	ʔ			

clear once all forms are considered (Table 16). *ɨ* is predominant and can tentatively be reconstructed, although languages throughout the family reflect **ə*. Since many languages only show *ɨ* in certain contexts, even forms like Tiriyó *wɨtənne* 'I went' (Meira 1999: 43) are ambiguous, since epenthetic *ɨ* breaks up CC clusters on prefix-verb boundaries.

3.2.3 **eti* and **a[p]* 'to be'

**a[p]* is the older copula and already had various irregularities in Proto-Cariban (Gildea 2018). **eti* is reconstructed by Meira & Gildea (2009) and Gildea (2018) as originally meaning 'to dwell, live', but serving as a copula already in Proto-Cariban.[12] Reflexes of these roots are used suppletively, conditioned by person and/or TAM. Both roots preserved $1S_A$ **w-* in Proto-Pekodian, Proto-Waiwaian, and Proto-Tiriyoan (Sections 3.1.1 to 3.1.3). Akuriyó *a* was not targeted by the extension of *k-* (Section 3.1.4), while no first person form of *e?i* is attested. Carijona innovated *j-*, but only in the reflex of **eti* (20); the *a*

[12]Such a stative, locative source is also suggested by the existence of *ɨ̄ʃ̄ɨ* 'to lie down' in Arara (Alves 2017: 196).

root preserves *w-* (Section 3.1.5). Yukpa introduced *j-* to the reflexes of both **a[p]* and **eti*, which are preserved as encliticized auxiliaries (21).

(20) Carijona (Robayo Moreno 1989: 177)
iretibə ɛ͡ʧinəme gərə jɛ͡ʧiɨ
ireti-bə ɛ͡ʧi-nə=me gərə j-ɛ͡ʧi-ɨ
then-from be-INF=ATTRZ still 1-be-PFV
'Then I was already grown up.'

(21) Yukpa (Meira 2006b: 143–144)

	NPST	PST
1	*=j-a(-s)*	*=j-e*
2	*=mak(o)*	*=m-e*
3	*=mak(o)*	*=n-e*

A comprehensive comparative overview of both roots is given by Gildea (2018: 375–382); they will not be discussed in detail here.

3.2.4 **(ət-)epɨ* 'to come'

Innovative **k-* was introduced on the Ikpeng and Bakairi reflexes of **ət-epɨ*, but not on the Arara reflex of **epɨ* (Section 3.1.1). The reflex of **ət-epɨ* resisted the introduction of Proto-Tiriyoan **t-* (Section 3.1.3). Carijona *ehɨ* shows innovative *j-*, rather than conservative *w-* (22). No Yukpa reflex of this verb is attested, and it was replaced in Proto-Waiwaian by **omokɨ* 'to come'.

(22) Carijona (Guerrero Beltrán 2019: 102)
əji-wa-e j-eh-ɨ
2-search-SUP 1-come-PFV
'I came looking for you.'

Gildea & D. Payne's (2007) reconstruction **ətepɨ* can be segmented into an optional prefix **ət-* and a root **epɨ*, since reflexes can be grouped into those with a reflex of **ət-* and those without (Table 17). As seen in Table 18, this division can exist within a single paradigm.

Long **ət-epɨ* lost **t* in Tiriyó, Apalaí, and Bakairi, but only the first development is due to regular sound changes (Meira 1998: 31–32). Akuriyó reflects **əte* as *ee*, where *eepɨ* cannot be a reflex of **epɨ* due the otherwise unexpected vowel length (Meira 1998: 114–115). Other languages coalesced **əte* to **əə*, with **əəpɨ* being reflected in Upper Xingu Carib (Gildea 2012: 452), Kari'ña, Werikyana, Panare, and Yawarana. **e*-initial transitive verbs detransitivized with **ət-* underwent the same development in Kari'ña, creating minimal pairs for length (Meira et al. 2010: 509–510), while **ət-e* > **əə* is irregular at least in Werikyana and Panare.

The Pemongan languages and Werikyana point to **jəpɨ* rather than **epɨ*, although Werikyana *johɨ* is very rare in contrast to more frequent *o(o)hɨ*. It only occurs in the third person form of the Progressive, meaning that the *j* may be a reflex of third person **i-*. However, regular *o*-initial Werikyana verbs have no third person prefix (Table 19), while *i-* occurs with C-initial *to[mo]* 'to go', suggesting that *j* is indeed part of the root. Putative Proto-Cariban **jəpɨ* could yield **epɨ* via the two widespread sound changes **ə* → **e* / **j*_ and **j* → ∅ / #_ (Meira et al. 2010). The argument against this reconstruction is that **ə* → **e* / **j*_ did not happen in Proto-Parukotoan, and only inconsistently in Proto-Panare-Pemongan and Mapoyo-Yawarana (Meira et al. 2010: 501–502; Gildea et al. 2010). Thus, it does not explain Werikyana *ehɨ*, nor the total absence of the sequence **jə* in Panare and Mapoyo-Yawarana.

The clear segmentability of **ət-* in combination with its form suggest that it is a detransitivizing prefix. Although the combination of a detransitivizer and an intransitive verb makes no sense semantically, some historical S_P verbs are attested as adding the detransitivizer to become S_A verbs. For example,

Table 17: Reflexes of **(ət-)epɨ* ‘to come’ (Abbott 1991: 32; Álvarez 2000: 102; Caesar-Fox 2003: 125; Cruz 2005: 299, 415; Cáceres 2011: 438; Robayo Moreno 2000: 178; Meira 1998: 168; M. C. Mattéi-Müller 1975: 74; Meira 1999: 294; Alves 2017: 113, 150; E. Koehn & S. Koehn 1986: 37; Pachêco 2001: 265; Stegeman & Hunter 2014: 160; Meira 2003a: 4; T. E. Payne & D. L. Payne 2013: 65, 451; Méndez-Arocha 1959: 68; Cáceres Arandia et al. 2017; Hoff 1968: 420, 437; Meira & Franchetto 2005: 182; Franchetto 1986: 42; p.c., Spike Gildea)

Language	Form							
Werikyana	*ehɨ*					e	h	ɨ
Werikyana	*johɨ*				j	o	h	ɨ
Werikyana	*o[o]hɨ*	oo		-			h	ɨ
Proto-Pekodian	**epɨ*					e	p	ɨ
Proto-Pekodian	**ədepɨ*	ə	d	-		e	p	ɨ
Arara	*ebɨ*					e	b	ɨ
Arara	*odebɨ*	o	d	-		e	b	ɨ
Ikpeng	*arep*	a	r	-		e	p	
Bakairi	*əewɨ*	ə		-		e	w	ɨ
Proto-Tiriyoan	**epɨ*					e	p	ɨ
Proto-Tiriyoan	**əʔepɨ*	ə	ʔ	-		e	p	ɨ
Tiriyó	*epɨ*					e	p	ɨ
Tiriyó	*əe[pɨ]*	ə		-		e	p	ɨ
Akuriyó	*eepɨ*	ee		-			p	ɨ
Carijona	*eh[ɨ]*					e	h	ɨ
Apalaí	*oepɨ*	o		-		e	p	ɨ
Kari’ña	*o[o]pɨ*	oo		-			p	ɨ
Ye’kwana	*ehə*					e	h	ə
Akawaio	*jepɨ*				j	e	p	ɨ
Akawaio	*əsipɨ*	ə	s	-		i	p	ɨ
Ingarikó	*jepə*				j	e	p	ə
Ingarikó	*jə*				j	ə		
Patamona	*jepɨ*				j	e	p	ɨ
Patamona	*jəpɨ*				j	ə	p	ɨ
Pemón	*jepɨ*				j	e	p	ɨ
Macushi	*ipɨ*					i	p	ɨ
Panare	*ə[ə]pɨ*	əə		-			p	ɨ
Yawarana	*epɨ*					e	p	ɨ
Yawarana	*əəpɨ*	əə		-			p	ɨ
Mapoyo	*epɨ*					e	p	ɨ
Upper Xingu Carib	*e[e]*	ee						
Upper Xingu Carib	*i*							i

Table 18: **(ət-)epɨ* ‘to come’ in paradigms (Alves 2017: 113, 150, 153, 156; Meira 1999: 294; p.c., Spike Gildea)

	Werikyana	Arara	Tiriyó
1	*w-oohɨ-*	*w-ebɨ-*	*w-əepɨ-*
2	*m-oohɨ-*	*m-odebɨ-*	*mən-epɨ*
1+2	*kɨs-ohɨ-*	*kud-ebɨ-*	*ke-epɨ*
3	*n-ehɨ-*	*t-ebɨ-*	*n-epɨ*

Table 19: Werikyana S_A verbs in the Progressive (Spike Gildea, p.c.)

	'to come'	'to dream'	'to go'
1	*∅-w-oohɨ-*	*∅-w-osone-*	*∅-wɨ-to-*
2	*o-w-ohɨ-*	*o-w-osone-*	*o-w-to-*
1+2	*ku-w-ohɨ-*	*ku-w-osone-*	*kɨ-w-to-*
3	*∅-johɨ-*	*∅-osone-*	*i-to-*

the Proto-Cariban S_P verb **wɨnɨkɨ* 'to sleep' becomes Tiriyó *əənɨkɨ* (Meira 1999: 252) and Kari'ña *əʔnɨkɨ* (Courtz 2008: 429), both S_A. Waiwai 'to sleep' can be *wɨnɨk* (R. E. Hawkins 1998: 30) or *et-wɨnɨk* (W. N. Hawkins & R. E. Hawkins 1953: 204). However, unlike 'to sleep', bare **epɨ* 'to come' apparently was an S_A verb already (although its reflexes in languages with split-S mostly co-occur with reflexes of **ət-epɨ*).

3.2.5 **ɨpɨtə* 'to go down'

Reflexes of this verb were not affected by the extensions of **k-* in Proto-Pekodian (Section 3.1.1) and *k-* in Akuriyó (Section 3.1.4). Its resistance against the former extension was later broken in Bakairi, while its fate in Ikpeng is unknown. When Akuriyó extended *k-*, the verb already had a first person form irregularly inflected with *p-*, inherited from Proto-Tiriyoan. At first sight, it was also affected by the extensions of *j-* in Carijona (23a) and Yukpa (23b).

(23) a. Carijona (David Felipe Guerrero, p.c.)
irə wat͡ʃinakano tae j-ehɨtə-e
INAN.ANA body.of.water along.bounded 1-go.down-NPST
'…I go down through that guachinacán.'

b. Yukpa (Meira 2003b)
aw yéwtu
aw j-ewuhtu
1PRO 1-go.down
'I went down.'

However, a family-wide perspective reveals a more complicated story (Table 20).[13] While a verb **ɨpɨtə* can be reconstructed to Proto-Cariban, different (proto-)languages do not agree about its class. Its reflexes in languages that preserve the split-S system are distributed fairly evenly between S_A and S_P.

The verb shows traits of both classes in Wayana, making it a "mixed" verb synchronically. It takes the first and second person S_P markers *j-* and *əw-* (Tavares 2005: 200), but the $1+2S_A$ marker *kut-* (Tavares 2005: 206). It also shows the S_A class marker *w-* in nominalizations (24a), but behaves like an S_P verb in taking a second person prefix in imperatives (24b).

(24) Wayana (Tavares 2005: 200)

a. *ïwïptëë*
ɨ-w-ɨptə-rɨ
1-S_A-go.down-NMLZ
'my going down'

b. *əw-ɨptə-k*
2-go.down-IMP
'Go down!'

[13] The cognacy status of parenthesized forms in Table 20 is uncertain. The reconstruction of Proto-Pekodian **ɨptə* treats the additional elements in daughter languages as non-cognate. Meira & Franchetto (2005) identify no correspondence between Bakairi *gɨ* and Ikpeng *ŋ*, and at least the addition of a final *ŋ* in Proto-Xinguan is attested elsewhere: a) Proto-Cariban **ənɛ* 'to see', Arara and Ikpeng *eneŋ*; b) Proto-Cariban **ətə* 'to hear', Arara *taŋ*, Ikpeng *iraŋ*; and c) Proto-Cariban **ənə* 'to eat meat', Arara *oŋoŋ* 'to bite' (Gildea & D. Payne 2007: 8; Alves 2017: 56, 144, 57; Pachêco 2001: 25, 270).

Table 20: Reflexes of **ɨpɨtə* ‘to go down’ (Meira 2003b; Derbyshire 1979: 196; R. E. Hawkins 1998: 55; Guerrero Beltrán 2019: 118; Camargo & Tapinkili 2010: 44; Camargo 2002: 99; Courtz 2008: 263; Cáceres 2011: 450; Stegeman & Hunter 2014: 139; Álvarez 2008: 139; Abbott 1991: 34; M.-C. Mattéi-Müller 1994: 88; Méndez-Arocha 1959: 68; Bruno 1996: 58; Gildea 1994: 84; Alves 2017: 153; von den Steinen 1892: 137; Meira 1998: 116; p.c., Angela Fabíola Alves Chagas, Spike Gildea)

Language	Form	Class														
Proto-Parukotoan	**ɨɸɨto*	S_P					ɨ	ɸ	ɨ	t	o					
Werikyana	*ɨhɨto*	S_P					ɨ	h	ɨ	t	o					
Proto-Waiwaian	**hto*	?						h		t	o					
Hixkaryána	*hto*	?						h		t	o					
Waiwai	*hto*	–						h		t	o					
Proto-Pekodian	**ɨptə*	S_A					ɨ	p		t	ə					
Arara	*ɨptoŋ*	S_A					i	p		t	o	-	ŋ			
Ikpeng	*ɨptoŋ*	?					i	p		t	o	-	ŋ			
Bakairi	*ɨtəgɨ*	S_A					ɨ			t	ə	-		g	ɨ	
Proto-Taranoan	**ɨpɨtə*	?					ɨ	p	ɨ	t	ə					
Proto-Tiriyoan	**ɨhtə*	S_A					ɨ	h		t	ə					
Tiriyó	*ɨhtə*	S_A					ɨ	h		t	ə					
Akuriyó	*ɨ[h]tə*	S_A					ɨ	h		t	ə					
Carijona	*ehɨtə*	–					e	h	ɨ	t	ə					
Wayana	*ɨptə*	S_A / S_P					ɨ	p		t	ə					
Apalaí	*ɨhto*	S_P					ɨ	h		t	o					
Kari’ña	*onɨʔto*	(S_A)	o	-	n	-	ɨ	ʔ		t	o					
Ye’kwana	*əʔtə*	S_P					ə	ʔ		t	ə					
Kapón	*(uʔtə)*	–														
Pemón	*(uʔtə)*	–														
Macushi	*(autɨ)*	–														
Panare	*əhtə*	S_A					ə	h		t	ə					
Yawarana	*əhtə*	–					ə	h		t	ə					
Yukpa	*(ew[uh]tu)*	–														
Waimiri-Atroari	*ɨtɨ*	–							ɨ	t	ɨ					

Its causativized form is *ɨptə-ka* (Tavares 2005: 255); the Proto-Cariban causativizer **-ka* was restricted to S_P verbs (Gildea & Cáceres in preparation). These patterns point to 'to go down' being a regular S_P verb in pre-Wayana, but partially switching to the S_A class by taking a $1{+}2S_A$ prefix and the S_A class marker. This in turn implies that (inflectionally defined) S_A reflexes in other languages fully switched from S_P.

Wayana-external evidence supports this hypothesis: The Arara causativized form is *eniptoŋ* (Alves 2017: 66), and Kari'ña has a cognate form *enɨʔto* (Courtz 2008: 263); *onɨʔto* 'to go down' in Table 20 is a detransitivized form thereof, lit. 'to get oneself down'. These forms reflect the transitivizer **en-*, occurring with S_P verbs in Proto-Cariban (Gildea & Cáceres in preparation). Tiriyó *ɨhtə* has irregular causativized forms, also with a reflex of **en-* (Meira 1999: 263). In conclusion, **ɨpɨtə* 'to go down' was an S_P verb in Proto-Cariban, but for unknown reasons switched classes in four "and a half" languages of the family.

This makes it impossible to tell whether it was affected by most extensions under discussion: For Proto-Tiriyoan, one cannot establish a relative chronology of the class switch, the introduction of idiosyncratic $1S_A$ **p-*, and the extension of **t-*. Its first person form and its inflectional class in Proto-Waiwaian are unknown. For Carijona and Yukpa, one cannot rule out a verb class switch before the breakdown of the split-S system. While no language-internal evidence supports such a switch, 'to go down' is clearly inclined to do so; Carijona may even have inherited it as S_A from Proto-Taranoan. In all four cases, the verb could have had S_A or S_P status at the time of the extension, so it is unknown whether it even was a potential target.

On the other hand, a class switch is reconstructible to Proto-Tiriyoan, so it was an S_A verb when Akuriyó introduced *k-*. Likewise, the class switch most likely took place before the extension of Proto-Pekodian **k-*. Otherwise, the newly-turned-S_A verb would have taken on conservative and lexically heavily restricted **w-*, either in Proto-Pekodian, Proto-Xinguan, or Arara.

3.2.6 **e-pɨ* 'to bathe'

This verb resisted the extensions of Proto-Pekodian **k-* (Section 3.1.1) and Akuriyó *k-* (Section 3.1.4). It took on new $1S_A$ prefixes in Proto-Tiriyoan (Tiriyó *s-epɨ-*, Akuriyó *ʧ-epɨ-* [Meira 1999: 292; Gildea 1994: 87]) and Proto-Waiwaian (Hixkaryána *k-ewehɨ-*, Waiwai *k-ejeɸu-* [Derbyshire 1985: 195; R. E. Hawkins 1998: 166]). The first person form of its Carijona reflex *ehɨ* (Koch-Grünberg 1908: 72) is unknown; an unattested Yukpa cognate may exist.

Verbs for intransitive 'to bathe' are typical S_A verbs in most Cariban languages, derived with a detransitivizer from a transitive root. These roots are reflexes of **[ɨ]pɨ*, or **kupi* in some Venezuelan languages (Table 21). Proto-Pekodian can be reconstructed as having the pair **ipɨ* (INTR) / **ɨpɨ* (TR) (Section 3.1.1). Thus, while Proto-Pekodian 'to bathe (TR)' has perfectly regular cognates in other languages of the family, 'to bathe (INTR)' changed **e-* to **i*. This is an irregular development, since there are no attested reflexes of a Pekodian detransitivizer **i-* (Meira et al. 2010: 506); its cause is unknown.

3.3 Summary

In Section 3.2, the verbs which were unaffected by the extensions in Section 3.1 were reconstructed, and affected reflexes in the languages under discussion were identified. Table 22 gives an overview of what verbs were affected by which extensions (except for *e-*initial Akuriyó verbs unaffected by the extension of *k-*, as they are a large and predictable group). In some cases, the verb does not occur, or at least not in a first person form (–), in others that form is unknown (?), and the question of affectedness is often not meaningfully answerable (N/A) for 'to go down'. Every ✓ stands for a verb affected by an extension, while × represents conservatively inflected verbs, making clear how strongly these verbs tend to resist person marker extensions in different languages. Section 4 will explore explanations for the fact that the same 1-7 verbs retained their old $1S_A$ marker in 6 independent developments, while a plethora of regular $S_{(A)}$ verbs took on new markers.

Table 21: Comparison of intransitive and transitive ‘to bathe’ (Derbyshire 1979: 198; R. E. Hawkins 1998: 192, 203; Alves 2017: 58, 150; Pachêco 1997: 103; Campetela 1997: 123; Meira 2003a: 4, 2005: 285, 1999: 697; Gildea 1994: 87; Camargo & Tapinkili 2010: 24, 52; Meira 2000: 218; Courtz 2008: 304; Cáceres 2011: 439, 454; Stegeman & Hunter 2014: 37; de Armellada 1944: 34, 129; M.-C. Mattéi-Müller 1994: 8, 294; p.c., Spike Gildea)

(a) Reflexes of **e-pɨ* ‘to bathe (INTR)’

Language	Form							
Werikyana	*eehɨ*	ee			-	h	ɨ
Hixkaryána	*ewehɨ*	e	w	e	-	h	ɨ
Waiwai	*ejeɸu*	e	j	e	-	ɸ	u
Arara	*ibɨ*	i			-	b	ɨ
Ikpeng	*ip*	i			-	p	
Bakairi	*i*	i					
Tiriyó	*epɨ*	e			-	p	ɨ
Akuriyó	*epɨ*	e			-	p	ɨ
Wayana	*epɨ*	e			-	p	ɨ
Apalaí	*epɨ*	e			-	p	ɨ

(b) Reflexes of **e-kupi* ‘to bathe (INTR)’

Language	Form							
Kari’ña	*ekupi*	e	-	k	u		p	i
Ye’kwana	*eʔhi*	e	-	ʔ			h	i
Kapón	*ekuʔpi*	e	-	k	u	ʔ	p	i
Pemón	*ekupɨ*	e	-	k	u		p	ɨ

(c) Reflexes of **ə-kupi* ‘to bathe (INTR)’

Language	Form						
Panare	*akupɨ*	a	-	k	u	p	ɨ

(d) Reflexes of **[ɨ]pɨ* ‘to bathe (TR)’

Language	Form			
Werikyana	*ɨhɨ*	ɨ	h	ɨ
Hixkaryána	*ɨhɨ*	ɨ	h	ɨ
Waiwai	*pɨ*		p	ɨ
Arara	*ɨbɨ*	ɨ	b	ɨ
Ikpeng	*ɨp*	ɨ	p	
Bakairi	*ɨ*			ɨ
Tiriyó	*pɨ*		p	ɨ
Akuriyó	*pɨ*		p	ɨ
Wayana	*upɨ*	u	p	ɨ
Apalaí	*pɨ*		p	ɨ
Ye’kwana	*ɨhɨ*	ɨ	h	ɨ
Pemón	*pɨ*		p	ɨ
Panare	*ɨpɨ*	ɨ	p	ɨ

(e) Reflexes of **kupi* ‘to bathe (TR)’

Language	Form					
Kari’ña	*kupi*	k	u		p	i
Kapón	*kuʔpi*	k	u	ʔ	p	i
Panare	*kupɨ*	k	u		p	ɨ

Table 22: Overview of extensions and (un-)affected verbs

	**ka[ti]* ‘say’	**ɨtə[mə]* ‘go’	**a[p]* ‘be-1’	**eti* ‘be-2’	**(ət-)epɨ* ‘come’	**ɨpɨtə* ‘go down’	**e-pɨ* ‘bathe’
Proto-Waiwaian **k-*	×	×	×	×	–	N/A	✓
Hixkaryána	×	×	×	×	–	N/A	✓
Waiwai	×	(✓)	×	×	–	N/A	✓
Proto-Pekodian **k-*	×	×	×	×	×	×	×
Arara	×	×	×	×	×	×	×
Ikpeng	×	✓	–	×	✓	?	×
Bakairi	×	×	×	×	✓	✓	×
Proto-Tiriyoan **t-*	×	×	×	×	×	N/A	✓
Tiriyó	×	×	×	×	×	N/A	✓
Akuriyó	×	×	×	?	×	N/A	✓
Akuriyó *k-*	×	×	×	?	×	×	×
Carijona *j-*	×	×	×	✓	✓	N/A	?
Yukpa *j-*	?	×	✓	✓	–	N/A	–

✓ affected; × not affected; ? unknown first person prefix; – does not occur; (✓) old and new marker combined; N/A not meaningfully answerable

4 Explaining conservativeness: a network morphology approach

A well-known approach to conservativeness, innovativeness, and (ir-)regularity in the lexicon is Bybee's (1985) network model of morphology, which seems well-suited for the data at hand. It aims to "account for cross-linguistic, diachronic and acquisition patterns in complex morphological systems" (Bybee 1995: 428). It does so by modeling shared morphological properties such as inflectional patterns as emerging from connections of differing strength between lexemes. A well-known example is a network of "strong" English verbs with *strɪŋ–strʌŋ* at the center and pairs like *rɪŋ–rʌŋ*, *spɪn–spʌn*, or *stɪk–stʌk* at its periphery. This network created new strong verbs in some dialects, like *sniːk–snʌk* or *brɪŋ–brʌŋ* (Bybee 1985: 129–130). These verbs formed a lexical connection with prototypical members of the group, developing new past tense forms as a result.

These verbs are an example for a phonologically motivated network; other possible bases of connections between lexemes are semantic or morphological similarity (Bybee 1985: 118). Another important factor in the model is frequency, since more frequent words have higher lexical strength (Bybee 1985: 119). This diminishes the influence from other lexemes, meaning that high-frequency items are more likely to resist innovations. For the Cariban first person patterns, the model predicts that semantically/phonologically/morphologically similar verbs are affected by person marker extensions, while high-frequency verbs resist extensions and thus remain conservative.

When considering the S_A verbs with innovative first person markers (those not in Table 22), several salient factors emerge as potential bases of lexical networks. The most obvious similarity is that they all have a reflex of the detransitivizer **ətə/e-* (see e.g. Meira [1998: 112] for Taranoan), a hallmark of S_A verbs (Section 2.3). Due to **ətə/e-* being prefixes, all derived S_A verbs begin with reflexes of **ə* or **e* (see e.g. Alves [2017: 153] for Arara), making phonologically based networks a second possibility. An inflectional commonality of innovative verbs is their S_A (sub-)class membership, represented by pre-extension $1S_A$ **w-* (*t-* in Akuriyó). There are no obvious semantic patterns, which is unsurprising considering the absence of a semantic basis in the split-S system (Section 2.3). Thus, there are three hypotheses for factors connecting members of the networks in which extensions spread: a reflex of **ətə/e-*, the stem-initial phoneme, or a specific $1S_A$ prefix.

It is intuitively obvious that many of the verbs in Table 22 are highly frequent, which would cause conservativeness according to the network model. Going beyond intuition is difficult due to the lack of frequency counts for individual lexemes for any Cariban language. The only statement in the literature is Courtz's (2008: 75) claim of Kari'ña underived S_A verbs being the most frequent ones: "It is difficult […] to imagine an intransitive or transitive origin for some of the most frequent middle verbs". This claim is neither supported by frequency counts nor accompanied by a list of verbs, although those verbs likely correspond to the five first columns of Table 22. To improve the data situation, a count of S_A verbs in three Apalaí texts from E. Koehn & S. Koehn (1994) will serve as a second source of frequency information (Table 23). The count data support the above interpretation of Courtz's (2008) claim, since defining "high frequency" as above average yields the exact same five verbs. It is questionable if the interpretation of Courtz's (2008) claim and this small Apalaí sample are representative of discourse patterns in the Cariban (proto-)languages discussed here, but their use is made necessary by the absence of alternatives.

Each of the three hypotheses for possible network factors can be combined with frequency: members of the lexical network connected by a factor are prone to innovation, but high-frequency verbs are exempt. The resulting six hypotheses were tested by predicting the expected behavior of each verb in each extension, illustrated in Table 31 for Proto-Tiriyoan. For example, **eʔi* 'to be' is expected to participate (✓) in innovations spreading in a phonologically defined (**e*-initial) or inflectionally defined ($1S_A$ **w-*) network. On the other hand, a network based on the detransitivizer predicts it to not undergo innovation (×). If frequency is taken into account, it is expected to remain conservative regardless of the basis of the network.

These predictions were checked against the data in Table 22 by tallying verbs with (in/)correctly predicted behavior. The resulting scores are illustrated for the extension of Proto-Tiriyoan **t-* in Table 32 and summed up for all extensions in Table 26. These scores crucially only consider the seven verbs in

Table 23: Frequency counts of S_A verbs in three Apalaí texts from E. Koehn & S. Koehn (1994) (163 S_A verbs, 1070 words)

Verb	Count	% S_A verb tokens
a 'be-1'	49	30.06%
eʃi 'be-2'	30	18.40%
ka 'say'	26	15.95%
ɨto 'go'	23	14.11%
oepɨ 'come'	13	7.98%
e-poreʔka 'arrive'	3	1.84%
ot-urupo 'ask'	2	1.23%
ot-uʔ 'eat'	2	1.23%
os-enakũnuʔ 'choke'	2	1.23%
e-unopɨ 'laugh'	1	0.61%
at-akĩma 'pack bags'	1	0.61%
at-ankɨema 'be happy'	1	0.61%
os-ereʔ 'be amazed'	1	0.61%
e-metɨka 'lose loincloth'	1	0.61%
e-tuarima 'suffer'	1	0.61%
e-puka 'fall'	1	0.61%
os-eporɨ 'meet'	1	0.61%
ot-ɨrɨiʔka 'land'	1	0.61%
ot-ɨiʔka 'finish'	1	0.61%
ot-uru 'talk'	1	0.61%
at-apiaka 'divide up'	1	0.61%
e-sɨrɨiʔma 'move'	1	0.61%

Table 24: Predictions for Proto-Tiriyoan

	**ka* 'say'	**[ə]tə[mɨ]* 'go'	**a* 'be-1'	**eʔi* 'be-2'	**əʔepɨ* 'come'	**epɨ* 'bathe'
DETRZ	×	×	×	×	✓	✓
DETRZ+freq	×	×	×	×	×	✓
phono (/ _**ə*, *e*)	×	×	×	✓	✓	✓
phono+freq	×	×	×	×	×	✓
infl (**w*-)	✓	✓	✓	✓	✓	✓
infl+freq	×	×	×	×	×	✓

Table 25: Evaluating predictions for Proto-Tiriyoan

	*ka 'say'	*[ə]tə[mɨ] 'go'	*a 'be-1'	*eʔi 'be-2'	*əʔepɨ 'come'	*epɨ 'bathe'	Score
DETRZ+freq	✓	✓	✓	✓	✓	✓	100.0%
phono+freq	✓	✓	✓	✓	✓	✓	100.0%
infl+freq	✓	✓	✓	✓	✓	✓	100.0%
DETRZ	✓	✓	✓	✓	×	✓	83.3%
phono	✓	✓	✓	×	×	✓	66.7%
infl	×	×	×	×	×	✓	16.7%

Table 26: Overview of prediction accuracy

	DETRZ	DETRZ+freq	phono	phono+freq	infl	infl+freq
Proto-Waiwaian *k-	100.0%	100.0%	60.0%	100.0%	20.0%	100.0%
Proto-Pekodian *k-	100.0%	100.0%	71.4%	100.0%	0.0%	71.4%
Proto-Tiriyoan *t-	83.3%	100.0%	66.7%	100.0%	16.7%	100.0%
Akuriyó *k-*	66.7%	83.3%	100.0%	100.0%	100.0%	100.0%
Carijona *j-*	60.0%	60.0%	100.0%	60.0%	40.0%	60.0%
Yukpa *j-*	33.3%	33.3%	100.0%	33.3%	66.7%	33.3%

Table 22, but each extension affected many run-of-the-mill S_A verbs.[14] If one simulates the addition of 1'000 regularly inflected detransitivized S_A verbs per language – a conservative estimate based on Courtz's (2008) Kari'ña dictionary – all six hypotheses consistently predict the behavior of 99.99+% verbs correctly. Since the descriptive coverage of Cariban languages makes using real such large-scale data impossible, the present investigation is restricted to the edge cases.

The extent of the Proto-Waiwaian and Proto-Pekodian and extensions is fully predicted by the presence or absence of a detransitivizer. In both cases, only the underived[15] S_A verbs were not affected, all other S_A verbs taking *k-. Not shown in Table 26 are subsequent evolutions in the Pekodian daughter languages, which largely support a detransitivizer-based explanation: First, both Ikpeng and Bakairi regularized the paradigm to use forms with detransitivizer for first person (Section 3.1.1), and both introduced *k-*.[16] Second, the development of Proto-Pekodian *ɨtən 'to go' to Ikpeng *aran* may have led to reanalysis of *ar* as a detransitivizer, accompanied by the introduction of *k-*.

The extent of three extensions (in Akuriyó, Carijona, and Yukpa) is correctly predicted by phonological criteria. As discussed in Section 3.1.4, Akuriyó *k-* only appears on *ə-*initial verbs. In Carijona, the extension of *j-* affected all *e-* and *ə-*initial verbs, including *eh* 'to come' and *e͡ʧi* 'to be', which do not have a detransitivizing prefix. Only *ka* 'say', *tə[mə]* 'go', and *a* 'be-1' did not take on *j-*. The extension of Yukpa *j-* can succinctly be characterized as affecting all vowel-initial verbs, as the only attested conservative verb is C-initial *to* 'to go'.

Inflectional morphology as a network basis only played a potential role in the case of Akuriyó. However, this explanation rests on the analysis of *t-* and *ʧ-* as distinct morphemes. Due to their phonologically conditioned distribution in Proto-Tiriyoan, the hypothetical inflectionally defined network is identical to the phonological one.

[14] While some non-detransitivized S_A verbs (Meira 1999: 252, 2000: 222; Gildea & D. Payne 2007: 30) are not in Table 22, these are mostly *ə-initial and were likely productively derived at some point, see Meira (1999: 252) for Tiriyó. The verbs to which this does not apply, like Tiriyó *wa* 'to dance' (Meira 1999: 252), are all instances of S_P verbs switching classes. Since none of them is attested as having S_A status at the time of a person marker extension, they are not relevant for this study.

[15] The assumption is made that the idiosyncratic evolution of *e-pɨ 'to bathe (INTR)' to *ipɨ Proto-Pekodian made the verb morphologically opaque.

[16] If one instead assumes that first person *w-ebɨ- and *k-əd-ebɨ- already co-existed in Proto-Pekodian, the clear correlation between *k- and the detransitivizer remains.

When adding the expected conservative effects of frequency, prediction accuracy was improved in 8 cases, stagnated in 7 cases, and worsened in three cases. These three cases are found in Carijona and Yukpa, the only languages to feature innovative markers on the reflexes of **eti* 'be-2' or **a[p]* 'be-1'. Overall, including the crude estimate of frequency in the model led to improvements, up to 100% prediction accuracy for all three factors in Proto-Tiriyoan, phonology and inflection in Proto-Waiwaian, and phonology in Proto-Pekodian.

One can conclude that the patterns of most extensions are correctly predicted not by a single explanation, but rather by 3 to 4 different ones. The exceptions are Carijona and Yukpa, where a phonologically defined lexical network emerges as the unambiguous winner, while frequency-based explanations fare worse. For the other extensions, the network model gives no unambiguous answer to the question of what (combination of) factors caused innovative first person markers to spread the way they did. This is due three of the factors in the model – frequency, detransitivizer, phonology – strongly converging in their predictions: The most frequent S_A verbs are at the same time those without a detransitivizer, and therefore mostly of a different phonological shape than regular S_A verbs.

5 Conclusion

The main research questions of this study were a) where did irregular first person prefixes originate? b) which verbs are irregular in what language? and c) why are they irregular? The main findings can be summarized as follows: Irregular first person inflections are conservative, leftovers of person marker extensions which did not affect some verbs. Some of these extensions are reconstructible to proto-languages, while others happened in pre-modern stages of single languages. The same 1-7 verbs are conservative in all languages, and are often irregular in other ways. Bybee's (1985) network model delivers explanations for distribution of innovative markers, but in in 4 of 6 cases, multiple explanations predict the same outcome. This is due to the situation reconstructible to Proto-Cariban, where only a small group of frequent S_A verbs had no detransitivizer **ətə/e-*.

It was the association of **ətə/e-*with A-oriented prefixes that led to the split-S system (Meira 2000). An open question is why the few underived S_A verbs that are reconstructible to Proto-Cariban are not only the most frequent S_A verbs, but arguably the most frequent intransitive verbs ('be', 'go', 'say'). At least the first two meanings could just as well be expressed with S_P verbs. Further comparative work on the detransitivizer might answer this question.

Bybee's (1985) network model held its promise of explaining irregularities in inflectional patterns. It did overshoot its goal somewhat in that only two of six cases had an unambiguous answer – phonology. Notably, the two phonology-based hypotheses (with and without frequency) together correctly predicted 100% of patterns. Since the crude frequency model increased overall prediction accuracy, but decreased it in the case of Carijona and Yukpa, language-specific counts of S_A verbs would be interesting. Regardless of the interpretation of the results, the inconclusiveness of the network model's answers is due to a specific pattern in Proto-Cariban. This raises the general question how applicable the model is when ambiguity exists.

As mentioned, additional and more extensive counts of S_A verbs in different Cariban languages would be important not only for comparison with the Apalaí sample, but also as better input for the network model. Apart from this specific purpose, there is a general need for corpora of Cariban languages accompanying the improving descriptive coverage, allowing studies like Sapién et al. (2021). Finally, this study could have benefitted from more extensive descriptive work on Yukpa and Carijona specifically, while such work on underdescribed languages would benefit Cariban linguistics as a whole.

6 Abbreviations

The following abbreviations were used in this paper: 1 first person, 1+2 first and second person, 1+3 first and third person, 2 second person, 3 third person, A agentive transitive argument, AGT agent, ALL

allative, ANA anaphoric, ATTRZ attributivizer, CAUS causative, CERT certainty, COM comitative, CONT continuative, COR coreference, DAT dative, DEM demonstrative, DES desiderative, DETRZ detransitivizer, DUB dubitative, EVID evidentiality, FRUST frustrative, HSY hearsay/indirect evidentiality, IMM immediate past, IMP imperative, INAN inanimate, INF infinitive, INT intermediate past, INTR intransitive, INTS intensifier, LK linker, LOC locative, MED medial, NEG negation, NMLZ nominalizer, NPST non-past, OBL oblique, P patientive transitive argument, PFV perfective, PL plural, PRO pronoun, PRS present, PST past, REC recent past, REM remote past, S intransitive argument, S_A S marked like A, S_P S marked like P, SAP speech act participant, SUP supine, TAM Tense-Aspect-Mood, TR transitive, UNCERT uncertainty.9

7 Acknowledgments

REDACTED

References

Abbott, Miriam (1991). "Macushi". In: *Handbook of Amazonian Languages*. Desmond C. Derbyshire & Geoffrey K. Pullum (eds.). Vol. 3. Berlin/New York: Mouton de Gruyter: 23–160.

Álvarez, José (2000). "Construcciones progresivas en pemón y otras lenguas caribes". In: *Opción* 32: 96–130.

— (2008). "Cláusulas relativas, nominalización y constituyentes en pemón (Caribe)". In: *Opción* 24.57: 114–143.

Alves, Ana Carolina Ferreira (2017). "Morfofonologia, morfossintaxe e o sistema de tempo, aspecto e modo em Arara (Karib)". Ph.D. dissertation. Universidade de São Paulo.

Alves Chagas, Angela Fabíola (2013). "O Verbo Ikpeng: estudo morfossintático e semântico-lexical". Ph.D. dissertation. Universidade Estadual de Campinas.

Bruno, Ana Carla dos Santos (1996). *Dicionário Waimiri-Atroari-Português*. Manaus: Programa Waimirí-Atroari / Conevênio FUNAI-Eletronorte.

Bybee, Joan (1985). *Morphology: A study of the relation between meaning and form*. Amsterdam/Philadelphia: John Benjamins.

— (1995). "Regular morphology and the lexicon". In: *Language and cognitive processes* 10.5: 425–455. DOI: 10.1080/01690969508407111.

Cáceres, Natalia (2011). "Grammaire fonctionnelle-typologique du ye'kwana, langue caribe du Venezuela". Ph.D. dissertation. Université Lumière Lyon 2.

Cáceres, Natalia & Spike Gildea (2018). "A first analysis of Tense-Aspect constructions in Yawarana (Cariban)". Paper presented at Amazónicas VI. Baños, Colombia.

Cáceres Arandia, Natalia, Spike Gildea & Marie-Claude Mattéi-Müller (2017). "Loss of Inflection in Yawarana (Cariban)". Paper presented at the International Conference on Historical Linguistics 23. San Antonio.

Caesar-Fox, Desrey Clementine (2003). "Zauro'nödok Agawayo Yau: Variants of Akawaio spoken at Waramadong". Ph.D. dissertation. Rice University.

Camargo, Eliane (2002). *Léxico bilingüe: aparai-português / português-aparai*. München: Lincom Europa.

Camargo, Eliane & Anaiman Tapinkili (2010). *Hakëne omijau eitop Wajana-Palasisi / Dictionnaire bilingue Wayana-Français*. Paris: CELIA/DRAC-Guyane, TEKUREMAI.

Campetela, Cilene (1997). "Análise do sistema de marcação de caso nas orações independentes da língua Ikpeng". MA thesis. Universidade Estadual de Campinas.

Carlin, Eithne B. (2004). *A Grammar of Trio: A Cariban Language of Suriname*. Vol. 55. Duisburger Arbeiten zur Sprach- und Kulturwissenschaft. Frankfurt am Main: Peter Lang.

Courtz, Henk (2008). *A Carib grammar and dictionary*. Toronto: Magoria Books.

Cruz, Maria Odileiz Sousa (2005). "Fonologia e gramática Ingarikó: Ka?pon-Brasil". Ph.D. dissertation. Vrije Universiteit Amsterdam.

de Armellada, Cesáreo (1944). *Gramática y diccionario de la lengua Pemón*. Vol. 2: *Diccionario*. Caracas: Artes Gráficas.
Derbyshire, Desmond C. (1965). *Textos Hixkaryâna*. Vol. 3. Publicações Avulsas. Belém: MPEG.
— (1979). *Hixkaryana*. Vol. 1. Lingua Descriptive Studies. Amsterdam: North-Holland.
— (1985). *Hixkaryana and Linguistic Typology*. Dallas: Summer Institute of Linguistics.
— (1999). "Carib". In: *The Amazonian Languages*. R.M.W. Dixon & Alexandra Y. Aikhenvald (eds.). Cambridge: Cambridge University Press: 23–64.
de Souza, Shirley Dias Cardoso (1993). "Alguns aspectos morfológicos da língua Arara (Karib)". MA thesis. Universidade de Brasília.
— (2010). "The morphology of nouns in the Ugoroŋmo language (Arara of Pará)". MA thesis. University of North Dakota.
Franchetto, Bruna (1986). "Falar Kuikúru: Estudo Etnolinguistico de um grupo Karíbe do Alto Xingu. Volume III: Fonologia e Textos". Ph.D. dissertation. Museu Nacional da Universidade Federal do Rio de Janeiro.
— (2002). "Kuikuro: uma língua ergativa no ramo meridional da família Karib (Alto Xingu)". In: *Ergatividade na Amazônia*. Vol. 1. Francesc Queixalós (ed.). Brasília, Paris: Laboratório de Línguas Indígenas, Centre d'études des langues indigènes d'Amérique.
— (2008). "Absolutivo e ergativo pensados a partir do Kuikuro". In: *Amerindia* 32: 37–56.
Franchetto, Bruna & Sérgio Meira (2016). "A phonological hypothesis on the historical origin of Kuikuro and Bakairi verb classes". Talk held at Amazónicas VI. Universidad Nacional de Colombia en Amazonia, Leticia.
García Ferrer, Donaldo José (2006). "Diccionario piloto pemón-español". BA thesis. Universidad del Zulia.
Gildea, Spike (1994). "Akuriyó field notes".
— (1998). *On reconstructing grammar: Comparative Cariban morphosyntax*. Oxford: Oxford University Press.
— (2012). "Linguistic studies in the Cariban family". In: *The Indigenous Languages of South America: A Comprehensive Guide*. Lyle Campbell & Verónica Grondona (eds.). Vol. 2. The World of Linguistics. Berlin/New York: Mouton de Gruyter: 441–494.
— (2015). "The evolution of valence categories in the Cariban family". Paper presented at the Venezuelan Cariban Conference. Lyon. URL: https://www.academia.edu/30561539/The_evolution_of_valence_categories_in_the_Cariban_family (visited on 10/02/2020).
— (2018). "Reconstructing the copula and nonverbal predicate constructions in Cariban". In: *Nonverbal predication in Amazonian languages*. Simon E. Overall, Rosa Vallejos & Spike Gildea (eds.). Amsterdam: John Benjamins: 365–402.
Gildea, Spike & Natalia Cáceres (in preparation). "The Cariban Language Family". In: *Amazonian Languages: An International Handbook*. Patience Epps & Lev Michael (eds.). Berlin: Mouton de Gruyter.
Gildea, Spike, Berend J. Hoff & Sérgio Meira (2010). "The story of *ô in the Cariban Family". In: *Fieldwork and linguistic analysis in Indigenous languages of the Americas*. Andrea L. Berez, Jean Mulder & Daisy Rosenblum (eds.). University of Hawai'i Press: 91–123.
Gildea, Spike & Doris Payne (2007). "Is Greenberg's "Macro-Carib" viable?" In: *Boletim do Museu Paraense Emílio Goeldi. Ciências Humanas* 2.2: 19–72.
Gildea, Spike & Fernando Zúñiga (2016). "Referential hierarchies: A new look at some historical and typological patterns". In: *Linguistics* 54.3: 483–529.
Girard, Victor James (1971). "Proto-Carib phonology". Ph.D. dissertation. University of California.
Guerrero Beltrán, David Felipe (2019). "The Grammar of Space in Karijona, a Cariban language from Northwest Amazonia". MA thesis. Universidad Nacional de Colombia.
Guerrero-Beltrán, David Felipe (2016). "The alignment and person-marking system in Karijona (Cariban, Northwest Amazonia)". Paper presented at the Language and Culture Research Centre, James Cook University.
Hammarström, Harald et al. (2020). "Glottolog 4.2.1". Online: http://glottolog.org (visited on 05/18/2020).

Harris, Alice C. & Lyle Campbell (1995). *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge University Press.

Haspelmath, Martin & Andrea Sims (2010). *Understanding morphology*. 2nd ed. London: Hodder Education.

Hawkins, Robert E. (1998). "Wai Wai". In: *Handbook of Amazonian Languages*. Desmond C. Derbyshire & Geoffrey K. Pullum (eds.). Vol. 4. Berlin/New York: Mouton de Gruyter: 25–224.

Hawkins, W. Neill & Robert E. Hawkins (1953). "Verb inflection in Waiwai (Carib)". In: *International Journal of American Linguistics* 19.3: 201–211.

Hoff, Berend J. (1968). *The Carib Language: Phonology, Morphophonology, Morphology, Texts and Word Index*. Vol. 55. Verhandelingen van het Koninklijk instituut voor taal-, land- en volkenkunde. The Hague: Martinus Nijhoff.

Howard, Catherine V. (2001). "Wrought Identities: The Waiwai expeditions in search of the "unseen tribes" of Northern Amazonia". Ph.D. dissertation. University of Chicago.

Koch-Grünberg, Theodor (1908). "Die Hianákoto-Umáua". In: *Anthropos* 3: 83–124, 297–335, 952–982.

Koehn, Edward & Sally Koehn (1986). "Apalai". In: *Handbook of Amazonian Languages*. Desmond C. Derbyshire & Geoffrey K. Pullum (eds.). Vol. 1. Berlin/New York: Mouton de Gruyter: 33–127.

— (1994). *14 textos Apalaí*. Brasília: Summer Institute of Linguistics.

Lacerda, Edith (1991). "Waimiri-Atroari: observações lingüísticas". MS. Manaus.

Largo, Wilson S. (2011). "Una gramática Yukpa". MS. Bogotá: Fundación para el Desarrollo de Pueblos Marginados.

List, Johann-Mattis et al. (2021). *LingPy: A Python library for quantitative tasks in historical linguistics*. Version v2.6.8. Leipzig. DOI: 10.5281/zenodo.5144474. URL: https://lingpy.org.

Mattéi-Müller, Marie Claude (1975). "Vocabulario básico de la lengua mapoya". In: *Antropológica* 42: 57–78.

Mattéi-Müller, Marie-Claude (1994). *Diccionario ilustrado Panare-Español, Índice Español-Panare: Un aporte al estudio de los Panares-E'ñepa*. Caracas: Graficas Armitano.

Matter, Florian (2021a). "Diachronic and typological aspects of Cariban verbal argument marking". Ph.D. dissertation. University of Bern.

— (2021b). "The origin of split third person marking in Arara (Cariban)". In: *Cadernos de Etnolingüística* 8.2: 1–22. URL: http://www.etnolinguistica.org/article:vol8n2-1.

Meira, Sérgio (1998). "A reconstruction of Proto-Taranoan: Phonology and Inflectional Morphology". MA thesis. Rice University.

— (1999). "A Grammar of Tiriyó". Ph.D. dissertation. Rice University.

— (2000). "The accidental intransitive split in the Cariban family". In: *Reconstructing Grammar: Comparative Linguistics and Grammaticalization*. Spike Gildea (ed.). Vol. 43. Typological Studies in Language. Amsterdam: John Benjamins: 201–230.

— (2002). "A first comparison of pronominal and demonstrative systems in the Cariban language family". In: *Current Studies on South American Languages: Selected papers from the 50th International Congress of Americanists in Warsaw and the Spinoza Workshop on Amerindian Languages in Leiden*. Mily Crevels et al. (eds.). Leiden: Research institution of Asian, African, and Amerindian Studies (CNWS): 255–275.

— (2003a). "A marcação de pessoa nos verbos em Bakairi (Karíb)". In: *Ergatividade na Amazônia*. Vol. 2. Francesc Queixalós (ed.). Brasília, Paris: Laboratório de Línguas Indígenas, Centre d'études des langues indigènes d'Amérique: 155–163.

— (2003b). "Primeras observaciones sobre la lengua yukpa". In: *Lingua Americana* 7.12.

— (2005). "Reconstructing Pre-Bakairi Segmental Phonology". In: *Anthropological Linguistics* 47.3: 261–291.

— (2006a). "Cariban Languages". In: *Encyclopedia of language and linguistics*. Keith Brown (ed.). Oxford: Elsevier: 199–203.

— (2006b). "Syntactic reanalysis in Yukpa (Cariban)". In: *International Journal of American Linguistics* 72.1: 135–150.

Meira, Sérgio & Bruna Franchetto (2005). “The Southern Cariban languages and the Cariban family”. In: *International Journal of American Linguistics* 71.2: 127–192.
Meira, Sérgio & Spike Gildea (2009). “Property concepts in the Cariban family: adjectives, adverbs, and/or nouns”. In: *The linguistics of endangered languages: Contributions to morphology and morphosyntax*. W. Leo Wetzels (ed.). Utrecht: LOT: 95–133.
Meira, Sérgio, Spike Gildea & Berend J. Hoff (2010). “On the Origin of Ablaut in the Cariban Family”. In: *International Journal of American Linguistics* 76.4: 477–515.
Méndez-Arocha, Alberto (1959). “Un vocabulario Yabarana con Apuntes Fonéticos”. In: *Antropológica* 7: 65–84.
Oquendo, Luis (2004). “Is Japreria a Yukpa Dialect?” In: *Lingua Americana* 8.14: 87–98.
Pachêco, Frantomé B. (1997). “Aspectos da gramática Ikpeng (Karíb)”. MA thesis. Universidade Estadual de Campinas.
— (2001). “Morfossintaxe do verbo Ikpeng (Karíb)”. Ph.D. dissertation. Universidade Estadual de Campinas.
Payne, Thomas E. & Doris L. Payne (2013). *A Typological Grammar of Panare: A Cariban Language of Venezuela*. Brill’s Studies in the Indigenous Languages of the Americas. Leiden: Brill.
Revell, Liam J. (2012). “phytools: An R package for phylogenetic comparative biology (and other things)”. In: *Methods in Ecology and Evolution* 3: 217–223.
Robayo Moreno, Camilo Alberto (1989). ““En faisant une rame”: texte carijona”. In: *Amerindia* 14: 189–199.
— (2000). “Avance sobre morfología carijona”. In: *Lenguas indígenas de Colombia: una visión descriptiva*. María Stella González de Pérez & María Luisa Rodríguez de Montes (eds.). Santafé de Bogotá: Instituto Caro y Cuervo: 171–180.
Sapién, Racquel-María et al. (2021). “Antipassive and semantic classes of verbs in the Cariban family”. In: *The Multifaceted Nature of Antipassive*. Katarzyna Janic & Alena Witzlack-Makarevich (eds.). Amsterdam: John Benjamins.
Stegeman, Ray & Rita Hunter (2014). *Akawaio-English Dictionary*. SIL International. URL: `https://akawaio.webonary.org/`.
Stolz, Thomas, Aina Urdze & Johan van der Auwera (2012). “Introduction: Irregularity – glimpses of a ubiquitous phenomenon”. In: *Irregularity in Morphology (and beyond)*. Thomas Stolz et al. (eds.). Berlin: Akademie Verlag: 7–38. DOI: `10.1524/9783050059587`.
Swiggers, Gisela (2010). “Gramática de Referencia del Pemón”. Ph.D. dissertation. Universidad del Zulia.
Tavares, Petronila da Silva (2005). “A grammar of Wayana”. Ph.D. dissertation. University of Houston.
von den Steinen, Karl (1892). *Die Bakaïri-Sprache: Wörterverzeichniss, Sätze, Sagen, Grammatik; mit Beiträgen zu einer Lautlehre der karaïbischen Grundsprache*. Leipzig: Koehler.
Wheatley, James (1969). “Bakairi verb structure”. In: *Linguistics* 47.47: 80–100.

A Predictions for the behavior of individual verbs

This appendix shows all predictions and the comparison with the attested data.

Table 27: Predictions for Proto-Waiwaian

	**ka[s]* 'say'	**[ɨ]to[m]* 'go'	**ah* 'be-1'	**eʃi* 'be-2'	**eeɸɨ* 'bathe'
DETRZ	×	×	×	×	✓
DETRZ+freq	×	×	×	×	✓
phono (/ _**o*, *e*, *a*)	×	×	✓	✓	✓
phono+freq	×	×	×	×	✓
infl (**w*-)	✓	✓	✓	✓	✓
infl+freq	×	×	×	×	✓

Table 28: Evaluating predictions for Proto-Waiwaian

	*ka[s] ‘say’	*[i]to[m] ‘go’	*ah ‘be-1’	*eʃi ‘be-2’	*eeɸɨ ‘bathe’	Score
DETRZ	✓	✓	✓	✓	✓	100.0%
DETRZ+freq	✓	✓	✓	✓	✓	100.0%
phono+freq	✓	✓	✓	✓	✓	100.0%
infl+freq	✓	✓	✓	✓	✓	100.0%
phono	✓	✓	×	×	✓	60.0%
infl	×	×	×	×	✓	20.0%

Table 29: Predictions for Proto-Pekodian

	*ke ‘say’	*ɨtə[n] ‘go’	*ap ‘be-1’	*eʧi ‘be-2’	*epɨ ‘come’	*ɨptə ‘go down’	*ipɨ ‘bathe’
DETRZ	×	×	×	×	×	×	×
DETRZ+freq	×	×	×	×	×	×	×
phono (/ _*ə, e)	×	×	×	✓	✓	×	×
phono+freq	×	×	×	×	×	×	×
infl (*w-)	✓	✓	✓	✓	✓	✓	✓
infl+freq	×	×	×	×	×	✓	✓

Table 30: Evaluating predictions for Proto-Pekodian

	*ke ‘say’	*ɨtə[n] ‘go’	*ap ‘be-1’	*eʧi ‘be-2’	*epɨ ‘come’	*ɨptə ‘go down’	*ipɨ ‘bathe’	Score
DETRZ	✓	✓	✓	✓	✓	✓	✓	100.0%
DETRZ+freq	✓	✓	✓	✓	✓	✓	✓	100.0%
phono+freq	✓	✓	✓	✓	✓	✓	✓	100.0%
phono	✓	✓	✓	×	×	✓	✓	71.4%
infl+freq	✓	✓	✓	✓	✓	×	×	71.4%
infl	×	×	×	×	×	×	×	0.0%

Table 31: Predictions for Proto-Tiriyoan

	*ka ‘say’	*[ə]tə[mɨ] ‘go’	*a ‘be-1’	*eʔi ‘be-2’	*əʔepɨ ‘come’	*epɨ ‘bathe’
DETRZ	×	×	×	×	✓	✓
DETRZ+freq	×	×	×	×	×	✓
phono (/ _*ə, e)	×	×	×	✓	✓	✓
phono+freq	×	×	×	×	×	✓
infl (*w-)	✓	✓	✓	✓	✓	✓
infl+freq	×	×	×	×	×	✓

Table 32: Evaluating predictions for Proto-Tiriyoan

	**ka* 'say'	**[ə]tə[mɨ]* 'go'	**a* 'be-1'	**eʔi* 'be-2'	**əʔepɨ* 'come'	**epɨ* 'bathe'	Score
DETRZ+freq	✓	✓	✓	✓	✓	✓	100.0%
phono+freq	✓	✓	✓	✓	✓	✓	100.0%
infl+freq	✓	✓	✓	✓	✓	✓	100.0%
DETRZ	✓	✓	✓	✓	×	✓	83.3%
phono	✓	✓	✓	×	×	✓	66.7%
infl	×	×	×	×	×	✓	16.7%

Table 33: Predictions for Akuriyó

	ka 'say'	*[ə]tə[mɨ]* 'go'	*a* 'be-1'	*eepɨ* 'come'	*ɨ[h]tə* 'go down'	*epɨ* 'bathe'
DETRZ	×	×	×	✓	×	✓
DETRZ+freq	×	×	×	×	×	✓
phono (/ _*ə*)	×	×	×	×	×	×
phono+freq	×	×	×	×	×	×
infl (*k*-)	×	×	×	×	×	×
infl+freq	×	×	×	×	×	×

Table 34: Evaluating predictions for Akuriyó

	ka 'say'	*[ə]tə[mɨ]* 'go'	*a* 'be-1'	*eepɨ* 'come'	*ɨ[h]tə* 'go down'	*epɨ* 'bathe'	Score
phono	✓	✓	✓	✓	✓	✓	100.0%
phono+freq	✓	✓	✓	✓	✓	✓	100.0%
infl	✓	✓	✓	✓	✓	✓	100.0%
infl+freq	✓	✓	✓	✓	✓	✓	100.0%
DETRZ+freq	✓	✓	✓	✓	✓	×	83.3%
DETRZ	✓	✓	✓	×	✓	×	66.7%

Table 35: Predictions for Carijona

	ka 'say'	*tə[mə]* 'go'	*a* 'be-1'	*eʧi* 'be-2'	*eh[ɨ]* 'come'
DETRZ	×	×	×	×	×
DETRZ+freq	×	×	×	×	×
phono (/ _*ə*, *e*)	×	×	×	✓	✓
phono+freq	×	×	×	×	×
infl (**w*-)	✓	✓	✓	✓	✓
infl+freq	×	×	×	×	×

Table 36: Evaluating predictions for Carijona

	ka ‘say’	*tə[mə]* ‘go’	*a* ‘be-1’	*eʧ͡ʃi* ‘be-2’	*eh[ɨ]* ‘come’	Score
phono	✓	✓	✓	✓	✓	100.0%
DETRZ	✓	✓	✓	×	×	60.0%
DETRZ+freq	✓	✓	✓	×	×	60.0%
phono+freq	✓	✓	✓	×	×	60.0%
infl+freq	✓	✓	✓	×	×	60.0%
infl	×	×	×	✓	✓	40.0%

Table 37: Predictions for Yukpa

	to ‘go’	*a* ‘be-1’	*e* ‘be-2’
DETRZ	×	×	×
DETRZ+freq	×	×	×
phono (/ _V)	×	✓	✓
phono+freq	×	×	×
infl (**w*-)	✓	✓	✓
infl+freq	×	×	×

Table 38: Evaluating predictions for Yukpa

	to ‘go’	*a* ‘be-1’	*e* ‘be-2’	Score
phono	✓	✓	✓	100.0%
infl	×	✓	✓	66.7%
DETRZ	✓	×	×	33.3%
DETRZ+freq	✓	×	×	33.3%
phono+freq	✓	×	×	33.3%
infl+freq	✓	×	×	33.3%