Verbs with conservative first person forms in Cariban languages

Florian Matter
University of Bern

December 9, 2021

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 two Western and four Southern outliers. Figure 1 gives an overview of the geographical distribution and genealogical affiliation of the extant Cariban languages. For linguistic 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 about here.]

[Table 1 about here.]

[Table 2 about here.]

Some Cariban languages show a small group of verbs which diverge in their first person inflection pattern, a topic which has not received much attention in the literature. This is illustrated for Hixkaryána in Table 1,[1]which shows person paradigms for four verbs, all members of the S_A inflectional class. In this language, the verb 'to be' diverges in its first person marker (*w-*), contrasting with other S_A verbs like 'to fall', which have *k-*. A similar pattern can be found in Tiriyó, 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* (Table 2). In both languages, the inflectional patterns of the verbs on the left of the table 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 analysis for Meira's (1998) reconstruction of Proto-Taranoan. In a synchronic analysis of a language, 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. However, 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. For such approaches, these 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 synchronic analysis, the reason for these inflectional patterns can be found in the diachrony of the languages in question. The purpose of this paper is to approach the patterns from

[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 the paradigms shown here. In Table 1 and other paradigm tables, any TAM suffixes found in the original forms found in the literature are omitted, since a) the focus lies on the prefixes and stems, and b) full paradigms containing the same TAM suffix are rarely found. Further, standard IPA symbols are used in the transcription of Cariban languages, with the exception of coronal rhotics, which are simply represented with ⟨r⟩, rather than ⟨ɽ⟩ for Wayana or ⟨r̲⟩ for Ye'kwana etc. In languages with strong morphophonological processes and/or subphonemic orthography the original transcription is shown in an additional surface line when presented in interlinearized examples. Gildea (2018) is followed in using ⟨ə⟩ for the proto-vowel reconstructed by Meira & Franchetto (2005), although it was likely more back (Gildea et al. 2010). Glossing abbreviations:

a comparative perspective and to provide a diachronic and functional account, proceeding as follows. In Section 2, readers are provided with necessary background knowledge about Proto-Cariban verbs, and it is shown that the mechanism of person marker extension is responsible for the Hixkaryána and Tiriyó patterns. In Section 3, six incomplete person marker extensions and their extent in the lexicon are described. Since there is a surprising amount of agreement about what verbs remain conservative, they are compared and reconstructed. Section 4 uses Bybee's (1985) network model of morphology to find explanations for the verbs (not) affected by each extension. Section 5 summarizes and discusses the results.

2 The origins of conservative first person inflections

The irregular first person prefixes introduced in Section 1 are relics, inherited from the ancestral Proto-Cariban system (Section 2.1). That system was subject to many different kinds of innovations; responsible for the patterns discussed are person marker extensions not spreading through the entire S_A lexicon (Section 2.2). A particular component of the system, the distinction between S_A and S_P verbs, plays a role in what extensions are incomplete and is the topic of 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. The choice of person marker in transitive verbs can be characterized as being conditioned by a basic person hierarchy $\boxed{1/2 > 3}$. The locuphoric markers had two forms, an A-oriented one for direct (SAP>3) scenarios and a 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 expressed in a non-transparent manner, both using the 1+2 prefix *k-.

[Table 3 about here.]

Formally identical or etymologically related markers occured in intransitive verbs, which showed a split-S system (Table 3b). That is, S_A verbs took similar markers as the A-oriented ones in transitive verbs, with the exception of first person (1>3 *$t(i)$- vs $1S_A$ *w-) 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. The third person marker in S_P verbs was identical to the one in 3>3 scenarios (*$n(i)$-).

Equipped with this knowledge about the ancestral system, it becomes 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- may be considered RELICS, old and restricted to specific lexically conditioned contexts, contrasting with the innovative prefixes found elsewhere. These verbs and their prefixes are comparable with the few English nouns like *ɒ̀ks*, which preserve the old plural suffix *-ən*. It, too, was once more widespread, being the normal plural suffix of the weak inflection, compare German *ɔks-n* 'ox-en', *naːmə-n* 'name-s', *haːzə-n* 'hare-s', *bɛːʁ-ən* 'bear-s'. Since the regular Hixkaryána and Tiriyó prefixes are innovative, the next question is where they came from.

2.2 Person marker extensions and lexical diffusion

In his discussion of the Proto-Cariban split-S system and reconstruction of the intransitive person prefixes, Gildea (1998: 88–96) shows that the system has undergone many different modifications in various languages. The main mechanism of change leading to these modifications are **person marker extensions**, i.e. the use of verbal person prefixes being extended to contexts previously occupied by other prefixes. There have been many different person marker extensions in Cariban languages, and some are still ongoing. This is illustrated by Gildea (1998), using the three Parukotoan languages as an example. Apart from segmental changes to individual morphemes, the following innovations happened in the Set I paradigm in Parukotoan:

1. Proto-Parukotoan
   (a) $1S_A$ *w- to 1>3
   (b) 1+2 *k- to $1S_P$ (completed in Proto-Waiwaian, ongoing in Werikyana)
   (c) 1+2 *$kɨt$- to $1+2S_P$ (completed in Proto-Waiwaian, ongoing in Werikyana)
2. Proto-Waiwaian
   (a) $1S_P$ *k- to $1S_A$
   (b) innovative *$owɨro\ j$- '1PRO LK' for 1P
3. Waiwai
   (a) $2S_A$ m- to $2S_P$

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 $\varnothing/j$- (Spike Gildea, p.c.).

[Figure 2 about here.]

Hixkaryána has preserved split-S only in the second person prefixes, while Werikyana still shows the variation in the first person and 1+2 prefixes that is reconstructible to Proto-Parukotoan. Waiwai, on the other hand, has lost the system entirely, which notably happened via distinct innovations at three different diachronic stages.

Gildea (1998) discusses person marker extensions in the context of the loss of the split-S system and the accompanying changes to indexing alignment, but this study focuses on a particular aspect of theirs. To begin, they most likely took place via lexical diffusion, characterized as a type of extension by Harris & Campbell (1995: 106–115); this hypothesis is supported by three facts. First, the variation in first person and 1+2 prefixes described above for Werikyana is not completely free. Rather, some verbs only allow for example first person k-, but not j-, while others can occur with both, which is the expected pattern in a lexical diffusion scenario. In addition, this is speaker-dependent (Spike Gildea, p.c.), which is what one would expect from a change in progress. Second, while there is no detailed diachronic scenario for the switch of 1>3 *t- and $1S_A$ in the Tiriyoan languages (Section 3.1.3), Meira (1998: 111–112) convincingly argues that it must have happened gradually rather than instantaneously,

and entailed both markers spreading at the same time. Whether this gradual switch was along ordered lines or not, lexical diffusion must have played a role.

The third argument in favor of the lexical diffusion scenario relates to the conservative Hixkaryána and Tiriyó forms in Tables 1 and 2. In both cases, the 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 best interpreted as the extension stopping short of these verbs, rather than spreading through the entire S_A lexicon. In a family-wide investigation of person marker extensions, 19 distinct extensions affecting intransitive verbs were identified, 6 of them incomplete. These incomplete extensions have left between 1 and 7 conservatively inflected verbs in 9 Cariban languages, and are treated individually in Section 3.

Interestingly, all six extensions 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 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), or the extension of the entire S_A set to S_P verbs in Waimiri-Atroari (Table 4c). This starkly different behavior of S_A and S_P verbs when subject to extensions points to the split-S system playing some role, so its main properties will be discussed in Section 2.3. Readers will also get an idea how it is possible for the S_A/S_P distinction to be lost only for a single person, or for S_P verbs to take on S_A markers with apparent semantic impunity.

[Table 4 about here.]

2.3 The Cariban split-S system

As seen in Section 2.1, the distinction between S_A and S_P verbs was instantiated by two inflection classes within the Proto-Cariban Set I person paradigm, but this was not the only inflectional difference: Many languages show an S_A class marker in deverbalized forms, originating in Proto-Cariban **w-*.[3] With S_A verbs, **w-* occurred immediately between the possessive prefixes and the verb stem, while S_P verbs took the bare prefixes (Tables 5 and 6).

[Table 5 about here.]

[Table 6 about here.]

[Table 7 about here.]

The distinction between S_A and S_P was also borne out in imperatives, where the latter took the P-oriented second person prefix **ə(j)-* while the former received no prefix (Table 7). Both the S_A marker **w-* and the prefix pattern in imperatives have been lost in some languages, but are reconstructible to Proto-Cariban.

[2] As an honorable mention, when Ikpeng introduced a new mixed person paradigm to Set I constructions, *a* 'to be' and *ke* 'to say' retained third person *n-* (Matter 2021b: 12). However, the spread of the innovative third person markers had an entirely different pathway from the extensions discussed here, spreading from subordinate to main clauses, rather than from verbs to verbs.

[3] See Meira (2000: 227), who identifies reflexes of this morpheme as having "no purpose other than being 'class markers', without any obvious semantic or functional load".

In the modern languages maintaining the split-S system, mismatches between the semantics of verbs and their A- or P-oriented inflectional morphology are common, exemplified with Kari'ña data in (1–2).

(1) Kari'ña

a. *mi-kupi-ja*
2>3-bathe-PRS
'You bathe him/her.' (Hoff 1968: 160)

b. *a-kupi-ja*
3>2-bathe-PRS
'S/he bathes you.' (Yamada 2011: 63)

In (1), the choice between the second person A- and P-oriented markers *mi-* and *a-* depends on the scenario: The transitive verb *kupi* 'to bathe' takes *mi-* in 2>3 scenarios (1a), but *a-* in 3>2 scenarios (1b). The intransitive verbs in (2) show the same person markers, but the choice of marker depends on the verb rather than semantics.

(2) Kari'ña

a. *sipi tɨnka-rɨ m-ekema-non hen*
net pull-NMLZ 2-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
2-get.up-PRS
'You are getting up.' (Hoff 1968: 167)

ekema 'to be afraid' takes an A-oriented marker, since it is an S_A verb (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) takes a sizable 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 detransitivizing prefix are treated as S_A verbs, while essentially all others are S_P verbs:

> 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)

Meira (2000: 221–223) also argues that the detransitivizing prefixes are indeed deriving S_A verbs, rather than being inflectional in nature: a) there are a few underived S_A verbs, with no detransitivizing

prefix; b) S_A verbs can develop irregular semantics compared to their transitive counterparts; c) it is unpredictable whether the A or P argument of the underlying transitive verb becomes the S of the derived S_A verb; d) some originally derived S_A verbs have lost their transitive counterparts; and e) "basic" concepts are expressed as derivations of more complex concepts, like 'to dance (S_A)' from 'to dance with (TR)'. He also 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 **əte-* and reflexive **e-*, although on verbs their reflexes have been merged in a single morpheme in modern languages. Modern reflexes of **əte*/*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	→	*e-nonta*,	'abandon each other'
'abandon'		*əi-nonta*	(reciprocal)
suka	→	*e-suka*,	'wash self'
'wash'		*əi-suka*	(reflexive)
pahka	→	*e-pahka*	'break (INTR)'
'break (TR)'			(anticausative)
puunəpɨ	→	*əh-puunəpɨ*,	'think, meditate'
'think about'		*əi-puunəpɨ*	(antipassive)

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

3 Inflections and verbs: innovation and resistance

As was shown in Section 2.2, irregularly inflected first person forms are leftovers from incomplete person marker extensions. Section 3.1 presents the six known incomplete extensions, the prefixes they introduced and the verbs they spared. Since the latter show considerable etymological overlap across languages, the resistant verbs are compared and reconstructed in Section 3.2. In addition, their reflexes which did get affected by one of the extensions are identified, where existent.

3.1 Incomplete extensions: the innovative 1S_A markers

As stated in Section 2.2, the six person marker extensions which did not affect all potential targets all have in common that they feature innovative first person markers on verbs that are (at least historically, in the case of loss of split-S) members of the S_A class. Of these extensions, half can be reconstructed to intermediate proto-languages, while the others happened in pre-modern stages of single languages. The sources 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 details the extension of *k- in Proto-Pekodian, reflected in the three daughter languages Arara, Ikpeng, and Bakairi. Section 3.1.2 treats the extension of *k- in Proto-Waiwaian, which was 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 only 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) focuses 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 8), allowing the reconstruction of a Proto-Pekodian 1S_A marker *k-.

[Table 8 about here.]

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 $w(i)$- rather than k-, shown in (4). There is also a reflex of the copula *$a[p]$, serving syntactically as a postposition introducing adverbial clauses meaning 'if' or 'when' (Alves 2017: 199–201). However, its inflectional morphology features verbal Set I prefixes, including first person w- (5).

(4) Arara (Alves 2017: 153)

wɨ-genɨ	'I said'
w-itʃinɨ	'I was, lied down'
w-ebɨnɨ	'I came'
w-ibɨnɨ	'I bathed'
w-iptoɲrɨ	'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 subclasses of S_A verbs, one taking first person w-, and one k-.[5] The verb used to illustrate the first group is *i* 'to bathe' (6).

[4] Seven under her analysis, which sees the two meanings of *itʃi* 'to be, to lie down' as different verbs.

[5] It should be noted at this point that 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).

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

While Meira (2003a: 4) lists some Bakairi cognates of 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-*. Inflected forms can be found in von den Steinen (1892), presented in (7) according to 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⟩
w-i-taki
$1S_A$-be-INT
'I was.'

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

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

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

f. ⟨töre-w-akine⟩
tərə w-a-kine
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 *i-* or Ø (8). The exception is 'to go', which has *k-* (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*] 'to be' do exist in Ikpeng, apparently only reflexes of **eti* 'to be' occur with first person inflectional prefixes (Gildea 2018: 401).

(8) Ikpeng

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

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

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.

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

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

[Table 9 about here.]

Reconstructed Proto-Pekodian forms of conservatively inflected verbs are given in Table 9. 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 10). Similarly, the change of **wɨ* to Bakairi *u* is found in other correspondences: *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'.

[Table 10 about here.]

Reconstructions of verb stems, Proto-Pekodian and other, are discussed in Section 3.2, but a brief comment on 'to come' is in order: The verb 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*. In contrast, the Arara first person form is directly based on this root **epɨ*. However, reflexes of **əd-epɨ* can be found elsewhere in the Arara paradigm (10).

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

In contrast, Ikpeng and Bakairi show reflexes of **əd-ebɨ* throughout the whole paradigm. Following the line of reasoning used by Meira (1998: 114) for a similar pattern in the three Taranoan languages (see also Section 3.1.3), it is likely that the idiosyncratic pattern in Arara is reconstructible to Proto-Pekodian, and that Bakairi and Ikpeng independently levelled the paradigm in favor of **əd-epɨ*.

3.1.2 Proto-Waiwaian **k*-

This extension, one of the Parukotoan innovations discussed in Section 2.2, resulted in the Hixkaryána patterns from Section 1. In Proto-Waiwaian, the $1S_P$ prefix **k*- innovated in Proto-Parukotoan was further extended to $1S_A$. For regularly inflected verbs, this created a unified 1S category (Table 11).

[Table 11 about here.]

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 *w(ɨ)-* (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 kajwerʲe 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 3.2.2).

There are three more verbs which did not take innovative **k-* in Proto-Waiwaian, shown alongside **ka* 'to say' in Table 12. The two forms for 'to be' are straightforwardly reconstructible, whereas 'to go' is a special case. While Hixkaryána has the expected *ɨ-*, Waiwai seems to have combined innovative *k-* with the old **w-*, an etymological analysis also considered by Gildea (1998: 90). Alternatively, this form may have been influenced by deverbalized forms of 'to go', where a reflex of the S_A class marker **w-* has become fossilized (e.g., *o-**w**to-topo-nho* 'my trip' [R. E. Hawkins 1998: 92]) In any case, the prefix pattern in Hixkaryána 'to go' clearly shows that a Proto-Waiwaian first person form **wɨ-tom-* can be reconstructed.

[Table 12 about here.]

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 already identified by Girard (1971), with Carijona as a more dis-

tant member. Meira (1998) contributes an extensive phonological, morphological, and lexical reconstruction of Proto-Taranoan, facing 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. This resulted in a regular $1S_A$ marker of the form **ʧ-* / *_e*, **t-* / *_ə* (Table 13).[6]

[Table 13 about here.]

The question of how this switch happened in detail (Meira 1998: 107–112) still has no answer, although it seems necessary to assume a scenario in which both **t-* and **w-* for a time occurred on both transitive and intransitive verbs (Meira 1998: 112).[7]

[Table 14 about here.]

Turning to verbs not affected by the spread of **t-*, Meira (1998) reconstructs four of the items in Table 14 as taking **w-* in Proto-Taranoan, for which reconstructed Proto-Tiriyoan forms are substituted here. To this list the second copular root **eʔi* can be added, whose Tiriyó reflex retains first person *w-*. The idiosyncratic Akuriyó first person prefix *ə-* on 'to go' is plausibly identified as a reflex of **wɨ-* by Meira (1998: 113), which is supported by additional evidence: Both components of the irregular change **wɨ- > ə-* (loss of **w* and lowering of **ɨ* to *ə*) are found in other person prefixes (12a–b).

(12) Akuriyó

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

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

For 'to come', Meira (1998: 114–115) reconstructs Proto-Taranoan **əepɨ* for first person, and **eepɨ* for other persons, based on an idiosyncratic paradigmatic pattern in Tiriyó and the vowel length in Akuriyó. Both Akuriyó and Carijona then levelled this original distribution, similar to what was suggested for the Pekodian languages (Section 3.1.1). This scenario is plausible, with the exception that Tiriyó *əepɨ* is a reflex of **ət-epɨ* (Section 3.2.4), meaning that the Proto-Tiriyoan form would have been **əʔepɨ* (Proto-Taranoan **əʧepɨ*).

3.1.4 Akuriyó *k-*

After the split-up of Proto-Tiriyoan, when **t-* had largely replaced **w-*, Akuriyó innovated a third $1S_A$ marker: *k-*. It seems to have replaced **t-* only in specific environments, with the two markers showing a clear phonologically conditioned distribution in Gildea's (1994) Akuriyó data (Table 15). Meira

[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. Here, a conservative stance is taken and the innovation is arbitrarily assumed to be Proto-Tiriyoan. This decision does not affect the results of this study.

(1998: 107) largely confirms the distribution shown here, but mentions "several cases of first person *t-* in Akuriyó" (on *ə*-initial verbs), albeit without any examples. He also suggests that *k-* might be more recent, which is very plausible: since the distribution **t-* / *_ə* / **ʧ-* / *_e* is reconstructible to Proto-Tiriyoan, the most straightforward scenario is *k-* replacing **t-* but not **ʧ-* in Akuriyó. The few *t-* mentioned by Meira (1998) were then perhaps reintroduced under Tiriyó influence. However, since there are no examples of *ə*-initial verbs with *t-*, or further information about them, these cases cannot be discussed further.

[Table 15 about here.]

The verbs listed for Proto-Tiriyoan in Table 14 in Section 3.1.3 did of course also not introduce *k-* in Akuriyó, although the first-person form of the copular verb *eʔi* is unknown. In addition, there is an S_A verb *ɨ(h)tə* 'to go down', which has an irregular first person marker *p-* in, which is apparently reconstructible to Proto-Tiriyoan (13). It was not affected by the extension of Akuriyó *k-*, but whether it was an S_A verb when Proto-Tiriyoan **t-* was extended is unclear (see Section 3.2.5).

(13) First person forms of 'to go down'

Tiriyó	*p-ɨhtə-*	(Meira 1999: 294)
Akuriyó	*p-ɨtə-*	(Gildea 1994: 84)

3.1.5 Carijona *j-*

Carijona, the cousin of the Tiriyoan languages, has extended the $1S_P$ marker *j(i)-*[8] to S_A verbs (Meira 1998: 105–107). Combined with the extension of $2S_A$ *m-* and $1+2S_A$ *kɨt-*/*kɨs-* to S_P verbs, this created a single unified S category for regularly inflected verbs (Table 16).

[Table 16 about here.]

Although the split-S system has been lost entirely, former S_A verbs can be identified by the presence of a detransitivizer, like *ehɨnəhɨ* 'to fight' (14), derived from *hɨnəhɨ* 'to kill' (Robayo Moreno 2000: 179).

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

As noted in Section 3.1.3, this extension also erased any traces of a putative Proto-Taranoan $1S_A$ marker **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).

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

(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-with-PL 1-be-NPST
'I am with you all.'

c. *dëmëmara kae ëwï ïya*
n-tə-mə=mara ∅-ka-e əwɨ i-ja
3-go-PST=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', one would expect *ka* 'to say' to 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 cases of loss were more regular, an already irregular marker undergoing idiosyncratic phonological erosion is not that surprising, see Akuriyó **wɨ- > ə-* in the preceding section. Alternatively, it is possible that the divergent development of **w-* on C-initial *ka* 'to say' and *tə* 'to go' is a result of the latter's originally V-initial nature (Section 3.2.3).

3.1.6 Yukpa *j-*

The divergent nature of the family-internal isolate Yukpa 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, where a unified intransitive paradigm (Table 17) is found. 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'.

[Table 17 about here.]

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-*. The same prefix occurs on transitive verbs in 3>1 scenarios (16a), so *j(ɨ)-* is the reflex of the Proto-Cariban $1S_{(P)}$ marker **u(j)-*. 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 elsewhere in the family (Sections 3.1.1 and 3.1.2), the zero marking in 1>3 scenarios can be identified as the Yukpa reflex of *w-, contrasting with $j(ɨ)$-, the reflex of *$u(j)$-.

In intransitive verbs, this first-person zero marking is attested in a single verb, *to* 'to go' (17). It diverges from regular C-initial verbs, which take *jɨ*-, like 'to sleep' (Table 17). 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 were presented, and the verbs resistant to these innovations were identified. The set of unaffected verbs is rather small in most cases, and (proto-)languages show a considerable degree of overlap in what verbs are unaffected. Therefore, this section investigates these verbs from a comparative perspective. Section 3.2.1 treats both roots of the copula *eti/$a[p]$ 'to be', Section 3.2.2 *$ka[ti]$ 'to say', Section 3.2.3 *$ɨtə[mə]$ 'to go', and Section 3.2.4 *$(ət)jəpɨ$ 'to come'. Section 3.2.5 takes a look at *$ɨpɨtə$ 'to go down', which is resistant in Proto-Tiriyoan and Proto-Pekodian, and Section 3.2.6 investigates *e-$pɨ$ 'to bathe', of which the Proto-Pekodian reflex *$ɨpɨ$ resisted the extension of *k-. The *e*-initial verbs not affected by the extension of *k*- in Akuriyó (Section 3.1.4) will not be discussed here, as they are a large and phonologically coherent group.

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

For a comprehensive comparative overview of these two roots, readers are referred to Gildea (2018: 375–382); they will not be discussed in detail here. *$a[p]$ is the original copula and can be reconstructed as already having various irregularities in Proto-Cariban. *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.[9] Various modern languages use reflexes of these two roots in a suppletive manner, 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 affected by the extension of *k*- (Section 3.1.4), while *e?i* is not attested in a first-person form. Carijona innovated *j*-, but only in the reflex of *eti (18); the *a* 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 in certain constructions (19).

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

(18) Carijona (Robayo Moreno 1989: 177)
*ireti*bə *efʃi*nəme gərə *jefʃiɨ*
ireti-bə *efʃi-nə=me* *gərə* *j-efʃi-ɨ*
then-from be-INF=ATTRZ still 1-be-PFV
'Then I was already grown up.'

(19) Yukpa (Meira 2006b: 143–144)

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

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

Most reflexes of this verb are simply *ka*, but a fleeting syllable **ti* is reconstructed by Gildea & D. Payne (2007), best visible in the imperative forms of some languages.Table 18 shows a comparison of the longest attested forms for each language.[10]

[Table 18 about here.]

This verb was not affected by any of the extensions found in Proto-Pekodian, Proto-Waiwaian, Proto-Tiriyoan, Akuriyó, or Carijona (Sections 3.1.1 to 3.1.5). The first person form of its Yukpa reflex *ka* is unattested.

As mentioned in Section 3.1.2, Derbyshire (1985) analyzes this verb as transitive in Hixkaryána. This analytical choice is not only motivated by a desire to avoid an idiosyncratic intransitive first person prefix *ɨ-* instead of regular *kɨ-*. Hixkaryána *ka* also shows the complementary distribution of third person *n-* and preceding objects typical of transitive verbs in Cariban (Gildea 1998: 60–81). Due to the semantics of 'to say', these objects are either ideophones or direct speech (20).

(20) Hixkaryána

a. *onɨ wyaro nkekonɨ bɨryekomo, tɨyonɨ wya*
onɨ *wjaro* *n-ka-jakonɨ* *bɨrʲekomo* *tɨ-jonɨ* *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-jatʃkonɨ* *hatɨ*
MED.DEM.INAN INTS say-REM.CONT.PL HSY
'"That one there" they said.' (Derbyshire 1965: 14)

[10] Cognate segments in Tables 18 to 23 were aligned automatically with LingPy (List et al. 2021), for easier recognition of correspondences.

In (20a), the prefix *n-* occurs because there is no preceding object ('he said it like this'). It does not occur in (20b), because 'they said' is preceded by direct speech. Looking beyond Hixkaryána, at least the Tiriyó cognate shows the same pattern, albeit inconsistently so (Carlin 2004: 267).

Derivational suffixes also point to **ka*[*ti*] 'to say' being transitive: Tiriyó *ka* is characterized as the only intransitive verb being able to take the causative suffix *-po* and the agentive nominalizer *-ne* (Meira 1999: 263, 169). The exceptionality of *ka* 'to say' taking *-po* 'CAUS.TR' has also been noted for Kari'ña (Courtz 2008: 82) and Wayana (Tavares 2005: 258). The agent nominalizer **-ne* gave rise to the Panare gnomic verbal suffix *-ɲe* on transitive verbs (Gildea 1998: 184–185). The occurrence of *-ɲe* on *ka* likely led T. E. Payne & D. L. Payne (2013: 214) to categorize it as transitive, contrasting with the intransitive analysis by M.-C. Mattéi-Müller (1994: 102). Finally, reflexes of the causativizer **-metipo*, usually restricted to transitive verbs (Gildea 2015), are found with *ka* in Apalaí (E. Koehn & S. Koehn 1986: 51) and Waiwai (R. E. Hawkins 1998: 52).

The categorization of 'to say' as an intransitive verb is supported primarily by its person prefixes. Kari'ña offers a minimal pair between transitive *ka* 'to remove' and intransitive *ka* 'to say', *sikai* 'I took it away' vs *wɨkai* 'I said' (Courtz 2008: 288, 45). Similarly, 'to say' has a reflex of $1S_A$ **w-* in Pekodian languages (Section 3.1.1), rather than 1>3 *s-* (Bakairi) or **ini-* (Proto-Xinguan). Additionally, languages which differentiate direct prefixes from S_A prefixes by the presence of *i* (Meira et al. 2010: 495) show *ɨ* instead of *i* for this verb, see Akuriyó in (21a), 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 (21b), and it is probably reflected as vowel length in the Tiriyó (Meira 1999: 333) and Wayana (Tavares 2005: 196) participles.

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

b. Kari'ña (Courtz 2008: 202)
Òmakon ʻwa oti ywykàpo kaiko.
o-ʔma-kon *ʔwa* *oti* *ɨ-wɨ-ka-ʔpo* *kai-ko*
2-child-PL OBL greeting $1\text{-}S_A$-say-PST.NMLZ say-IMP
'Pass my greetings to your children.'

In summary, this verb can be reconstructed as being 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.

3.2.3 **ɨtə*[*mə*] 'to go'

This verb is reconstructed by Gildea & D. Payne (2007) as **tə*[*mə*], like **ka*[*ti*] 'to say' with a fleeting second syllable. It is true that many reflexes are clearly *t*-initial, for example Hixkaryána *ntoje* 'he went' (Derbyshire 1985: 27), Tiriyó *təkə* 'go!' (Meira 1999: 246), or Wayana *kuptəm* 'we went' (Tavares 2005: 195). However, once one considers all forms of the various reflexes of this verb (Table 19), an initial

vowel **i* must clearly be reconstructed – in contrast to unambiguously C-initial **ka*[*ti*] 'to say'.[11] This verb was not affected by the any of the extensions in Section 3.1.

[Table 19 about here.]

3.2.4 **(ət-)jəpɨ* 'to come'

[Table 20 about here.]

This verb is reconstructed as **ətepɨ* by Gildea & D. Payne (2007: 30), but an inspection of all attested reflexes (Table 20) points to a more complex story. Crucially, the majority do not reflect the **ət* part of their reconstruction, and many forms are ostensibly reflexes of **əpɨ*, **jepɨ*, or **jəpɨ*, rather than **epɨ*. A unifying account of all these forms is achieved by reconstructing a Proto-Cariban form **(ət-)jəpɨ*, morphologically segmentable into a detransitivizing prefix and a root **jəpɨ*.

Only the Pemongan languages and Werikyana point to an **j*-initial root, and 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-*. As Table 21 shows, regular *o*-initial verbs have no third person prefix, while *i-* occurs with C-initial *to*[*mo*] 'to go'. Even if 'to come' once shared this *i-*, the circumstances are strongly in favor of a *j*-initial Werikyana root, allowing the reconstruction of Proto-Cariban **jəpɨ*.

[Table 21 about here.]

Most of the (non-Pemongan) morphologically complex forms corresponding to Gildea & D. Payne's (2007) **ətepɨ* show no segmental trace of **j*, but the *i* in Akawaio *əsipɨ* is likely a reflex of the sequence **jə*. This analysis is supported by the reflex *ipɨ* from bare **jəpɨ* in very closely related Macushi.

As for the many forms seemingly reflecting **əpɨ* and **epɨ* rather than *jəpɨ*, they are distributed widely in the family, sometimes even co-occurring in the same language. A unifying account requires the root **jəpɨ* to undergo two major sound changes: a) **j*-loss; and b) **ə*-umlaut after **j*. While both sound changes are found in other contexts throughout the family (Meira et al. 2010), they appear to have applied irregularly to this verb, and not always in the same order. For example, the Kari'ña form *opɨ* can only be explained if **j* was lost before the umlaut of **ə* to **e*, which would have been triggered by **j*. On the other hand, forms like Ye'kwana *ehə* must be the result of **ə* → **e* / **j_*, with subsequent loss of **j*. The Akuriyó form *eepɨ* looks like a reflex thereof as well, but the length is unexpected, and is analyzed by Meira (1998) as a coalescence of **e* and **əe*.

While a root **jəpɨ*, the two sound changes, and the optional addition of **ət-* do account for the majority of the forms in Table 20,[12] the distribution within the family is rather chaotic. In addition

[11] As indicated by the brackets in Table 19, there are many languages where the initial vowel is only present in some forms. Also, the prefix-verb boundary in many inflected forms like e.g. Tiriyó *wɨtənne* or Arara *wɨdolɨ* 'I went' (Meira 1999: 43; Alves 2017: 153) is ambiguous, since an epenthetic *ɨ* breaks up potential CC clusters. Still, when one considers unambiguous forms, the contrast with **ka*[*ti*] becomes very clear.

[12] Apart from aforementioned Akuriyó *eepɨ*, another diachronically irregular form is Apalaí *oepɨ*, where one would expect **ət-epɨ* to yield *os-epɨ* (Meira et al. 2010: 506). Similarly, while *oepɨ* would be a regular outcome of a hypothetical **ə-jəpɨ*, the / _C allomorph of the detransitivizer is *e-* in Apalaí. One possibility is that the form is due to borrowing from Tiriyó, which has lost intervocalic **t* to create *əepɨ*. Alternatively, Apalaí *oepɨ* could be a fossilized loan from Wayana, which has replaced its reflex of **ətjəpɨ*, but where intervocalic **t* was also regularly lost (Tavares 2005: 63).

to the seemingly unordered distribution of **əpɨ* and **epɨ*, forms with and without **ət-* can be found within the same language, usually conditioned by different prefixes. This was briefly discussed in Section 3.1.1 for Arara (and Proto-Pekodian) and in Section 3.1.3 for Tiriyó (and Proto-Taranoan). To illustrate, the Tiriyó Set I paradigm shows a reflex of **ətepɨ* (< **ətjəpɨ*) for first, but of **epɨ* (< **jəpɨ*) for the other persons (22).[13] It should be noted that forms with and without **ət-* in different languages are not triggered by the same person values.

(22) Tiriyó (Meira 1999: 294)
1 *w-əepɨ*
2 *mən-epɨ*
1+2 *ke-epɨ*
3 *n-epɨ*

The interpretation of the **ət* part as a detransitivizer is based on its form and its paradigmatically conditioned occurrence in some languages. Although the combination of a detransitivizer and an intransitive verb seems semantically illogical, some historical S_P verbs are attested as adding the detransitivizer to become S_A verbs. For example, the Proto-Cariban S_P verb **wɨnɨkɨ* 'to sleep' becomes Tiriyó *əənɨkɨ* (Meira 1999: 252) and Kari'ña *ə?nɨkɨ* (Courtz 2008: 429), both S_A. Also, Waiwai 'go to sleep' can be *wɨnɨk* (R. E. Hawkins 1998: 30) or *et-wɨnɨk* (W. N. Hawkins & R. E. Hawkins 1953: 204). The parallels to 'to sleep' end here, since bare **jəpɨ* 'to come' apparently already was an S_A verb, as evidenced the class membership of its reflexes in Werikyana, Kari'ña, Arara, Tiriyó, and Panare (23).

(23) Panare (T. E. Payne & D. L. Payne 2013: 65)
ju-w-əəpɨ-n *ka=m* *kano?*
3-S_A-come-NSPEC Q=2.AUX rain
'Do you think it is gonna rain?'

Summing up, this verb is highly irregular, both from a synchronic and diachronic perspective. The scenario suggested here involves reflexes of the detransitivizer **ət(e)-* being optionally added to an S_A verb root **jəpɨ*, which further underwent umlaut and loss of **j*, but in no systematic manner, resulting in the chaotic picture in Table 20. As discussed in Section 3.1.1, innovative **k-* was introduced on the Ikpeng and Bakairi reflexes of **ətjəpɨ*, but not on the Arara reflex of **jəpɨ*. Reflexes of **ətjəpɨ* (Tiriyó) and of **ətjəpɨ*/**jəpɨ* (Akuriyó) resisted the introduction of Proto-Tiriyoan **t-*. Carijona *ehɨ* shows innovative *j-*, rather than conservative *w-* (24). It is unknown whether there is a Yukpa reflex of this verb, and it was fully replaced in Proto-Waiwaian by **omokɨ* 'to come'.

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

[13] While the 1+2 form is a regular outcome of **kɨt-epɨ*, the second person form is mysterious (Meira 1998: 115).

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). The resistance against the former extension was subsequently overcome in Bakairi; 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. One might think that it was also affected by the extensions of *j-* in Carijona (25a) and Yukpa (25b).

(25) a. Carijona (David Felipe Guerrero, p.c.)
irə *waʧinakano* *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 broader comparative perspective reveals a much more complicated story (Table 22).[14] It turns out that 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.

[Table 22 about here.]

The verb shows traits of both classes in Wayana, necessitating an analysis as a "mixed" verb in a synchronic description of that language. 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 (26a), but behaves like an S_P verb in taking a second person prefix in imperatives (26b).

(26) 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!'

[14] In Table 22, parenthesized forms indicate uncertainty about cognacy status. The reconstruction of Proto-Pekodian **ɨptə* treats the suffixed elements found in the 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 **ənə* '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).

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 a scenario where the verb was a regular member of the S_P class in pre-Wayana, but partially switched to the S_A class, taking a $1+2S_A$ prefix and the S_A class marker. This in turn implies that S_A reflexes of this verb in other languages fully switched from S_P in their inflectional patterns.

Wayana-external comparative evidence supports this hypothesis: The Arara causativized form is *eniptoŋ* (Alves 2017: 66), and Kari'ña has a cognate form *eniʔto* (Courtz 2008: 263); *oniʔto* 'to go down' in Table 22 is a detransitivized form thereof, lit. 'to get oneself down'. Both causativized forms contain a reflex of the transitivizer **en-*, which was usually found with S_P verbs (Gildea & Cáceres in preparation). Tiriyó *ɨhtə* has irregular causativized forms that also feature a reflex of **en-* (Meira 1999: 263). Thus, it appears that this verb was originally S_P, but then switched its class in four "and a half" languages of the family, for so far unknown reasons.

These circumstances make it impossible to answer the question of whether 'to go down' was affected by some extensions. For Proto-Tiriyoan, one cannot establish a relative chronology of the verb class switch, the introduction of the idiosyncratic marker **p-*, and the extension of **t-*. For Proto-Waiwaian, both its first person form and its inflectional class are unknown. For Carijona and Yukpa, one cannot rule out the possibility that the verb switched classes before the breakdown of the split-S system. While no language-internal evidence supports that scenario, 'to go down' clearly has an inclination to switch classes; in the case of Carijona, that could have already happened at the Proto-Taranoan stage. In all four cases, the verb could have had S_A status at the time of the extension, resisting it and preserving the old prefix, but it could also have had S_P status and thus not even have been a potential target.

On the other hand, the class switch happened before the split of Proto-Tiriyoan, and therefore this verb resisted the extension of Akuriyó *k-* as an S_A verb. Likewise, it is very likely that the class switch 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, trivially, Akuriyó *k-* (Section 3.1.4). It took on new $1S_A$ prefixes in Proto-Tiriyoan (Tiriyó *s-epi-*, Akuriyó *ʧe+pi-* [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 usually typical S_A verbs in Cariban languages, derived with a detransitivizer from a transitive root. These roots are reflexes of **pɨ*, or **kupi* in some Venezuelan languages (Table 23). As was shown in Section 3.1.1, Proto-Pekodian can be reconstructed as having the pair **ɨpɨ* (INTR) / **ɨp(ɨ)* (TR). Thus, while Proto-Pekodian 'to bathe (TR)' has perfectly regular cognates in other languages of the family, intransitive 'to bathe' is divergent in this branch, changing **e-* to **i-*. This is an irregular development, since there are no attested reflexes of the detransitivizer in Pekodian that would go back to Proto-Pekodian **i-* (Meira et al. 2010: 506); its cause is unknown. However, it should be noted that other languages also show unexpected developments in this verb, considering the glide insertions in Waiwaian or the distribution of **pɨ* and **kupi* in Venezuelan languages.

[Table 23 about here.]

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, where possible. Table 24 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 at all or just not in first person forms (–), in others the first person form is unknown (?), and in the case of 'to go down' the question of affectedness is not meaningfully answerable (N/A), for any of the reasons discussed in Section 3.2.5. Every ✓ stands for a verb affected by an extension, while × represents conservatively inflected verbs. This overview makes clear just how pervasive the tendency for these verbs to resist innovative markers is, as they do so in different languages.

[Table 24 about here.]

It is astonishing that the same 1-7 verbs retain their old first person marker in 6 distinct developments, while a plethora of regular $S_{(A)}$ verbs take on innovative markers. This suggest that there is some strong motivation for these verbs to not be affected by innovative markers. The question arises what properties unite these verbs and make them so conservative across different Cariban languages. Possible answers to this question will be discussed in Section 4, using Bybee's (1985) network model of morphology.

4 Explaining conservativeness: a network morphology approach

As stated in Section 3.3, the fact that the same verbs retain conservative first person prefixes in six individual innovations is in need of explanation. Perhaps the most well-known contribution regarding 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 classic 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 is attracting new verbs in certain dialects, like *sniːk–snʌk* or *brɪŋ–brʌŋ* (Bybee 1985: 129–130). These verbs are recruited based on the lexical connection they form with prototypical members of the group, and accordingly develop irregular or "strong" past tense forms.

As possible bases of these connections between lexemes, Bybee (1985: 118) suggests the criteria of semantic, phonological, and morphological similarity; the English strong verbs are an example for a phonologically motivated network. Another important factor in the model is frequency, since more frequent words have a higher lexical strength (Bybee 1985: 119). This higher lexical strength diminishes the influence from other lexemes, meaning that high-frequency items are more likely to resist innovations. For the Cariban inflectional patterns under investigation, the model would predict that a) semantically/phonologically/morphologically similar verbs will be affected by person marker extensions, and b) high-frequency verbs will tend to resist these extensions and thus remain conservative.

When one considers the groups of verbs with innovative first person markers (those not in Table 24), one can easily perceive several factors which could serve as the thread connecting a lexical network. Perhaps the most obvious similarity is that they all have a reflex of the detransitivizer **ətə/e-*, a hallmark of S_A verbs (Section 2.3). This also has a phonological consequence: all derived S_A verbs begin with reflexes of **ə* or **e*, meaning that networks with a phonological basis are also plausible. A more trivial connection between these other verbs is that they are all S_A verbs, and thus share inflectional (morphological) patterns. In the following analysis, the $1S_A$ (pre-innovation) prefix will serve as a proxy for fhese inflectional patterns. This is made possible by the fact that S_A verbs all had a reflex of **w-* at the time of innovation, making the first person prefix representative for the inflectional class. The exception to this is Akuriyó, where innovative $1S_A$ *k-* and idiosyncratic *p-* on 'to go down' created inflectional subclasses of S_A verbs. There are no semantic patterns that would be obvious, which is not unsurprising given the lack of semantic patterns in the split-S system overall (Section 2.3). Thus, for each extension, there are three hypotheses as to what connected the members of the responsible network: a reflex of DETRZ, their stem-initial phoneme, or a specific $1S_A$ prefix.

It is intuitively obvious that many of the conservative verbs in Table 24 are high-frequency verbs, which would mean high lexical strength and conservativeness according to the network model. A major obstacle to confirming this intuition is 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 supported neither by frequency counts nor accompanied by a list of verbs, but the unmentioned verbs likely correspond to the five first columns of Table 24, all underived S_A verbs in Kari'ña. To ameliorate the dearth of data, a count of S_A verbs in three glossed texts from E. Koehn & S. Koehn (1994) will serve as a second source of frequency information, the results of which are shown in Table 25. The Apalaí data agree with the above interpretation of Courtz's (2008) claim; defining "high frequency" as having an above average count produces the exact same five verbs. While it is not at all certain that the interpretation of Courtz's (2008) claim and the small Apalaí sample are representative of discourse patterns in the Cariban (proto-)languages under discussion, their use as a tool for identifying high-frequency verbs is necessitated by the absence of alternatives.

[Table 25 about here.]

Thus, in addition to the three hypotheses for possible network factors, each can be combined with frequency; high-frequency verbs are predicted to not undergo innovation, even though the factor in question would put them in the same lexical network as regular S_A verbs. This leaves each extension with six possible explanations for which verbs are affected and which are not. To test the hypotheses, for each explanation the expected behavior of each verb was predicted, illustrated in Table 26 for Proto-Tiriyoan. For example, **e?i* 'to be' is expected to participate in innovations spreading in a phonologically defined network, being **e-*initial, as well as in an inflectionally defined one, since it shared **w-* with other S_A verbs. However, it would not have belonged to a network defined by the presence of a detransitivizer. If frequency is taken into account, it is expected to remain conservative regardless of the nature of the network. These predictions were then checked against the data in Table 24, to see how many potentially conservative verb each explanation predicted correctly. This yielded a score of

what proportion of potentially conservative verbs had their behavior accurately predicted, illustrated for Proto-Tiriyoan in Table 27 and summed up in Table 28.

[Table 26 about here.]

[Table 27 about here.]

[Table 28 about here.]

It is important to understand that the scores in Table 28 only refer to the group of seven verbs in Table 24, i.e. those that are attested as resisting at least one extension. For each extension, there were also many run-of-the-mill S_A verbs, all taking on the new person marker, except for the Akuriyó *e*-initial verbs.[15] To illustrate, if one adds 1'000 regular S_A verbs per language to the data – a conservative estimate based on Courtz's (2008) Kari'ña dictionary – all six explanations consistently predict the behavior of 99.99+% verbs correctly. However, the available data simply does not allow such large-scale tests for Cariban languages, so the present investigation is restricted to the edge cases.

The extent of the extensions in both Proto-Waiwaian and Proto-Pekodian is fully predicted by the presence or absence of a detransitivizer. In both cases, only the underived[16] S_A verbs were not affected, all other S_A verbs taking **k*-. Not shown in Table 28 are subsequent evolutions in the Pekodian daughter languages, which can largely be argued to also be due to the detransitivizer: First, both Ikpeng and Bakairi regularized the paradigm to use forms with detransitivizer for first person (Section 3.1.1), and both introduced *k*-.[17] Second, the subsequent introduction of *k*- to Ikpeng *aran* 'to go' (< **ɨtən*) suggests a potential reanalysis of *ar* as a detransitivizer.

Three extensions are fully predicted by phonological criteria, those in Akuriyó, Carijona, and Yukpa. As already discussed, Akuriyó *k*- only appears on *ə*-initial verbs (Section 3.1.4). In Carijona, the extension of *j*- affected all *e*- and *ə*-initial verbs, including *eh* 'to come' or *eʃʃi* 'to be', which do not have a detransitivizing prefix. Only *ka* 'say', *təmə* 'go', and *a* 'be-1' did not take on *j*-. Similarly, the extension of Yukpa *j*- can succinctly be characterized as affecting all vowel-initial verbs; 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 only works if *t*- and *ʃʃ*-, which were of course also phonologically conditioned, are analyzed as distinct morphemes.

When additionally considering putative conservatory frequency effects, prediction accuracy was improved in 8 cases, stagnated in 7 cases, and worsened in 3 cases. The three cases where the tentative model of verb frequency arrives at incorrect predictions 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'. Including

[15] While there are a few S_A verbs not transparently derived from transitive verbs (Meira 1999: 252, 2000: 222; Gildea & D. Payne 2007: 30), which are not featured in Table 24, these are mostly **ə*-initial and were likely productively derived at some point. 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 being an S_A verb at the time of a person marker extension, they are not relevant for this study.

[16] Note that for Proto-Pekodian, the idiosyncratic evolution of **e-pɨ* 'to bathe (INTR)' to **ipɨ* made the verb morphologically opaque.

[17] 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.

frequency in the model led to overall improvements, up to 100% prediction accuracy for all three potential factors in Proto-Tiriyoan, as well as for the phonological criterion in Proto-Pekodian and the inflection criterion in Proto-Waiwaian.

Overall, one must 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 lexical network with a phonological basis emerges as the unambiguous winner, while frequency-based explanations fare much 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 in turn is due to the fact that three of the factors used in the model – detransitivizer, phonology, frequency – largely converge 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 Discussion

The first research question of this study asked where the irregularly inflected first person forms in some Cariban languages came from. A second question was what verbs are irregular in what languages. Finally, reasons for their irregularness were sought.

The main findings can be summarized as follows: Verbs irregularly inflected for first person are conservative, leftovers of person marker extensions which left some verbs untouched. Some of these extensions are reconstructible to proto-languages, while others happened in pre-modern stages of single languages. Conservatively inflected verbs show a great degree of overlap between languages, and often behave irregularly in other ways, too. While Bybee's (1985) network model offers eplanations for the lexical extent of innovative markers, in 4 of 6 cases it gives no unambiguous answer, as multiple factors 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 prefix **ətə/e-*.

While the origins of the split-S system clearly have to do wih the detransitivizer, the question why **ətə/e-*is associated A-oriented prefixes (Meira 2000) is still not answered. Another question to be asked 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', 'say', 'go'). There is no reason why they could not be simple S_P verbs, except maybe in the case of **ka*[*ti*] 'to say' with its transitive tendencies. The answer might be found in the origins of **ətə/e-*and its association with A-oriented prefixes.

As for Bybee's (1985) network model of morphology, it fulfilled the promise of delivering attractive explanations for irregularities in inflectional patterns. If anything, it overshot its goal somewhat; only two of the four investigated innovations had an unambiguous answer – phonology. It may be noted that across all six hypotheses, the two featuring phonology (one with, once without frequency) together correctly predicted 100% of patterns. Since frequency decreased prediction accuracy in the case of Carijona and Yukpa, it would be interesting to see S_A verb frequency statistics from corpora of these languages. However one interprets the results of the network model investigation, it needs to be pointed out that the ambiguity of its answers are due to the special circumstances in Proto-Cariban, rather than a flaw in the model. Still, the results raise the question how applicable the model is in circumstances where ambiguity arises.

As mentioned, other 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 providing better input for the network model. Apart from this specific purpose, corpus-based investigations like Sapién et al. (2021) in Cariban languages are direly needed, to accompany the improving descriptive side. Concerning the latter, more extensive descriptive work on Yukpa and Carijona would not only tremendously benefit this paper, but Cariban studies as a whole.

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.

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: Kaʔpon-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 (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.

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.

Meira, Sérgio (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.
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.
Schuring, Phillip Arthur (n.d.). "Uma gramática morfossintática bâsica da língua Katxuyana". MS.
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.
Yamada, Racquel-María (2011). “A New Approach to *ky-* and *-ng* in Kari’nja: Evidentiality or Something Else?” In: *International Journal of American Linguistics* 77.1: 59–89.

A Predictions for the behavior of individual verbs

[Table 29 about here.]

[Table 30 about here.]

[Table 31 about here.]

[Table 32 about here.]

[Table 33 about here.]

[Table 34 about here.]

[Table 35 about here.]

[Table 36 about here.]

[Table 37 about here.]

[Table 38 about here.]

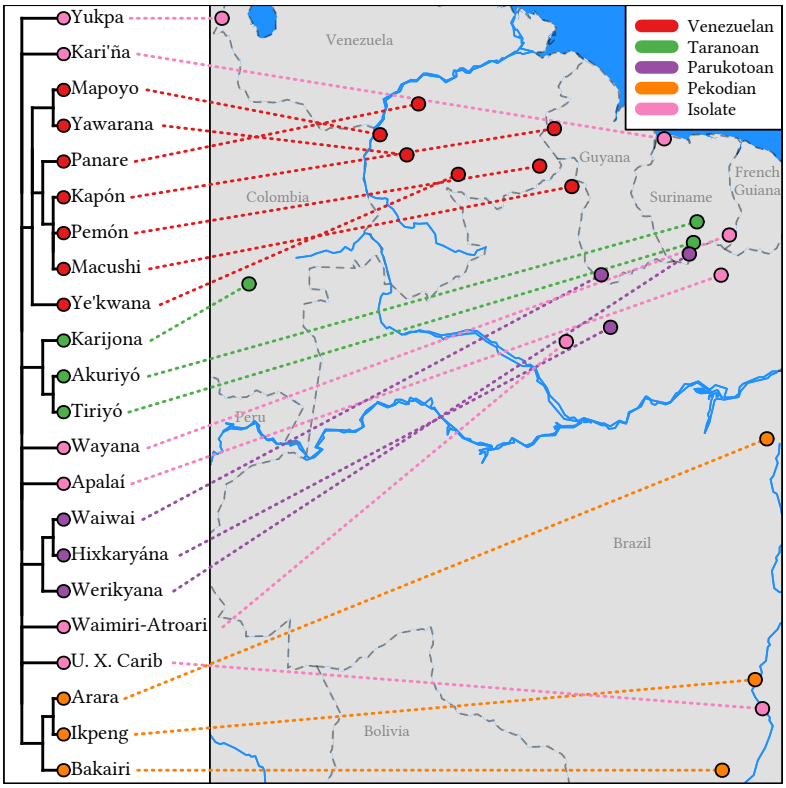

Figure 1: The Cariban language family. Classification by Matter (2021a), positions from Hammarström et al. (2020). Created with a modified version of `phytools` (Revell 2012).

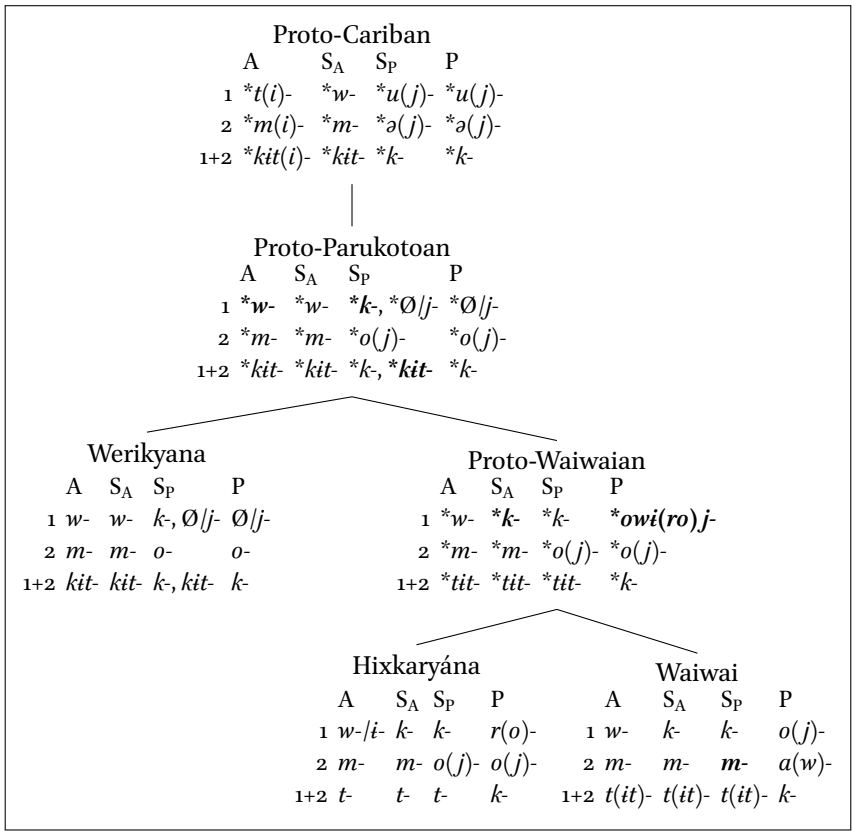

Figure 2: Person marking extensions in Parukotoan, after Gildea (1998: 94)

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-oserʲehɨ-*	*k-atarʲeknohɨ-*	*k-atama-*	*w-eʃe-*
2	*m-ehurka-*	*m-oserʲehɨ-*	*m-atarʲeknohɨ-*	*m-atama-*	*m-eʃe-*
1+2	*t-ehurka-*	*t-oserʲehɨ-*	*t-atarʲeknohɨ-*	*t-atama-*	*t-eʃe-*
3	*ɲ-ehurka-*	*n-oserʲehɨ-*	*n-atarʲeknohɨ-*	*n-atama-*	*n-eʃe-*

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-əəniki-*	*t-əene-*	*s-epi-*	*s-entapo-*	*wi-tən-*
2	*m-əəniki-*	*m-əene-*	*m-epi-*	*m-entapo-*	*mi-tən-*
1+2	*kit-əəniki-*	*k-əene-*	*ke-epi-*	*k-entapo-*	*ki-tən-*
3	*n-əəniki-*	*n-əene-*	*n-epi-*	*n-entapo-*	*ni-tən-*

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)-

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)-*[1]	
3	*n(ɨ)-*	

[1] Due to the wholesale loss of 1+2 as an person value.

(c) Waimiri-Atroari

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

Table 5: Participles of S_A and S_P verbs (Schuring n.d.: 39; Alves 2017: 118, 207; Meira 1999: 333, 334; Tavares 2005: 400; E. Koehn & S. Koehn 1986: 35, 1994: kuruaz-154; Hoff 1968: 430, 433; T. E. Payne & D. L. Payne 2013: 232, 244)

Language	S_A	S_P
Werikyana	*t-ehurka-ʧe* 'fallen'	*ti-jaʔ-so* 'burnt'
Arara	*t-**o**-ep-te* 'come'	*t-oregrum-te* 'sad'
Tiriyó	*tɨ-**w**-əturu-e* 'talked'	*t-əpəə-se* 'tired'
Wayana	*tə-**w**-epɨ-he* 'bathed'	*t-onopɨ-he* 'painted'
Apalaí	*t-**o**-ɨto-se* 'gone'	*t-ɨhto-se* 'gone down'
Kari'ña	*tu-**w**-oʔka-se* 'come out'	*t-okari-se* 'told'
Panare	*t-**o**-tatɨhpə-se* 'wailed'	*tɨ-sɨrɨke-ʧe* 'tired'

Table 6: Nominalizations of S_A and S_P verbs (Schuring n.d.: 49, 74; Alves 2017: 97; Meira 1999: 246; Tavares 2005: 130, 409; E. Koehn & S. Koehn 1986: 90, 1994: ner2-003; Hoff 1968: 135, 392; T. E. Payne & D. L. Payne 2013: 390; M.-C. Mattéi-Müller 1994: 23)

Language	S_A	S_P
Werikyana	*o-**w**-ehurka-tpɨrɨ* 'your fall'	*o-onenmehɨ-tpɨrɨ* 'your waking up'
Arara	***w**-orik-tubo* 'dancing place'	*ereŋmɨ-tpo* 'killing instrument'
Tiriyó	*ji-**w**-əturu-to* '(for) my talking'	*j-emamina-to* '(for) my playing'
Wayana	*ɨ-**w**-əturu-topo* 'my story'	*j-ɨnɨkɨ-topo* 'my object for sleeping'
Apalaí	*j-epɨ-topo* 'my bathing place'	*j-enuru-topõ-pɨrɨ* 'the place of my birth'
Kari'ña	*a-**w**-ekupi-rɨ* 'your taking a bath'	*aj-ereʔna-∅* 'your fainting'
Panare	*j-**u**-ʧireema-n* 'their eating'	*tj-arunkampəti-n* 'his hair standing on end'

Table 7: Imperatives of S_A and S_P verbs (Derbyshire 1965: 44, 89; Alves 2017: 161; Meira 1999: 323; Tavares 2005: 227; E. Koehn & S. Koehn 1986: 62, 1994: Mopo/20; Hoff 1968: 190; M.-C. Mattéi-Müller 1994: 5, 17)

Language	S_A	S_P
Hixkaryána	*omoh-ko* 'come!'	***oj**-okajɨm-ko* 'go up!'
Arara	*odotpot-ko* 'come back!'	***o**-alum-ko* 'jump!'
Tiriyó	*epɨ-kə* 'bathe!'	***ə**-eremina-kə* 'sing!'
Wayana	*əməm-kə* 'enter!'	***əw**-eremi-kə* 'sing!'
Apalaí	*otuʔ-ko* 'eat!'	***o**-nɨʔ-ko* 'sleep!'
Kari'ña	*oʔmaʔ-ko* 'stop!'	***aj**-awon-ko* 'get up!'
Panare	*ape-ʔ* 'flee!'	*ahpən-kə* 'jump!'

Table 8: 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-*

Table 9: Verbs preserving $1S_A$ **w-* in Proto-Pekodian (Alves 2017: 153, 200; von den Steinen 1892: 76, 130, 131, 374, 397**aŋ-ka**; Pachêco 2001: 42, 80, 139, 209, 1997: 68; Meira 2003a: 4)

	Proto-Pekodian	Arara	Ikpeng	Bakairi
'be-1'	**w-ap-*	*w-ap-*	–	*w-a-*
'be-2'	**w-eʧi-*	*w-itʃi-*	∅*-eʧi-*	*w-i-*
'say'	**wɨ-ge-*	*wɨ-ge-*	*ɨ-ge-*	*u-ge-*
'go'	**w-ɨtən-*	*w-ɨdo-*	*k-aran-*	*u-tə-*
'come'	**w-epɨ-*	*w-ebɨ-*	*k-arep-*	*k-əewɨ-*
'go down'	**w-ɨptə-*	*w-ɨptoŋ-*	*?-ɨptoŋ-*	*k-ɨtəgɨ-*
'bathe'	**w-ɨpɨ-*	*w-ɨbɨ-*	∅*-ɨp-*	*w-ɨ-*

Table 10: 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*

Table 11: Regular 'to fall' (S_A) and 'to sleep' (S_P) in Proto-Waiwaian (W. N. Hawkins & R. E. Hawkins 1953: 209, 211; Howard 2001: 150; Derbyshire 1985: 189, 190, 196; R. E. Hawkins 1998: 30)

	Proto-Waiwaian 'to fall'	'to sleep'	Hixkaryána 'to fall'	'to sleep'	Waiwai 'to fall'	'to sleep'
1	**k-eɸurka-*	**ki-winiki-*	*k-ehurka-*	*ki-niki-*	*k-eɸirka-*	*ki-winiki-*
2	**m-eɸurka-*	**o-winiki-*	*m-ehurka-*	*o-wniki-*	*m-eɸirka-*	*mi-winiki-*
1+2	**t-eɸurka-*	**tit-winiki-*	*t-ehurka-*	*ti-niki-*	*ʧ-eɸirka-*	*tit-winiki-*
3	**ɲ-eɸurka-*	**ni-winiki-*	*ɲ-ehurka-*	*ni-niki-*	*ɲ-eɸirka-*	*ni-winiki-*

Table 12: Verbs preserving $1S_A$ *w- in Proto-Waiwaian (Derbyshire 1985: 70, 197, 198, 1979: 4; R. E. Hawkins 1998: 71, 85; p.c., Spike Gildea)

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

Table 13: Regular Proto-Tiriyoan S_A verbs (Meira 1999: 292, 294; Gildea 1994: 87)

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

Table 14: Verbs preserving $1S_A$ *w- in Proto-Tiriyoan (Meira 1998: 112, 113, 114, 115, 165, 1999: 292, 294, 339)

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

Table 15: Regular Akuriyó 1S$_A$ markers (Gildea 1994: 77, 79, 82, 84, 85, 86, 87)

first person *k-*	first person *ʧ-*
əempa- 'to learn'	*epɨ-* 'to bathe (INTR)'
əəʧena- 'to cry'	*ekɨrika-* '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'	

Table 16: 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 17: Regular Yukpa verbs (Largo 2011: 72, 76; Meira 2006b: 139)

	'to sleep'	'to fall'	'to wash self'
1	*ji-ni-*	*j-ata-*	*j-otum-*
2	*mi-ni-*	*m-ata-*	*m-otum-*
3	*ni-ni-*	*n-ata-*	*n-otum-*

Table 18: Reflexes of *ka[ti] 'to say' (Meira 2005: 267, 2003a: 4; Franchetto 2008: 48; Pachêco 2001: 209, 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	*ke*	k	e		
Bakairi	*ge*	g	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		

Table 19: 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; 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	**ɨtən*	ɨ	t	ə	n	
Arara	*ɨdo*	ɨ	d	o		
Arara	*to*		t	o		
Ikpeng	*aran*	a	r	a	n	
Ikpeng	*ero*	e	r	o		
Bakairi	[*ɨ*]*tə*	ɨ	t	ə		
Proto-Tiriyoan	**təmɨ*		t	ə	m	ɨ
Tiriyó	*tə*[*n*]		t	ə	n	
Akuriyó	[*ə*]*tə*[*mɨ*]	ə	t	ə	m	ɨ
Carijona	*təmə*		t	ə	m	ə
Wayana	[*ɨ*]*tə*[*m*]	ɨ	t	ə	m	
Apalaí	*ɨto*	ɨ	t	o		
Kari’ña	*to*		t	o		
Kari’ña	[*ɨ*]*ʔ*	ɨ	ʔ			
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		

Table 20: Reflexes of *(*ət-*)*jəpɨ* '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; Méndez-Arocha 1959: 68; Courtz 2008: 429; Meira & Franchetto 2005: 182; p.c., Spike Gildea)

Language	Form							
Werikyana	*o*[*o*]*hɨ*					oo	h	ɨ
Werikyana	*johɨ*				j	o	h	ɨ
Werikyana	*ehɨ*					e	h	ɨ
Proto-Pekodian	**epɨ*					e	p	ɨ
Proto-Pekodian	**ədepɨ*	ə	d	-		e	p	ɨ
Arara	*ep*					e	p	
Arara	*odebɨ*	o	d	-		e	b	ɨ
Arara	*ebɨ*					e	b	ɨ
Ikpeng	*arep*	a	r	-		e	p	
Bakairi	*əewɨ*	ə		-		e	w	ɨ
Proto-Tiriyoan	**əʔepɨ*	ə	ʔ	-		e	p	ɨ
Proto-Tiriyoan	**epɨ*					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	*opɨ*					o	p	ɨ
Ye'kwana	*ehə*					e	h	ə
Akawaio	*jepɨ*				j	e	p	ɨ
Akawaio	*əsɨpɨ*	ə	s	-		i	p	ɨ
Ingarikó	*jepə*				j	e	p	ə
Ingarikó	*jə*				j	ə		
Patamona	*jəpɨ*				j	ə	p	ɨ
Patamona	*jepɨ*				j	e	p	ɨ
Pemón	*jepɨ*				j	e	p	ɨ
Panare	*əpɨ*					ə	p	ɨ
Yawarana	*əpɨ*					ə	p	ɨ
Mapoyo	*epɨ*					e	p	ɨ
Upper Xingu Carib	*ee*					ee		

Table 21: Werikyana *johɨ* 'to come' compared with other verbs (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-*	*ɨ-to-*

Table 22: 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	*oniʔto*	(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					

Table 23: Comparison of intransitive and transitive 'to bathe' (Derbyshire 1979: 198; R. E. Hawkins 1998: 192, 203; Alves 2017: 150, 162; 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	*ɨp*	ɨ	p	
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 24: Overview of extensions and (un-)affected verbs

	*ka[ti] 'say'	*ɨtə[mə] 'go'	*a[p] 'be-1'	*eti 'be-2'	*(ət-)jəpɨ '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

Table 25: 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-ereh '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-ɨrɨʔka 'land'	1	0.61%
ot-ɨʔka 'finish'	1	0.61%
ot-uru 'talk'	1	0.61%
at-apiaka 'divide up'	1	0.61%
e-sɨrɨʔma 'move'	1	0.61%

Table 26: Predictions for Proto-Tiriyoan

	**a* 'be'	**eʔi* 'be'	**əʔepɨ* 'come'	**təmɨ* 'go'	**ka* 'say'	**epɨ* 'bathe (INTR)'
DETRZ	×	×	✓	×	×	✓
DETRZ+freq	×	×	×	×	×	✓
phono (/ _*ə, *e*)	×	✓	✓	×	×	✓
phono+freq	×	×	×	×	×	✓
infl (**w*-)	✓	✓	✓	✓	✓	✓
infl+freq	×	×	×	×	×	✓

Table 27: Evaluating predictions for Proto-Tiriyoan

	*a 'be'	*eʔi 'be'	*əʔepɨ 'come'	*təmɨ 'go'	*ka 'say'	*epɨ 'bathe (INTR)'	Score
DETRZ+freq	✓	✓	✓	✓	✓	✓	100.0%
phono+freq	✓	✓	✓	✓	✓	✓	100.0%
infl+freq	✓	✓	✓	✓	✓	✓	100.0%
DETRZ	✓	✓	×	✓	✓	✓	83.3%
phono	✓	×	×	✓	✓	✓	66.7%
infl	×	×	×	×	×	✓	16.7%

Table 28: 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-*	83.3%	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 29: Predictions for Proto-Waiwaian

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

Table 30: Evaluating predictions for Proto-Waiwaian

	*ka[s] 'say'	*eʃi 'be'	*ah 'be'	*[i]to[m] 'go'	*eeɸi 'bathe (INTR)'	Score
DETRZ	✓	✓	✓	✓	✓	100.0%
DETRZ+freq	✓	✓	✓	✓	✓	100.0%
phono+freq	✓	✓	✓	✓	✓	100.0%
infl+freq	✓	✓	✓	✓	✓	100.0%
phono	✓	×	×	✓	✓	60.0%
infl	×	×	×	×	✓	20.0%

Table 31: Predictions for Proto-Pekodian

	*ap 'be'	*eʧi 'be'	*epɨ 'come'	*ɨtən 'go'	*ɨptə 'go down'	*ke 'say'	*ipɨ 'bathe (INTR)'
DETRZ	×	×	×	×	×	×	×
DETRZ+freq	×	×	×	×	×	×	×
phono (/ _*ə, e)	×	✓	✓	×	×	×	×
phono+freq	×	×	×	×	×	×	×
infl (*w-)	✓	✓	✓	✓	✓	✓	✓
infl+freq	×	×	×	×	✓	×	✓

Table 32: Evaluating predictions for Proto-Pekodian

	*ap 'be'	*eʧi 'be'	*epɨ 'come'	*ɨtən 'go'	*ɨptə 'go down'	*ke 'say'	*ipɨ 'bathe (INTR)'	Score
DETRZ	✓	✓	✓	✓	✓	✓	✓	100.0%
DETRZ+freq	✓	✓	✓	✓	✓	✓	✓	100.0%
phono+freq	✓	✓	✓	✓	✓	✓	✓	100.0%
phono	✓	×	×	✓	✓	✓	✓	71.4%
infl+freq	✓	✓	✓	✓	×	✓	×	71.4%
infl	×	×	×	×	×	×	×	0.0%

Table 33: Predictions for Akuriyó

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

Table 34: Evaluating predictions for Akuriyó

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

Table 35: Predictions for Carijona

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

Table 36: Evaluating predictions for Carijona

	eh[ɨ] 'come'	*ka* 'say'	*eʧi* 'be'	*təmə* 'go'	*a* 'be'	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'	*e* 'be'
DETRZ	×	×	×
DETRZ+freq	×	×	×
phono (/ _V)	×	✓	✓
phono+freq	×	×	×
infl (**w-*)	✓	✓	✓
infl+freq	×	×	×

Table 38: Evaluating predictions for Yukpa

	to 'go'	*a* 'be'	*e* 'be'	Score
phono	✓	✓	✓	100.0%
infl	×	✓	✓	66.7%
DETRZ	✓	×	×	33.3%
DETRZ+freq	✓	×	×	33.3%
phono+freq	✓	×	×	33.3%
infl+freq	✓	×	×	33.3%