

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

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

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

Contents

1	Introduction	1
2	The origins of irregular first person inflections	2
2.1	Proto-Cariban person marking and inflectional relics	2
2.2	Person marker extensions and lexical diffusion	3
2.3	The Cariban split-s system	4
3	Prefixes and verbs: innovation and resistance	6
3.1	Incomplete extensions: the innovative 1S _A markers	6
3.1.1	Proto-Pekodian * <i>k</i> -	6
3.1.2	Proto-Waiwaian * <i>k</i> -	9
3.1.3	Proto-Tiriyoan * <i>t</i> -	10
3.1.4	Akuriyó <i>k</i> -	11
3.1.5	Carijona <i>j</i> -	12
3.1.6	Yukpa <i>j</i> -	13
3.2	Conservative verbs in comparison	13
3.2.1	* <i>ka</i> [<i>ti</i>] ‘to say’	14
3.2.2	* <i>itə</i> [<i>mə</i>] ‘to go’	15
3.2.3	* <i>eti</i> and * <i>a</i> [<i>p</i>] ‘to be’	17
3.2.4	*(<i>ət</i> -) <i>epi</i> ‘to come’	17
3.2.5	* <i>ipitə</i> ‘to go down’	19
3.2.6	* <i>e-pi</i> ‘to bathe’	21
3.3	Summary	21
4	Explaining conservativeness: a network morphology approach	23
5	Conclusion	26
6	Abbreviations	26
7	Acknowledgments	28
A	Predictions for the behavior of individual verbs	32

List of Figures

1	The Cariban language family	1
2	Person marker extensions in Parukotoan	4

List of Tables

1	Some Hixkaryána verbs	1
2	Some Tiriyo verbs	2
3	Proto-Cariban Set I (main clause) person markers	3
4	Some examples for completed extensions	5
5	Regular Pekodian S _A verbs	7
6	Verbs preserving 1S _A *w- in Proto-Pekodian	8
7	Loss of *w in Ikpeng	9
8	Regular Proto-Waiwaian verbs	9
9	Verbs preserving 1S _A *w- in Proto-Waiwaian	10
10	Regular Proto-Tiriyoan S _A verbs	11
11	Verbs preserving 1S _A *w- in Proto-Tiriyoan	11
12	Regular Akuriyó 1S _A markers	12
13	Regular Carijona verbs	12
14	Regular Yukpa verbs	13
15	Reflexes of *ka[ti] ‘to say’	14
16	Reflexes of *itə[mə] ‘to go’	16
17	Reflexes of *(ət-)epi ‘to come’	18
18	*(ət-)epi ‘to come’ in paradigms	18
19	Werikyana S _A verbs in the Progressive	19
20	Reflexes of *ipitə ‘to go down’	20
21	Comparison of intransitive and transitive ‘to bathe’	22
22	Overview of extensions and (un-)affected verbs	22
23	Frequency counts of S _A verbs in Apalaí	24
24	Predictions for Proto-Tiriyoan	24
25	Evaluating predictions for Proto-Tiriyoan	25
26	Overview of prediction accuracy	25
27	Predictions for Proto-Waiwaian	32
28	Evaluating predictions for Proto-Waiwaian	33
29	Predictions for Proto-Pekodian	33
30	Evaluating predictions for Proto-Pekodian	33
31	Predictions for Proto-Tiriyoan	33
32	Evaluating predictions for Proto-Tiriyoan	34
33	Predictions for Akuriyó	34
34	Evaluating predictions for Akuriyó	34
35	Predictions for Carijona	34
36	Evaluating predictions for Carijona	35
37	Predictions for Yukpa	35
38	Evaluating predictions for Yukpa	35

1 Introduction

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

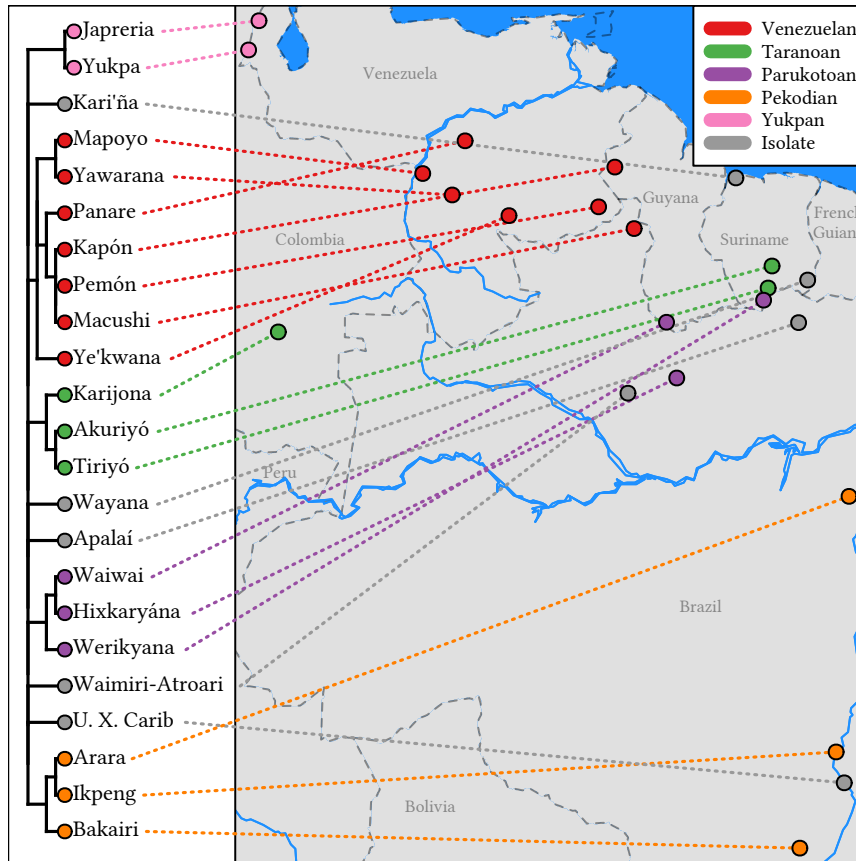


Figure 1: The Cariban language family. Classification from Matter (2021a), positions from Hammarström et al. (2020). Made with a modified version of phytools (Revell 2012)

In some Cariban languages, a small group of verbs show a divergent first person inflection pattern, a topic which has not received much attention in the literature. This is illustrated for Hixkaryána in Table 1,¹ with

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	<i>k-ehurka-</i>	<i>k-oserjehi-</i>	<i>k-atarjeknohi-</i>	<i>k-atama-</i>	<i>w-eje-</i>
2	<i>m-ehurka-</i>	<i>m-oserjehi-</i>	<i>m-atarjeknohi-</i>	<i>m-atama-</i>	<i>m-eje-</i>
1+2	<i>t-ehurka-</i>	<i>t-oserjehi-</i>	<i>t-atarjeknohi-</i>	<i>t-atama-</i>	<i>t-eje-</i>
3	<i>n-ehurka-</i>	<i>n-oserjehi-</i>	<i>n-atarjeknohi-</i>	<i>n-atama-</i>	<i>n-eje-</i>

¹The presence of a 1+2 person value implies that of a 1+3 value. This is usually expressed with a free pronoun combined with third person morphology in Cariban languages, so it is not represented as a distinct value in Table 1 and other paradigm tables. Further, TAM suffixes in the attested forms are omitted, since a) the focus lies on the prefixes and stems, and b) full paradigms with the same TAM suffix are rare in the available sources. Standard IPA symbols are used in the transcription of Cariban languages, except for coronal rhotics,

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

	‘to sleep’	‘to see self’	‘to bathe (INTR)’	‘to yawn’	‘to go’
1	<i>t-æniki-</i>	<i>t-æene-</i>	<i>s-epi-</i>	<i>s-entapo-</i>	<i>wi-tən-</i>
2	<i>m-æniki-</i>	<i>m-æene-</i>	<i>m-epi-</i>	<i>m-entapo-</i>	<i>mi-tən-</i>
1+2	<i>kit-æniki-</i>	<i>k-æene-</i>	<i>ke-epi-</i>	<i>k-entapo-</i>	<i>ki-tən-</i>
3	<i>n-æniki-</i>	<i>n-æene-</i>	<i>n-epi-</i>	<i>n-entapo-</i>	<i>ni-tən-</i>

paradigms of four verbs, all members of the S_A inflectional class. In this language, the verb ‘to be’ diverges from other S_A verbs like ‘to fall’ by having a first person marker *w-*, rather than *k-*. A similar pattern exists in Tiriyo (Table 2), where S_A verbs have a prefix with phonologically conditioned allomorphs *t-* / $_ə$ and *s-* / $_e$, while the verb ‘to go’ has a first-person prefix *wi-*. For both languages, the first person prefix on the first four verbs in the table is representative for the vast majority of S_A verbs.

Such divergent verbs have been identified for Hixkaryana (Derbyshire 1985: 188), Waiwai (Gildea 1998: 90), the three Taranoan languages (Meira 1998: 112–115), Bakairi (Meira 2003a), and Arara (Alves 2017: 153), but have only been subject to comparative scrutiny in Meira’s (1998) reconstruction of Proto-Taranoan. In synchronic analyses, these verbs and their first person prefixes may be called **IRREGULAR**, contrasting with regular prefixes (like Hixkaryana *ki-* and Tiriyo *t-/s-*) on regular verbs. Note that there is no widely accepted definition of irregularity (Stolz et al. 2012), and many stricter definitions (e.g., Haspelmath & Sims 2010) require the pattern to occur at a single place in the grammar. In such approaches, 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 details of synchronic analysis, the cause for the divergent inflectional patterns is found in the diachrony of the languages in question. The goal of this study is to approach these patterns from a comparative perspective and to provide a diachronic and functional account, proceeding as follows: In Section 2, relevant aspects of the Proto-Cariban verbal system are introduced, and it is shown that the mechanism of person marker extensions is responsible for patterns like in Tables 1 and 2. In Section 3, six incomplete person marker extensions and the verbs unaffected by them are described. Since languages show considerable etymological overlap in their conservative verbs, these are further discussed and reconstructed. Section 4 uses **bybee1985morphologyempty citation** network model of morphology to search explanations for the verbs (un)affected by each extension. Section 5 summarizes and discusses the results of the study.

2 The origins of irregular first person inflections

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

2.1 Proto-Cariban person marking and inflectional relics

Proto-Cariban is reconstructed by Gildea (1998) as using a person paradigm called Set I in its independent verb forms, shown in Table 3. Person indexation in transitive verbs was conditioned by a basic hierarchy $1/2 > 3$. 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 marked identically and non-transparently with the 1+2 prefix $*k-$.

which are simply represented with $\langle r \rangle$, rather than $\langle t \rangle$ for Wayana or $\langle t \rangle$ for Ye’kwana etc. In languages with strong morphophonological processes and/or non-phonemic orthography the original form is shown in an additional line when presented in interlinearized examples. Gildea (2018) is followed in using $\langle ə \rangle$ for the Proto-Cariban reconstructed by Meira & Franchetto (2005), although it was likely more back (Gildea et al. 2010).

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

(a) Transitive					(b) Intransitive		
A/P	1	2	1+2	3	S _A	S _P	
1		* <i>k</i> -		* <i>t(i)</i> -	1	* <i>w</i> -	* <i>u(j)</i> -
2	* <i>k</i> -			* <i>m(i)</i> -	2	* <i>m</i> -	* <i>ə(j)</i> -
1+2				* <i>kit(i)</i> -	1+2	* <i>kit</i> -	* <i>k</i> -
3	* <i>u(j)</i> -	* <i>ə(j)</i> -	* <i>k</i> -	* <i>n(i)</i> -	3	* <i>n</i> -	* <i>n(i)</i> -

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

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

2.2 Person marker extensions and lexical diffusion

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

- (1) Proto-Parukotoan 1S_A **w*- to 1>3
 1+2 **k*- to 1S_P^a
 1+2 **kit*- to 1+2S_P^a
 Proto-Waiwaian 1S_P **k*- to 1S_A
 **owi(ro)j*- ‘1PRO LK’ for 1P
 Waiwai 2S_A *m*- to 2S_P

^a Completed in Proto-Waiwaian, ongoing in Werikyana.

All innovations are person marker extensions except the innovative first person P form, 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 Ø/*j*- (Spike Gildea, p.c.). Hixkaryána has preserved split-s only in the second person prefixes, while Werikyana still shows the variation reconstructible to Proto-Parukotoan in its first person and 1+2 prefixes. Waiwai has lost the system entirely, which notably happened via three diachronically distinct innovations.

For Gildea (1998), person marker extensions are relevant for the loss of split-s and the resulting changes to indexing alignment, whereas this study focuses on a different aspect. Namely, they most likely took place via lexical diffusion, characterized as a type of extension by **harris1995historicaleempty citation**, a hypothesis supported by three facts. First, the variation in first person and 1+2 prefixes described above for Werikyana is not completely free. Some verbs only allow e.g. first person *k*-, but not *j*-, while others can occur with both,

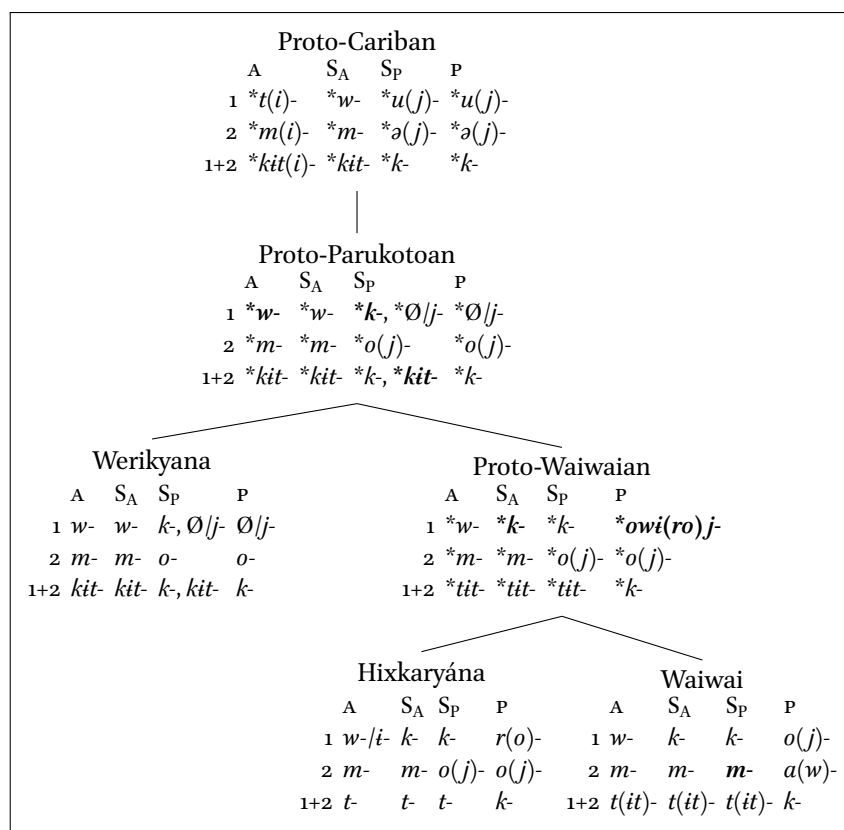


Figure 2: Person marker extensions in Parukotoan, after Gildea (1998: 94)

a pattern expected in a lexical diffusion scenario. In addition, this is speaker-dependent (Spike Gildea, p.c.), which points to an ongoing change. Second, while there is no detailed diachronic account of the switch of 1>3 *t- and 1S_A *w- in the Tiriyoan languages (Section 3.1.3), Meira (1998: 111–112) argues that it must have happened gradually rather than instantaneously, and entailed both markers spreading simultaneously. Whether or not this gradual switch followed ordered lines, lexical diffusion must have played a role.

The third argument in favor of the lexical diffusion scenario goes back to conservative forms like those in Hixkaryána and Tiriyo. In both languages, 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 the result of the extension stopping short of these verbs, rather than spreading through the entire S_A lexicon. In a family-wide search, 18 distinct extensions affecting intransitive verbs were identified, 6 of them incomplete. The latter have left between 1 and 7 conservatively inflected verbs in 9 Cariban languages (Section 3).

Interestingly, all six featured innovative first person markers on S_A verbs. All other (complete) extensions² 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(i)- (< *kit-) to S_P verbs in Apalaí (Table 4a), of 2S_A m(i)- in to S_P verbs in Panare (Table 4b), and the extension of the entire S_A set to S_P verbs in Waimiri-Atroari (Table 4c). The starkly different behavior of S_A and S_P verbs regarding extensions points to the split-s system playing a role, so its main properties will be discussed in Section 2.3. This will also clarify how the S_A/S_P distinction can be lost for a single person, or how S_P verbs can take on S_A markers with apparent semantic impunity.

2.3 The Cariban split-s system

As seen in Section 2.1, the split between S_A and S_P verbs was instantiated by inflection classes within the Proto-Cariban Set I person paradigm, but this was not the only difference: In deverbalized forms, S_A verbs took a class

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

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

(a) Apalaí			(b) Panare			(c) Waimiri-Atroari		
	S _A	S _P		S _A	S _P		S	
1	<i>i-/∅</i>	<i>i-/j-</i>	1	<i>w(i)-</i>	<i>∅/j-</i>	1	<i>w(i)-/i-</i>	
2	<i>m(i)-</i>	<i>o-</i>	2	<i>m(i)-</i>		2	<i>m(i)-</i>	
1+2	<i>s(i)-</i>		1+2	<i>n(i)-^a</i>		1+2	<i>h(i)-</i>	
3	<i>n(i)-</i>		3	<i>n(i)-</i>		3	<i>n-/∅</i>	

^a 1+2 was lost as a person value.

98 marker **w-*, while S_P verbs lacked that prefix (Gildea 1998: 89, 141–142; Meira 2000: 208). Also, S_P verbs took the
 99 2S_P prefix **a(j)-* in imperatives, while S_A verbs were unprefixes (Gildea 1998: 89; Meira 2000: 208).

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

100 (2) Kari’ña

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

102 b. *aya:woiya*
aj-awomi-ja
 2S_P-get.up-PRS
 ‘You are getting up.’ (Hoff 1968: 167)

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

110 Rather, the reason for a verb to take A- or P-oriented prefixes is (at least diachronically) a morphological one.
 111 Meira (2000: 217–221) demonstrates that those intransitive verbs which (etymologically) have a derivational
 112 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 detransitized forms of transitive verbs, either synchron-
 ically (with still existing transitive sources) or diachronically (with reconstructible but no longer
 existing transitive sources) (Meira 2000: 201)

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

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

114 Regarding the form of this marker, Meira et al. (2010: 505–512) reconstruct two distinct prefixes for Proto-
 115 Cariban: reciprocal **ate-* and reflexive **e-*, although their reflexes on verbs have been merged into a single mor-

116 pHEME in modern languages.³ Reflexes of **ate/e-* show a range of meanings summarizable as “detransitive”, illus-
 117 trated with Tiriyó S_A verbs in (3).

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

119	<i>nonta</i>	→	<i>e-nonta</i> ,	‘abandon each other’
	‘abandon’		<i>ai-nonta</i>	(reciprocal)
	<i>suka</i>	→	<i>e-suka</i> ,	‘wash self’
	‘wash’		<i>ai-suka</i>	(reflexive)
	<i>pahka</i>	→	<i>e-pahka</i>	‘break (INTR)’
	‘break (TR)’			(anticausative)
	<i>puunapi</i>	→	<i>ah-puunapi</i> ,	‘think, meditate’
	‘think about’		<i>ai-puunapi</i>	(antipassive)

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

125 3 Prefixes and verbs: innovation and resistance

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

131 3.1 Incomplete extensions: the innovative 1S_A markers

132 As stated in Section 2.2, the six person marker extensions which left a group of verbs untouched all introduced
 133 innovative first person markers on S_A verbs. Of these extensions, half can be reconstructed to intermediate proto-
 134 languages, and half happened in pre-modern stages of single languages. The sources of innovative markers vary,
 135 but not much: the innovative 1S_A prefix is formally identical to the 1+2P/S_P marker (Proto-Cariban **k-*) in three
 136 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
 137 case. For each extension, regular (innovative) verbs are contrasted with irregular (conservative) ones, and verb
 138 forms are reconstructed where necessary. Section 3.1.1 demonstrates the extension of **k-* in Proto-Pekodian,
 139 reflected in the three daughter languages Arara, Ikpeng, and Bakairi. Section 3.1.2 details the extension of **k-* in
 140 Proto-Waiwaian, briefly shown in Section 2.2. Section 3.1.3 focuses on innovative **t-* in Proto-Tiriyóan, reflected
 141 in modern Tiriyó and Akuriyó. The topic of Sections 3.1.4 to 3.1.6 are innovative 1S_A markers found in single
 142 languages: *k-* in Akuriyó, and *j-* in Carijona and Yukpa.

143 3.1.1 Proto-Pekodian **k-*

144 The Pekodian branch consists of closely related Arara and Ikpeng, with Bakairi as a more distant member. Meira
 145 & Franchetto (2005), who established the branch, focused on phonology and lexicon, so no reconstructions of

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

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

	Bakairi ‘to go up’	Arara ‘to dance’	Ikpeng ‘to run’
1	<i>k-əku-</i>	<i>k-origu-</i>	<i>k-aranme-</i>
2	<i>m-əku-</i>	<i>m-origu-</i>	<i>m-aranme-</i>
1+2	<i>kid-əku-</i>	<i>kud-origu-</i>	<i>kw-aranme-</i>
3	<i>n-əku-</i>	<i>Ø-origu</i>	<i>Ø-aranme-</i>

146 Proto-Pekodian morphosyntax are found in the literature. However, all three Pekodian languages have a regular
 147 1S_A marker *k-* (Table 5), allowing the reconstruction of a Proto-Pekodian 1S_A marker **k-*.

148 The most detailed description of a Pekodian language (Alves 2017) names six⁴ Arara S_A verbs forming a
 149 subclass defined by a first person marker *w(i)-* rather than *k-*, shown in (4). There is also a reflex of the copula
 150 **a[p]* (Section 3.2.3), serving syntactically as a postposition introducing adverbial clauses meaning ‘if’ or ‘when’
 151 (Alves 2017: 199–201). However, it inflects with verbal Set I prefixes, including first person *w-* (5).

152 (4) Arara (Alves 2017: 153)

153 *wi-geṇi* ‘I said’
w-iḡḡini ‘I was, lied down’
w-ebini ‘I came’
w-ibini ‘I bathed’
w-iptonri ‘I went down’
w-idoli ‘I went’

154 (5) Arara (Alves 2017: 200)

155 1 *w-aptam* ‘when/if I was’
2 *m-od-aptam* ‘when/if you were’
1+2 *kud-aptam* ‘when/if we were’
3 *Ø-aptam* ‘when/if s/he was’

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

158 (6) Bakairi (Meira 2003a: 4)

159 *w-i-də*
1S_A-bathe-IMM
‘I bathed’

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

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

165 a. ⟨u-ḡépa⟩
u-ge-pa
1S_A-say-NEG
‘I don’t say.’

166 b. ⟨wi-táki⟩ / ⟨wi-tági⟩
w-i-taki
1S_A-be-INT
‘I was.’

⁴Seven in her analysis, treating the two meanings of *iḡḡi* ‘to be / to lie down’ as different verbs.

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

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

	Proto-Pekodian	Arara	Ikpeng	Bakairi
‘say’	* <i>wi-ge-</i>	<i>wi-ge-</i>	<i>i-ge-</i>	<i>u-ge-</i>
‘go’	* <i>w-itən-</i>	<i>w-ido-</i>	<i>k-aran-</i>	<i>u-tə-</i>
‘be-1’	* <i>w-ap-</i>	<i>w-ap-</i>	–	<i>w-a-</i>
‘be-2’	* <i>w-efi-</i>	<i>w-ifi-</i>	Ø- <i>efi-</i>	<i>w-i-</i>
‘come’	* <i>w-epi-</i>	<i>w-ebi-</i>	<i>k-arep-</i>	<i>k-æwi-</i>
‘go down’	* <i>w-iptə-</i>	<i>w-ipton-</i>	?- <i>ipton-</i>	<i>k-itagi-</i>
‘bathe’	* <i>w-ipi-</i>	<i>w-ibi-</i>	Ø- <i>ip-</i>	<i>w-i-</i>

- 167 c. ⟨kχaewi-le⟩ 169 e. ⟨úta⟩ / ⟨uúta⟩
k-æwi-li
 1_{SA}-come-IMM
 ‘I came.’
u-tə
 1_{SA}-go
 ‘I go.’
- 168 d. ⟨kχ-itaké-he⟩ 170 f. ⟨töre-w-akine⟩
k-itagi-se
 1_{SA}-go.down-NPST?
 ‘I go down.’
tərə w-a-kine
 there 1_{SA}-be-PST.CONT
 ‘I was there.’

171 All descriptions of Ikpeng list *k-* as the only 1_{SA} marker (Pachêco 1997: 55; Campetela 1997: 105; Pachêco
 172 2001: 64; Chagas 2013: 205). However, most Ikpeng cognates of the verbs in question do not take *k-*, but rather *i-*
 173 or Ø (8), with the exception of *k*-prefixed ‘to go’ (9). There is a formally identical Ikpeng cognate of Arara *ipton*
 174 ‘to go down’, but no first person forms are attested (Angela Chagas, p.c.). While reflexes of **a[p]* ‘be-1’ do exist in
 175 Ikpeng, apparently only reflexes of **eti* ‘be-2’ occur with first person prefixes (Gildea 2018: 401).

176 (8) Ikpeng

- 177 a. *i-ge-li*
 1-say-REC
 ‘I said.’ (Pachêco 2001: 209)
- 178 b. Ø-*efi-li*
 1-be-REC
 ‘I was.’ (Pachêco 2001: 139)
- 179 c. *afagotpop* Ø-*ip-ffi* *ik-gwa-kffi*
 always 1-bathe-NPST river-LOC.aquatic-ALL
 ‘I always bathe in this river.’ (Pachêco 1997: 68)

180 (9) Ikpeng (Pachêco 2001: 80)

- 181 *k-aran-ffi*
 1-go-NPST
 ‘I’m going.’

182 Reconstructed Proto-Pekodian forms of conservatively inflected verbs are given in Table 6. Newly identified
 183 Ikpeng *i-/Ø* is demonstrably a reflex of Proto-Xinguan **w(i)-*, based on other (albeit irregular) cases of loss of
 184 **w* (Table 7). Similarly, the change of **wi* to Bakairi *u* is found in correspondences like *udo* (Meira & Franchetto
 185 2005) from Proto-Cariban **witoto* ‘person’ (Gildea & D. Payne 2007: 4). Thus, a 1_{SA} prefix **w(i)-* can securely

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

Meaning	Arara	Ikpeng
‘to defecate’	<i>watke</i>	<i>atke</i>
‘DAT’	<i>wina</i>	<i>ina</i>
‘dog’	<i>wokori</i>	<i>akari</i>
‘capuchin monkey’	<i>tawe</i>	<i>tae</i>
‘to sleep’	<i>winki</i>	<i>inki</i>

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

	‘to fall’			‘to sleep’		
	Proto-Waiwaian	Hixkaryána	Waiwai	Proto-Waiwaian	Hixkaryána	Waiwai
1	<i>k-epurka-</i>	<i>k-ehurka-</i>	<i>k-epirka-</i>	<i>ki-winiki-</i>	<i>ki-niki-</i>	<i>ki-winiki-</i>
2	<i>m-epurka-</i>	<i>m-ehurka-</i>	<i>m-epirka-</i>	<i>o-winiki-</i>	<i>o-wniki-</i>	<i>mi-winiki-</i>
1+2	<i>t-epurka-</i>	<i>t-ehurka-</i>	<i>tf-epirka-</i>	<i>tit-winiki-</i>	<i>ti-niki-</i>	<i>tit-winiki-</i>
3	<i>n-epurka-</i>	<i>n-ehurka-</i>	<i>n-epirka-</i>	<i>ni-winiki-</i>	<i>ni-niki-</i>	<i>ni-winiki-</i>

be reconstructed to Proto-Pekodian, identical to its Arara reflex in form and distribution. In later, individual developments, Bakairi extended *k-* to ‘to go down’, and Ikpeng to ‘to go’.

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

(10) Arara (Alves 2017: 150)

m-odebi-ni

2S_A-come-REC

‘You came.’

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

3.1.2 Proto-Waiwaian **k-*

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

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

(11) a. Waiwai (R. E. Hawkins 1998: 71)

wiikekne

wi-ka-jakne

1-say-PST

‘I said.’

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

	Proto-Waiwaian	Hixkaryána	Waiwai
'say'	*wi-ka-	i-ka-	wi-ka-
'go'	*wi-tom-	i-to-	kíw-tom-
'be-1'	*w-ah-	w-ah-	w-a-
'be-2'	*w-efi-	w-efe-	w-eefi-

- b. Hixkaryána (Derbyshire 1985: 124)

roxehra nay hami Kaywerye ikekoni

ro-fe-hira n-a-je hami kajwer'e i-ka-jakoni

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)

i-koroka-no

1>3-wash-IMM

'I washed him.'

- d. Waiwai (R. E. Hawkins 1998: 192)

wîyesî

wi-jo-jasi

1>3-boil-NPST

'I will boil it.'

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

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

3.1.3 Proto-Tiriyoan *t-

The moniker Tiriyoan (Hammarström et al. 2020) subsumes Tiriyo and Akuriyo, the more closely related of the three Taranoan languages identified by Girard (1971), with Carijona as a sister. Meira (1998) provides an extensive phonological, morphological, and lexical reconstruction of Proto-Taranoan, facing an interesting puzzle in the Set I paradigms of Tiriyo and Akuriyo: Proto-Cariban 1>3 *t- and 1S_A *w- seem to have switched places, creating a regular 1S_A marker of the form *ff- / _e, *t- / _ə (Table 10).⁶ The question of how this switch happened in detail (Meira 1998: 107–112) is still open, although a scenario seems necessary in which both *t- and *w- for a time occurred on both transitive and intransitive verbs (Meira 1998: 112).⁷

As for verbs unaffected by the spread of *t-, Meira (1998) reconstructs four of the items in Table 11 as taking *w- in Proto-Taranoan, for which reconstructed Proto-Tiriyoan forms are substituted here.⁸ As a fifth verb *e?i

⁶The latter allomorph was subsequently replaced by *k-* in Akuriyo (Section 3.1.4).

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

⁸The present reconstruction of 'to come' diverges from Meira's (1998: 114–115), who reconstructs Proto-Taranoan *æpi for first, but

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

	‘to bathe (INTR)’ Proto-Tiriyoan	Tiriyo	Akuriyo	‘to sleep’ Proto-Tiriyoan	Tiriyo	Akuriyo
1	* <i>ʃ-epi-</i>	<i>s-epi-</i>	<i>ʃ-epi-</i>	* <i>t-əəniki-</i>	<i>t-əəniki-</i>	<i>k-əəniki-</i>
2	* <i>m-epi-</i>	<i>m-epi-</i>	<i>m-epi-</i>	* <i>m-əəniki-</i>	<i>m-əəniki-</i>	<i>m-əəniki-</i>
1+2	* <i>ke-epi-</i>	<i>ke-epi-</i>	<i>ke-epi-</i>	* <i>kit-əəniki-</i>	<i>kit-əəniki-</i>	<i>kəʔ-əəniki-</i>
3	* <i>n-epi-</i>	<i>n-epi-</i>	<i>n-epi-</i>	* <i>n-əəniki-</i>	<i>n-əəniki-</i>	<i>n-əəniki-</i>

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

	Proto-Tiriyoan	Tiriyo	Akuriyo
‘say’	* <i>wi-ka-</i>	<i>wi-ka-</i>	<i>wi-ka-</i>
‘go’	* <i>w-itəmi-</i>	<i>wi-tən-</i>	<i>ə-təmi-</i>
‘be-1’	* <i>w-a-</i>	<i>w-a-</i>	Ø- <i>a-</i>
‘be-2’	* <i>w-eʔi-</i>	<i>w-ei-</i>	?- <i>eʔi-</i>
‘come’	* <i>w-əʔepi-</i>	<i>w-əepi-</i>	Ø- <i>eepi-</i>

232 ‘be-1’ can be added, whose Tiriyo reflex retains *w-*. The idiosyncratic Akuriyo first person prefix *ə-* on ‘to go’
 233 is identified as a reflex of **wi-* by Meira (1998: 113), which is supported the fact that both components of the
 234 idiosyncratic change **wi-* > *ə-* (**w* → Ø and **i* → *ə*) are found in other person prefixes (12a–b).

235 (12) Akuriyo

236	a. <i>wi-toka</i> 1>3-hit ‘I hit him/her.’ (Gildea 1994: 86)	237	b. <i>kəʔ-eepi</i> 1+2-come ‘We came.’ (Meira 1998: 114)
-----	--	-----	--

238 3.1.4 Akuriyo *k-*

239 After the split of Proto-Tiriyoan, when **t-* / **ʃ-* had largely replaced **w-*, Akuriyo innovated a third 1S_A marker
 240 *k-*. *k-* and *ʃ-* show a clear phonologically conditioned distribution in Gildea’s (1994) Akuriyo data (Table 12).
 241 Meira (1998: 107) largely confirms that distribution, but mentions “several cases of first person *t-* in Akuriyo” (on
 242 *ə*-initial verbs), albeit without any examples. He also suggests that *k-* may be more recent, which is plausible:
 243 since the distribution **t-* / *ə* / **ʃ-* / *ə* is reconstructible to Proto-Tiriyoan, the most straightforward scenario
 244 is *k-* replacing **t-* but not **ʃ-* in Akuriyo. The few *t-* mentioned by Meira (1998) were then perhaps reintroduced
 245 under Tiriyo influence. However, since there are no examples of – or more information about – *ə*-initial verbs
 246 with *t-*, these cases cannot be discussed further.

247 The verbs listed for Proto-Tiriyoan in Table 11 in Section 3.1.3 of course also resisted the extension of *k-* in
 248 Akuriyo, although the first-person form of the copular verb *eʔi* is unknown. In addition, there is an S_A verb *i(h)tə*
 249 ‘to go down’, which has an irregular first person marker *p-*, apparently reconstructible to Proto-Tiriyoan (13).
 250 It was not affected by the extension of Akuriyo *k-*, but whether it was an S_A verb when Proto-Tiriyoan **t-* was
 251 introduced is unclear (see Section 3.2.5).

252	(13) First person forms of ‘to go down’
253	Tiriyo <i>p-ihətə-</i> (Meira 1999: 294)
	Akuriyo <i>p-itə-</i> (Gildea 1994: 84)

**eepi* for other persons, based on the paradigmatic pattern in Tiriyo and the vowel length in Akuriyo. Akuriyo and Carijona would then have levelled that pattern, similar to what was suggested for the Pekodian languages (Section 3.1.1). Here, both Tiriyo *əepi* and Akuriyo *eepi* are identified as reflecting **ətepi* (Section 3.2.4), via Proto-Taranoan *(*əʃ-*)*epi* and Proto-Tiriyoan *(*əʔ-*)*epi*.

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

first person <i>k</i> -	first person <i>ff</i> -
<i>aempa</i> - ‘to learn’	<i>epi</i> - ‘to bathe (INTR)’
<i>aəffəna</i> - ‘to cry’	<i>ekirika</i> - ‘to stay back’
<i>aiwa</i> - ‘to tremble’	<i>entapo</i> - ‘to yawn’
<i>aməmi</i> - ‘to enter’	<i>etonema</i> - ‘to lie down’
<i>atajiŋka</i> - ‘to run’	<i>ewai</i> - ‘to sit down’
<i>aturu</i> - ‘to talk’	<i>ehpa</i> - ‘to bathe (INTR)’
<i>aəniki</i> - ‘to sleep’	

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

	‘to dance’	‘to arrive’
1	<i>j-eharaga</i> -	<i>ji-tuda</i> -
2	<i>m-eharaga</i> -	<i>mi-tuda</i> -
1+2	<i>kis-eharaga</i> -	<i>kisi-tuda</i> -
3	<i>n-eharaga</i> -	<i>ni-tuda</i> -

3.1.5 Carijona *j*-

Carijona, the third Taranoan language, extended the 1S_P marker *j(i)*-⁹ to S_A verbs (Meira 1998: 105–107). Combined with the extension of 2S_A *m*- and 1+2S_A *kit*-/*kis*- to S_P verbs, this created a single unified s category for regular verbs (Table 13). Although the split-s system was lost entirely, former S_A verbs can be identified by the presence of a detransitivizer, like *ehinəhi* ‘to fight’ (14), derived from *hinəhi* ‘to kill’ (Robayo Moreno 2000: 179).

- (14) Carijona (Koch-Grünberg 1908: 79)
həne(x)tónoko-máŋe y-e-hənə(x)yaɿ
hinəhono-ko=marə j-e-hinəhi-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 putative Proto-Taranoan 1S_A **t*-. However, it did not fully eclipse the old 1S_A marker **w*-, which is attested as being preserved in the verbs *tə* ‘to go’ (15a) and *a* ‘to be’ (15b). In addition, the verb *ka* ‘to say’ has a zero-marked first-person form (15c).

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

- a. *wi-tə-e=rehe*
 1-go-NPST=FRUST
 ‘I almost go (but I am not going to go).’
 b. *aji-marə-ne w-a-e*
 2-COM-PL 1-be-NPST
 ‘I am with you all.’
 c. *dēmēmara kae ewi iya*
n-təmə=mara Ø-ka-e əwi i-ja
 3-go=DUB 1-say-NPST.CERT 1PRO 3-OBL
 ‘“Did s/he leave?”, I say to him.’

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

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

	‘to wash self’	‘to sleep’	‘to fall’
1	<i>j-otum-</i>	<i>ji-ni-</i>	<i>j-ata-</i>
2	<i>m-otum-</i>	<i>mi-ni-</i>	<i>m-ata-</i>
3	<i>n-otum-</i>	<i>ni-ni-</i>	<i>n-ata-</i>

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

3.1.6 Yukpa *j-*

The divergent nature of Yukpa¹⁰ is *inter alia* visible in the loss of many Set I forms and the formation of non-cognate innovative main clause verbal constructions (Meira 2006b). However, it does preserve the Set I prefixes in the immediate past, which shows a unified intransitive paradigm (Table 14). The wholesale loss of 1+2 as an inflectional value was combined with the extension of 2S_A *m(i)-* to (now former) S_P verbs like *ni* ‘to sleep’. These verbs share their first person marker *j(i)-* with former S_A verbs like *otum* ‘to wash self’, identifiable by their semantics and the reflex of **əte/e-*. Since *j(i)-* is the reflex of the Proto-Cariban 1S_(P) marker **u(j)-* (Gildea 1998: 92), it also occurs on transitive verbs in 3>1 scenarios (16a). In contrast, 1>3 scenarios are zero-marked (16b).

(16) Yukpa (Meira 2006b: 139)

- a. *aw j-esare*
1PRO 3>1-see
‘S/he saw me.’
- b. *aw Ø-esare*
1PRO 1>3-see
‘I saw it.’

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

Zero-marked first person is attested for a single intransitive verb, *to* ‘to go’ (17). It diverges from regular C-initial verbs with *ji-*, like ‘to sleep’ (Table 14). It can thus be identified as having resisted the extension of *j-* in Yukpa.

(17) Yukpa (Meira 2006b: 139)

- aw Ø-to*
1PRO 1S_A-go
‘I went.’

3.2 Conservative verbs in comparison

In Section 3.1, six incomplete extensions of personal prefixes into 1S_A territory and the verbs unaffected by them were identified. This set is rather small in most cases, and many of the verbs remain conservative in differ-

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

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

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

ent (proto-)languages. Here, these verbs are investigated from a comparative perspective: **ka[ti]* ‘to say’ (Section 3.2.1), **itə[mə]* ‘to go’ (Section 3.2.2), both roots of the copula **eti* and **a[p]* (Section 3.2.3), **(ət)jəpɪ* ‘to come’ (Section 3.2.4), **ipitə* ‘to go down’ (Section 3.2.5), and **e-pi* ‘to bathe’ (Section 3.2.6). The large and phonologically coherent group of *e*-initial verbs not affected by the extension of Akuriyó *k-* (Section 3.1.4) will not be discussed.

3.2.1 **ka[ti]* ‘to say’

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

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

(18) Hixkaryána

¹¹Cognate segments in Tables 15 to 21 were aligned automatically with LingPy (lingpy268), for improved exposition of correspondences. Brackets indicate segments not present in all forms.

- 310 a. *oní wyaro nkekoni biryekomo, tiyoni wya*
oní wjaro n-ka-jakoni birjekomo ti-joni wja
 this like 3-say-REM.CONT boy COR-mother OBL
 ‘This is what the boy said to his mother.’ (Derbyshire 1985: 36)
- 311 b. *moro ha, ketxkoná hatá.*
moro ha ka-jaffkoni hati
 MED.DEM.INAN INTS say-REM.CONT.PL HSY
 ‘‘That one there’’ they said.’ (Derbyshire 1965: 14)

312 In (18a), the prefix *n-* occurs because there is no preceding object (‘he said it like this’), while does not occur in
 313 (18b) where the verb is preceded by direct speech. In the rest of the family, at least the Tiriyo reflex shows the
 314 same pattern, albeit inconsistently so (Carlin 2004: 267).

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

322 Arguments in favor of intransitive ‘to say’ primarily come from its inflectional prefixes. Kari’ña has a minimal
 323 pair between transitive *ka* ‘to remove’ and intransitive *ka* ‘to say’, *sikai* ‘I took it away’ vs *wikai* ‘I said’ (Courtz
 324 2008: 288, 45). Similarly, Proto-Pekodian **ke* ‘to say’ had 1S_A **w-* (Section 3.1.1), rather than 1>3 *s-* (Bakairi) or
 325 **ini-* (Proto-Xinguan). Additionally, languages differentiating transitive from S_A prefixes by the presence of *i*
 326 (Meira et al. 2010: 495) have *i*-final prefixes, see Akuriyo in (19a), as well as Meira (1999: 294), Tavares (2005: 195),
 327 Pachêco (2001: 288), Alves (2017: 150), and Hoff (1968: 168) for cognate forms in other such languages. Finally,
 328 the S_A class marker *w-* occurs on nominalizations in Kari’ña (19b), and is probably reflected as vowel length in
 329 the Tiriyo (Meira 1999: 333) and Wayana (Tavares 2005: 196) participles.

- 330 (19) a. Akuriyo (Meira 1998: 113)
 331 *mi-ka*
 2-say
 ‘You said.’
- 332 b. Kari’ña (Courtz 2008: 202)
 333 *Ômakon ‘wa oti ywykàpo kaiko.*
o-?ma-kon ?wa oti i-wi-ka-?po kai-ko
 2-child-PL OBL greeting 1-S_A-say-PST.NMLZ say-IMP
 ‘Pass my greetings to your children.’

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

337 3.2.2 **itə[mə]* ‘to go’

338 This verb was not affected by the any of the extensions in Section 3.1. Gildea & D. Payne (2007) reconstruct it
 339 as C-initial **tə[mə]*, like **ka[tí]* ‘to say’ with a fleeting second syllable. While many reflexes are unambiguously
 340 *t*-initial (e.g. Hixkaryana *ntoje* ‘he went’ [Derbyshire 1985: 27], Tiriyo *təkə* ‘go’ [Meira 1999: 246], or Wayana *kup-*
 341 *təm* ‘we went’ [Tavares 2005: 195]), the contrast with C-initial **ka[tí]* becomes clear once all forms are considered
 342 (Table 16). *i*-initial forms are predominant, so the vowel can tentatively be reconstructed, although languages
 343 throughout the family reflect **ə*. Since many languages only show *i* in certain contexts, even forms like Tiriyo
 344 *witanne* ‘I went’ (Meira 1999: 43) are ambiguous, since epenthetic *i* breaks up CC clusters on prefix-verb bound-
 345 aries.

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

Language	Form				
Werikyana	<i>to[mə]</i>	t	o	m	o
Proto-Waiwaian	<i>*[ɪ]to[m]</i>	i	t	o	m
Hixkaryána	<i>[ɪ]to</i>	i	t	o	
Waiwai	<i>[e]to[m]</i>	e	t	o	m
Proto-Pekodian	<i>*itə[n]</i>	i	t	ə	n
Arara	<i>to</i>	t	o		
Arara	<i>ido</i>	i	d	o	
Ikpeng	<i>aran</i>	a	r	a	n
Ikpeng	<i>ero</i>	e	r	o	
Bakairi	<i>[ɪ]tə</i>	i	t	ə	
Proto-Tiriyóan	<i>*[ə]tə[mɪ]</i>	ə	t	ə	m i
Tiriyó	<i>tə[n]</i>	t	ə	n	
Akuriyó	<i>[ə]tə[mɪ]</i>	ə	t	ə	m i
Carijona	<i>tə[mə]</i>	t	ə	m	ə
Wayana	<i>[ɪ]tə[m]</i>	i	t	ə	m
Apalaí	<i>ito</i>	i	t	o	
Kari’ña	<i>[ɪ]ʔ</i>	i	ʔ		
Kari’ña	<i>to</i>	t	o		
Ye’kwana	<i>itə[mə]</i>	i	t	ə	m ə
Ingarikó	<i>ətə</i>	ə	t	ə	
Pemón	<i>[e]tə</i>	e	t	ə	
Macushi	<i>[a]tɪ</i>	a	t	ɪ	
Panare	<i>tə[n]</i>	t	ə	n	
Yawarana	<i>tə</i>	t	ə		
Mapoyo	<i>tə</i>	t	ə		
Upper Xingu Carib	<i>[e]te</i>	e	t	e	
Yukpa	<i>to</i>	t	o		
Waimiri-Atroari	<i>ɪʔ</i>	i	ʔ		

3.2.3 *eti and *a[p] ‘to be’

*a[p] is the older copula and already had various irregularities in Proto-Cariban (Gildea 2018). *eti is reconstructed by Meira & Gildea (2009) and Gildea (2018) as originally meaning ‘to dwell, live’, but serving as a copula already in Proto-Cariban.¹² Reflexes of these roots are used suppletively, conditioned by person and/or TAM. Both roots preserved 1SA *w- in Proto-Pekodian, Proto-Waiwaian, and Proto-Tiriyoan (Sections 3.1.1 to 3.1.3). Akuriyó *a* was not targeted by the extension of *k-* (Section 3.1.4), while no first person form of *e?i* is attested. Carijona innovated *j-*, but only in the reflex of *eti (20); the *a* root preserves *w-* (Section 3.1.5). Yukpa introduced *j-* to the reflexes of both *a[p] and *eti, which are preserved as encliticized auxiliaries (21).

- (20) Carijona (Robayo Moreno 1989: 177)
- iretibə effinəme gərə jeffi*
ireti-bə effi-nə=me gərə j-effi-i
 then-from be-INF=ATTRZ still 1-be-PFV
 ‘Then I was already grown up.’

- (21) Yukpa (Meira 2006b: 143–144)

	NPST	PST
1	= <i>j-a(-s)</i>	= <i>j-e</i>
2	= <i>mak(o)</i>	= <i>m-e</i>
3	= <i>mak(o)</i>	= <i>n-e</i>

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

3.2.4 *(ət-)epi ‘to come’

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

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

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

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

The Pemongan languages and Werikyana point to *jəpi rather than *epi, although Werikyana *johi* is very rare in contrast to more frequent *o(o)hi*. It only occurs in the third person form of the Progressive, meaning that the *j* may be a reflex of third person **i-*. However, regular *o*-initial Werikyana verbs have no third person prefix (Table 19), while *i-* occurs with C-initial *to[mo]* ‘to go’, suggesting that *j* is indeed part of the root. Putative Proto-Cariban *jəpi could yield *epi via the two widespread sound changes *ə → *e / *j_ and *j → ∅ / #_ (Meira

¹²Such a stative, locative source is also suggested by the meaning ‘to lie down’ of Arara *iffi* (Alves 2017: 196).

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

Language	Form					
Werikyana	<i>ehi</i>				e	h i
Werikyana	<i>johi</i>			j	o	h i
Werikyana	<i>o[o]hi</i>	oo	-			h i
Proto-Pekodian	<i>*epi</i>				e	p i
Proto-Pekodian	<i>*ədepi</i>	ə	d	-	e	p i
Arara	<i>ebi</i>				e	b i
Arara	<i>odebi</i>	o	d	-	e	b i
Ikpeng	<i>arep</i>	a	r	-	e	p
Bakairi	<i>æwi</i>	ə		-	e	w i
Proto-Tiriyóan	<i>*epi</i>				e	p i
Proto-Tiriyóan	<i>*əʔepi</i>	ə	ʔ	-	e	p i
Tiriyó	<i>epi</i>				e	p i
Tiriyó	<i>æ[pɪ]</i>	ə		-	e	p i
Akuriyó	<i>eepe</i>	ee		-		p i
Carijona	<i>eh[i]</i>				e	h i
Apalaí	<i>oepe</i>	o		-	e	p i
Kari'ña	<i>o[o]pi</i>	oo		-		p i
Ye'kwana	<i>ehə</i>				e	h ə
Akawaio	<i>jepi</i>			j	e	p i
Akawaio	<i>əsipi</i>	ə	s	-	i	p i
Ingarikó	<i>jepə</i>			j	e	p ə
Ingarikó	<i>jə</i>			j	ə	
Patamona	<i>jepi</i>			j	e	p i
Patamona	<i>jəpi</i>			j	ə	p i
Pemón	<i>jepi</i>			j	e	p i
Macushi	<i>ipi</i>				i	p i
Panare	<i>ə[ə]pi</i>	əə		-		p i
Yawarana	<i>epi</i>				e	p i
Yawarana	<i>əəpi</i>	əə		-		p i
Mapoyo	<i>epi</i>				e	p i
Upper Xingu Carib	<i>e[e]</i>	ee				
Upper Xingu Carib	<i>i</i>					i

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

	Werikyana	Arara	Tiriyó
1	<i>w-oohi-</i>	<i>w-ebi-</i>	<i>w-əepi-</i>
2	<i>m-oohi-</i>	<i>m-odebi-</i>	<i>mən-epi</i>
1+2	<i>kis-ohi-</i>	<i>kud-ebi-</i>	<i>ke-epi</i>
3	<i>n-ehi-</i>	<i>t-ebi-</i>	<i>n-epi</i>

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

	‘to come’	‘to dream’	‘to go’
1	Ø- <i>w-ooht</i> -	Ø- <i>w-osone</i> -	Ø- <i>wi-to</i> -
2	<i>o-w-ohi</i> -	<i>o-w-osone</i> -	<i>o-w-to</i> -
1+2	<i>ku-w-ohi</i> -	<i>ku-w-osone</i> -	<i>ki-w-to</i> -
3	Ø- <i>johi</i> -	Ø- <i>osone</i> -	<i>i-to</i> -

et al. 2010). The argument against this reconstruction is that $*\partial \rightarrow *e / *j_$ did not happen in Proto-Parukotoan, and only inconsistently in Proto-Panare-Pemongan and Mapoyo-Yawarana (Meira et al. 2010: 501–502; Gildea et al. 2010). Thus, it does not explain Werikyana *ehi*, nor the total absence of the sequence $*j\partial$ in Panare and Mapoyo-Yawarana.

The clear segmentability of $*\partial t$ - in combination with its form suggest that it is a detransitivizing prefix. Although the combination of a detransitivizer and an intransitive verb makes no sense semantically, some historical S_P verbs are attested as adding the detransitivizer to become S_A verbs. For example, the Proto-Cariban S_P verb $*winiki$ ‘to sleep’ becomes Tiriyo *əniki* (Meira 1999: 252) and Kari’ña *əʔniki* (Courtz 2008: 429), both S_A. Waiwai ‘to sleep’ can be *winik* (R. E. Hawkins 1998: 30) or *et-winik* (W. N. Hawkins & R. E. Hawkins 1953: 204). However, unlike ‘to sleep’, bare $*epi$ ‘to come’ apparently was an S_A verb already (although its reflexes in languages preserving the split-s system mostly co-occur with reflexes of $*\partial t-epi$).

3.2.5 $*ipit\partial$ ‘to go down’

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

- (23) a. Carijona (David Felipe Guerrero, p.c.)
irə waffinakano tae j-ehitə-e
 INAN.ANA body.of.water along.bounded 1-go.down-NPST
 ‘...I go down through that guachinacán.’
- b. Yukpa (Meira 2003b)
aw yéwtu
aw j-ewuhtu
 1PRO 1-go.down
 ‘I went down.’

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

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

- (24) Wayana (Tavares 2005: 200)

¹³The cognacy status of parenthesized forms in Table 20 is uncertain. The reconstruction of Proto-Pekodian $*ipit\partial$ treats the additional elements in daughter languages as non-cognate. Meira & Franchetto (2005) identify no correspondence between Bakairi *gi* and Ikpeng *ŋ*, and at least the addition of a final *ŋ* in Proto-Xinguan is attested elsewhere: a) Proto-Cariban $*\partial ne$ ‘to see’, Arara and Ikpeng *eneŋ*; b) Proto-Cariban $*\partial ta$ ‘to hear’, Arara *taŋ*, Ikpeng *iraŋ*; and c) Proto-Cariban $*\partial na$ ‘to eat meat’, Arara *oŋoŋ* ‘to bite’ (Gildea & D. Payne 2007: 8; Alves 2017: 56, 144, 57; Pacheco 2001: 25, 270).

Table 20: Reflexes of **ipitə* ‘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	<i>*iφito</i>	S _P		i	φ	i	t	o	
Werikyana	<i>ihito</i>	S _P		i	h	i	t	o	
Proto-Waiwaian	<i>*hto</i>	?			h		t	o	
Hixkaryána	<i>hto</i>	?			h		t	o	
Waiwai	<i>hto</i>	–			h		t	o	
Proto-Pekodian	<i>*iptə</i>	S _A		i	p		t	ə	
Arara	<i>ipton</i>	S _A		i	p		t	o	- ŋ
Ikpeng	<i>ipton</i>	?		i	p		t	o	- ŋ
Bakairi	<i>itəgi</i>	S _A		i			t	ə	- g i
Proto-Taranoan	<i>*ipitə</i>	?		i	p	i	t	ə	
Proto-Tiriyoan	<i>*ihtə</i>	S _A		i	h		t	ə	
Tiriyó	<i>ih̥tə</i>	S _A		i	h		t	ə	
Akuriyó	<i>i[h̥]tə</i>	S _A		i	h		t	ə	
Carijona	<i>ehitə</i>	–		e	h	i	t	ə	
Wayana	<i>iptə</i>	S _A / S _P		i	p		t	ə	
Apalaí	<i>ih̥to</i>	S _P		i	h		t	o	
Kari’ña	<i>oniʔto</i>	(S _A)	o - n -	i	ʔ		t	o	
Ye’kwana	<i>əʔtə</i>	S _P		ə	ʔ		t	ə	
Kapón	<i>(uʔtə)</i>	–							
Pemón	<i>(uʔtə)</i>	–							
Macushi	<i>(auti)</i>	–							
Panare	<i>əhtə</i>	S _A		ə	h		t	ə	
Yawarana	<i>əhtə</i>	–		ə	h		t	ə	
Yukpa	<i>(ew[uh]tu)</i>	–							
Waimiri-Atroari	<i>iti</i>	–				i	t	i	

410	a.	<i>īwiptēē</i>	411	b.	<i>əw-iptā-k</i>
		<i>i-w-iptā-ri</i>			2-go.down-IMP
		1-S _A -go.down-NMLZ			'Go down!'
		'my going down'			

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

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

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

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

433 3.2.6 **e-pi* 'to bathe'

434 This verb resisted the extensions of Proto-Pekodian **k-* (Section 3.1.1) and Akuriyó *k-* (Section 3.1.4). It took on
 435 new 1S_A prefixes in Proto-Tiriyoan (Tiriyo *s-epi-*, Akuriyó *f̥-epi-* [Meira 1999: 292; Gildea 1994: 87]) and Proto-
 436 Waiwaian (Hixkaryána *k-ewehi-*, Waiwai *k-ejeφu-* [Derbyshire 1985: 195; R. E. Hawkins 1998: 166]). The first per-
 437 son form of its Carijona reflex *ehi* (Koch-Grünberg 1908: 72) is unknown; an unattested Yukpa cognate may exist.

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

444 3.3 Summary

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

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

(a) Reflexes of * <i>e-pi</i> ‘to bathe (INTR)’						
Language	Form					
Werikyana	<i>eehi</i>	ee		-	h	i
Hixkaryána	<i>ewehi</i>	e	w	e	-	h i
Waiwai	<i>ejeɸu</i>	e	j	e	-	ɸ u
Arara	<i>ibi</i>	i			-	b i
Ikpeng	<i>ip</i>	i			-	p
Bakairi	<i>i</i>	i				
Tiriyó	<i>epi</i>	e			-	p i
Akuriyó	<i>epi</i>	e			-	p i
Wayana	<i>epi</i>	e			-	p i
Apalaí	<i>epi</i>	e			-	p i

(b) Reflexes of * <i>e-kupi</i> ‘to bathe (INTR)’						
Language	Form					
Kari’ña	<i>ekupi</i>	e	-	k	u	p i
Ye’kwana	<i>eʔhi</i>	e	-	ʔ		h i
Kapón	<i>ekuʔpi</i>	e	-	k	u	ʔ p i
Pemón	<i>ekupi</i>	e	-	k	u	p i

(c) Reflexes of * <i>a-kupi</i> ‘to bathe (INTR)’						
Language	Form					
Panare	<i>akupi</i>	a	-	k	u	p i

(d) Reflexes of * <i>[i]pi</i> ‘to bathe (TR)’						
Language	Form					
Werikyana	<i>ihi</i>	i		h	i	
Hixkaryána	<i>ihi</i>	i		h	i	
Waiwai	<i>pi</i>			p	i	
Arara	<i>ibi</i>	i		b	i	
Ikpeng	<i>ip</i>	i		p		
Bakairi	<i>i</i>				i	
Tiriyó	<i>pi</i>			p	i	
Akuriyó	<i>pi</i>			p	i	
Wayana	<i>upi</i>	u		p	i	
Apalaí	<i>pi</i>			p	i	
Ye’kwana	<i>ihi</i>	i		h	i	
Pemón	<i>pi</i>			p	i	
Panare	<i>ipi</i>	i		p	i	

(e) Reflexes of * <i>kupi</i> ‘to bathe (TR)’						
Language	Form					
Kari’ña	<i>kupi</i>	k	u		p	i
Kapón	<i>kuʔpi</i>	k	u	ʔ	p	i
Panare	<i>kupi</i>	k	u		p	i

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

	* <i>ka</i> [ti]	* <i>itə</i> [mə]	* <i>a</i> [p]	* <i>eti</i>	* <i>(ət-)epi</i>	* <i>ipitə</i>	* <i>e-pi</i>	✓ affected; × not
	‘say’	‘go’	‘be-1’	‘be-2’	‘come’	‘go down’	‘bathe’	
Proto-Waiwaian * <i>k-</i>	×	×	×	×	–	N/A	✓	
Hixkaryána	×	×	×	×	–	N/A	✓	
Waiwai	×	(✓)	×	×	–	N/A	✓	
Proto-Pekodian * <i>k-</i>	×	×	×	×	×	×	×	
Arara	×	×	×	×	×	×	×	
Ikpeng	×	✓	–	×	✓	?	×	
Bakairi	×	×	×	×	✓	✓	×	
Proto-Tiriyóan * <i>t-</i>	×	×	×	×	×	N/A	✓	
Tiriyó	×	×	×	×	×	N/A	✓	
Akuriyó	×	×	×	?	×	N/A	✓	
Akuriyó <i>k-</i>	×	×	×	?	×	×	×	
Carijona <i>j-</i>	×	×	×	✓	✓	N/A	?	
Yukpa <i>j-</i>	?	×	✓	✓	–	N/A	–	

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

4 Explaining conservativeness: a network morphology approach

A well-known approach to conservativeness, innovativeness, and (ir-)regularity in the lexicon is bybee1985morphologyempty cit 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” (bybee1995regular). It does so by modeling shared morphological properties such as inflectional patterns as emerging from connections of differing strength between lexemes. A well-known example is a network of “strong” English verbs with *strɪŋ-stray* at the center and pairs like *rɪŋ-ray*, *spɪn-span*, or *stɪk-stak* at its periphery. This network created new strong verbs in some dialects, like *snɪk-snak* or *brɪŋ-bray* (bybee1985morphology). These verbs formed a lexical connection with prototypical members of the group, developing new past tense forms as a result.

These verbs are an example for a phonologically motivated network; other possible bases of connections between lexemes are semantic or morphological similarity (bybee1985morphology). Another important factor in the model is frequency, since more frequent words have higher lexical strength (bybee1985morphology). This diminishes the influence from other lexemes, meaning that high-frequency items are more likely to resist innovations. For the Cariban first person patterns, the model predicts that semantically/phonologically/morphologically similar verbs are affected by person marker extensions, while high-frequency verbs resist extensions and thus remain conservative.

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

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

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

These predictions were checked against the data in Table 22 by tallying verbs with (in/)correctly predicted behavior. The resulting scores are illustrated for the extension of Proto-Tiriyóan **t-* in Table 32 and summed up for all extensions in Table 26. These scores crucially only consider the seven verbs in Table 22, but each extension affected many run-of-the-mill S_A verbs.¹⁴ If one simulates the addition of 1’000 regularly inflected

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

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

Verb	Count	% S _A verb tokens
<i>a</i> ‘be-1’	49	30.06%
<i>efi</i> ‘be-2’	30	18.40%
<i>ka</i> ‘say’	26	15.95%
<i>ito</i> ‘go’	23	14.11%
<i>oeṗi</i> ‘come’	13	7.98%
<i>e-poreʔka</i> ‘arrive’	3	1.84%
<i>ot-urupo</i> ‘ask’	2	1.23%
<i>ot-uʔ</i> ‘eat’	2	1.23%
<i>os-enakũnuʔ</i> ‘choke’	2	1.23%
<i>e-unopi</i> ‘laugh’	1	0.61%
<i>at-akĩma</i> ‘pack bags’	1	0.61%
<i>at-ankiema</i> ‘be happy’	1	0.61%
<i>os-ereʔ</i> ‘be amazed’	1	0.61%
<i>e-metika</i> ‘lose loincloth’	1	0.61%
<i>e-tuarima</i> ‘suffer’	1	0.61%
<i>e-puka</i> ‘fall’	1	0.61%
<i>os-epori</i> ‘meet’	1	0.61%
<i>ot-iriʔka</i> ‘land’	1	0.61%
<i>ot-iʔka</i> ‘finish’	1	0.61%
<i>ot-uru</i> ‘talk’	1	0.61%
<i>at-apiaka</i> ‘divide up’	1	0.61%
<i>e-siriʔma</i> ‘move’	1	0.61%

Table 24: Predictions for Proto-Tiriyoan

	<i>*ka</i> ‘say’	<i>*[ə]tə[mi]</i> ‘go’	<i>*a</i> ‘be-1’	<i>*eʔi</i> ‘be-2’	<i>*əʔepi</i> ‘come’	<i>*epi</i> ‘bathe’
DETRZ	×	×	×	×	✓	✓
DETRZ+freq	×	×	×	×	×	✓
phono (/ _ *ə, e)	×	×	×	✓	✓	✓
phono+freq	×	×	×	×	×	✓
infl (*w-)	✓	✓	✓	✓	✓	✓
infl+freq	×	×	×	×	×	✓

Table 25: Evaluating predictions for Proto-Tiriyoan

	* <i>ka</i> 'say'	*[ə] <i>tə</i> [<i>mi</i>] 'go'	* <i>a</i> 'be-1'	* <i>eʔi</i> 'be-2'	* <i>əʔepi</i> 'come'	* <i>epi</i> 'bathe'	Score
DETRZ+freq	✓	✓	✓	✓	✓	✓	100.0%
phono+freq	✓	✓	✓	✓	✓	✓	100.0%
infl+freq	✓	✓	✓	✓	✓	✓	100.0%
DETRZ	✓	✓	✓	✓	×	✓	83.3%
phono	✓	✓	✓	×	×	✓	66.7%
infl	×	×	×	×	×	✓	16.7%

Table 26: Overview of prediction accuracy

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

detransitivized S_A verbs per language – a conservative estimate based on Courtz's (2008) Kari'ña dictionary – all six hypotheses consistently predict the behavior of 99.99% verbs correctly. Since the descriptive coverage of Cariban languages makes using real such large-scale data impossible, the present investigation is restricted to the edge cases.

The extent of the Proto-Waiwaian and Proto-Pekodian and extensions is fully predicted by the presence or absence of a detransitivizer. In both cases, only the underived¹⁵ S_A verbs were not affected, all other S_A verbs taking **k-*. Not shown in Table 26 are subsequent evolutions in the Pekodian daughter languages, which largely support a detransitivizer-based explanation: First, both Ikpeng and Bakairi regularized the paradigm to use forms with detransitivizer for first person (Section 3.1.1), and both introduced *k-*.¹⁶ Second, the development of Proto-Pekodian **itən* 'to go' to Ikpeng *aran* may have led to reanalysis of *ar* as a detransitivizer, accompanied by the introduction of *k-*.

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

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

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

S_A status at the time of a person marker extension, they are not relevant for this study.

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

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

Proto-Tiriyoan, phonology and inflection in Proto-Waiwaian, and phonology in Proto-Pekodian.

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

5 Conclusion

The main research questions of this study were a) where did irregular first person prefixes originate? b) which verbs are irregular in what language? and c) why are they irregular? The main findings can be summarized as follows: Irregular first person inflections are conservative, leftovers of person marker extensions which did not affect some verbs. Some of these extensions are reconstructible to proto-languages, while others happened in pre-modern stages of single languages. The same 1-7 verbs are conservative in all languages, and are often irregular in other ways. **bybee1985morphologyempty citation** network model delivers explanations for distribution of innovative markers, but in 4 of 6 cases, multiple explanations predict the same outcome. This is due to the situation reconstructible to Proto-Cariban, where only a small group of frequent S_A verbs had no detransitivizer **ate/e-*.

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

bybee1985morphologyempty citation network model held its promise of explaining irregularities in inflectional patterns. It did overshoot its goal somewhat in that only two of six cases had an unambiguous answer – phonology. Notably, the two phonology-based hypotheses (with and without frequency) together correctly predicted 100% of patterns. Since the crude frequency model increased overall prediction accuracy, but decreased it in the case of Carijona and Yukpa, language-specific counts of S_A verbs would be interesting. Regardless of the interpretation of the results, the inconclusiveness of the network model's answers is due to a specific pattern in Proto-Cariban. This raises the general question how applicable the model is when ambiguity exists.

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

6 Abbreviations

The following abbreviations were used in this paper:

1 first person

1+2 first and second person

1+3 first and third person

2 second person

3 third person

- 570 A agent-like argument of canonical transitive verb
- 571 AGT agent
- 572 ALL allative
- 573 ANA anaphoric
- 574 ATTRZ attributivizer
- 575 CAUS causative
- 576 CERT certainty
- 577 COM comitative
- 578 CONT continuative
- 579 COR coreference
- 580 DEM demonstrative
- 581 DES desiderative
- 582 DETRZ detransitivizer
- 583 DUB dubitative
- 584 EVID evidentiality
- 585 FRUST frustrative
- 586 HSY hearsay/indirect evidentiality
- 587 IMM immediate past
- 588 IMP imperative
- 589 INAN inanimate
- 590 INF infinitive
- 591 INT intermediate past
- 592 INTR intransitive
- 593 INTS intensifier
- 594 LK linker
- 595 LOC locative
- 596 MED medial
- 597 NEG negation
- 598 NMLZ nominalizer/nominalization
- 599 NPST non-past
- 600 OBL oblique
- 601 P patient-like argument of canonical transitive verb

- 602 PFV perfective
 603 PL plural
 604 PRO pronoun
 605 PRS present
 606 PST past
 607 REC recent past
 608 REM remote past
 609 s single argument of canonical intransitive verb
 610 S_A S marked like A
 611 S_P S marked like P
 612 SAP speech act participant
 613 SUP supine
 614 TAM Tense-Aspect-Mood
 615 TR transitive
 616 UNCERT uncertainty

617 7 Acknowledgments

618 REDACTED

619 References

- 620 Abbott, Miriam (1991). "Macushi". In: *Handbook of Amazonian Languages*. Desmond C. Derbyshire & Geoffrey K.
 621 Pullum (eds.). Vol. 3. Berlin/New York: Mouton de Gruyter: 23–160.
 622 Álvarez, José (2000). "Construcciones progresivas en pemón y otras lenguas caribes". In: *Opción* 32: 96–130.
 623 — (2008). "Cláusulas relativas, nominalización y constituyentes en pemón (Caribe)". In: *Opción* 24.57: 114–143.
 624 Alves, Ana Carolina Ferreira (2017). "Morfofonologia, morfossintaxe e o sistema de tempo, aspecto e modo em
 625 Arara (Karib)". Ph.D. dissertation. Universidade de São Paulo.
 626 Bruno, Ana Carla dos Santos (1996). *Dicionário Waimiri-Atroari-Português*. Manaus: Programa Waimiri-Atroari /
 627 Conevênio FUNAI-Eletronorte.
 628 Cáceres, Natalia (2011). "Grammaire fonctionnelle-typologique du ye'kwana, langue caribe du Venezuela". Ph.D.
 629 dissertation. Université Lumière Lyon 2.
 630 Cáceres, Natalia & Spike Gildea (2018). "A first analysis of Tense-Aspect constructions in Yawarana (Cariban)".
 631 Paper presented at Amazónicas VI. Baños, Colombia.
 632 Cáceres Arandia, Natalia, Spike Gildea & Marie-Claude Mattéi-Müller (2017). "Loss of Inflection in Yawarana
 633 (Cariban)". Paper presented at the International Conference on Historical Linguistics 23. San Antonio.
 634 Caesar-Fox, Desrey Clementine (2003). "Zauro'nödok Agawayo Yau: Variants of Akawaio spoken at Waramadong".
 635 Ph.D. dissertation. Rice University.
 636 Camargo, Eliane (2002). *Léxico bilingüe: aparaí-português / português-aparaí*. München: Lincom Europa.
 637 Camargo, Eliane & Anaiman Tapinkili (2010). *Hakëne omijau eitop Wajana-Palasisi / Dictionnaire bilingue Wayana-*
 638 *Français*. Paris: CELIA/DRAC-Guyane, TEKUREMAI.
 639 Campetela, Cilene (1997). "Análise do sistema de marcação de caso nas orações independentes da língua Ikpeng".
 640 MA thesis. Universidade Estadual de Campinas.

- 641 Carlin, Eithne B. (2004). *A Grammar of Trio: A Cariban Language of Suriname*. Vol. 55. Duisburger Arbeiten zur
642 Sprach- und Kulturwissenschaft. Frankfurt am Main: Peter Lang.
- 643 Chagas, Angela Fabíola Alves (2013). “O Verbo Ikpeng: estudo morfossintático e semântico-lexical”. Ph.D. disser-
644 tation. Universidade Estadual de Campinas.
- 645 Courtz, Henk (2008). *A Carib grammar and dictionary*. Toronto: Magoria Books.
- 646 Cruz, Maria Odileiz Sousa (2005). “Fonologia e gramática Ingarikó: Ka?pon-Brasil”. Ph.D. dissertation. Vrije Uni-
647 versiteit Amsterdam.
- 648 de Armellada, Cesáreo (1944). *Gramática y diccionario de la lengua Pemón*. Vol. 2: *Diccionario*. Caracas: Artes
649 Gráficas.
- 650 Derbyshire, Desmond C. (1965). *Textos Hixkaryana*. Vol. 3. Publicações Avulsas. Belém: MPEG.
- 651 — (1979). *Hixkaryana*. Vol. 1. Lingua Descriptive Studies. Amsterdam: North-Holland.
- 652 — (1985). *Hixkaryana and Linguistic Typology*. Dallas: Summer Institute of Linguistics.
- 653 — (1999). “Carib”. In: *The Amazonian Languages*. R.M.W. Dixon & Alexandra Y. Aikhenvald (eds.). Cambridge:
654 Cambridge University Press: 23–64.
- 655 de Souza, Shirley Dias Cardoso (1993). “Alguns aspectos morfológicos da língua Arara (Karib)”. MA thesis. Uni-
656 versidade de Brasília.
- 657 — (2010). “The morphology of nouns in the Ugoronmo language (Arara of Pará)”. MA thesis. University of North
658 Dakota.
- 659 Franchetto, Bruna (1986). “Falar Kuikúru: Estudo Etnolinguístico de um grupo Karíbe do Alto Xingu. Volume III:
660 Fonologia e Textos”. Ph.D. dissertation. Museu Nacional da Universidade Federal do Rio de Janeiro.
- 661 — (2002). “Kuikuro: uma língua ergativa no ramo meridional da família Karib (Alto Xingu)”. In: *Ergatividade na*
662 *Amazônia*. Vol. 1. Francesc Queixalós (ed.). Brasília, Paris: Laboratório de Línguas Indígenas, Centre d’études
663 des langues indigènes d’Amérique.
- 664 — (2008). “Absolutivo e ergativo pensados a partir do Kuikuro”. In: *Amerindia* 32: 37–56.
- 665 Franchetto, Bruna & Sérgio Meira (2016). “A phonological hypothesis on the historical origin of Kuikuro and
666 Bakairi verb classes”. Talk held at Amazônicas VI. Universidad Nacional de Colombia en Amazonia, Leticia.
- 667 García Ferrer, Donaldo José (2006). “Diccionario piloto pemón-español”. BA thesis. Universidad del Zulia.
- 668 Gildea, Spike (1994). “Akuriyó field notes”.
- 669 — (1998). *On reconstructing grammar: Comparative Cariban morphosyntax*. Oxford: Oxford University Press.
- 670 — (2012). “Linguistic studies in the Cariban family”. In: *The Indigenous Languages of South America: A Compre-*
671 *hensive Guide*. Lyle Campbell & Verónica Grondona (eds.). Vol. 2. The World of Linguistics. Berlin/New York:
672 Mouton de Gruyter: 441–494.
- 673 — (2015). “The evolution of valence categories in the Cariban family”. Paper presented at the Venezuelan Cariban
674 Conference. Lyon. URL: [https://www.academia.edu/30561539/The_evolution_of_valence_](https://www.academia.edu/30561539/The_evolution_of_valence_categories_in_the_Cariban_family)
675 [categories_in_the_Cariban_family](https://www.academia.edu/30561539/The_evolution_of_valence_categories_in_the_Cariban_family) (visited on 10/02/2020).
- 676 — (2018). “Reconstructing the copula and nonverbal predicate constructions in Cariban”. In: *Nonverbal pred-*
677 *ication in Amazonian languages*. Simon E. Overall, Rosa Vallejos & Spike Gildea (eds.). Amsterdam: John
678 Benjamins: 365–402.
- 679 Gildea, Spike & Natalia Cáceres (in preparation). “The Cariban Language Family”. In: *Amazonian Languages: An*
680 *International Handbook*. Patience Epps & Lev Michael (eds.). Berlin: Mouton de Gruyter.
- 681 Gildea, Spike, Berend J. Hoff & Sérgio Meira (2010). “The story of *ô in the Cariban Family”. In: *Fieldwork and*
682 *linguistic analysis in Indigenous languages of the Americas*. Andrea L. Berez, Jean Mulder & Daisy Rosenblum
683 (eds.). University of Hawai’i Press: 91–123.
- 684 Gildea, Spike & Doris Payne (2007). “Is Greenberg’s “Macro-Carib” viable?” In: *Boletim do Museu Paraense Emílio*
685 *Goeldi. Ciências Humanas* 2.2: 19–72.
- 686 Gildea, Spike & Fernando Zúñiga (2016). “Referential hierarchies: A new look at some historical and typological
687 patterns”. In: *Linguistics* 54.3: 483–529.
- 688 Girard, Victor James (1971). “Proto-Carib phonology”. Ph.D. dissertation. University of California.
- 689 Guerrero Beltrán, David Felipe (2019). “The Grammar of Space in Karijona, a Cariban language from Northwest
690 Amazonia”. MA thesis. Universidad Nacional de Colombia.
- 691 Guerrero-Beltrán, David Felipe (2016). “The alignment and person-marking system in Karijona (Cariban, North-
692 west 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).
- Haspelmath, Martin & Andrea Sims (2010). *Understanding morphology*. 2nd ed. London: Hodder Education.
- Hawkins, Robert E. (1998). "Wai Wai". In: *Handbook of Amazonian Languages*. Desmond C. Derbyshire & Geoffrey K. Pullum (eds.). Vol. 4. Berlin/New York: Mouton de Gruyter: 25–224.
- Hawkins, W. Neill & Robert E. Hawkins (1953). "Verb inflection in Waiwai (Carib)". In: *International Journal of American Linguistics* 19.3: 201–211.
- Hoff, Berend J. (1968). *The Carib Language: Phonology, Morphophonology, Morphology, Texts and Word Index*. Vol. 55. Verhandelingen van het Koninklijk instituut voor taal-, land- en volkenkunde. The Hague: Martinus Nijhoff.
- Howard, Catherine V. (2001). "Wrought Identities: The Waiwai expeditions in search of the "unseen tribes" of Northern Amazonia". Ph.D. dissertation. University of Chicago.
- Koch-Grünberg, Theodor (1908). "Die Hianákoto-Umáua". In: *Anthropos* 3: 83–124, 297–335, 952–982.
- Koehn, Edward & Sally Koehn (1986). "Apalai". In: *Handbook of Amazonian Languages*. Desmond C. Derbyshire & Geoffrey K. Pullum (eds.). Vol. 1. Berlin/New York: Mouton de Gruyter: 33–127.
- (1994). *14 textos Apalaí*. Brasília: Summer Institute of Linguistics.
- Lacerda, Edith (1991). "Waimiri-Atroari: observações lingüísticas". ms. Manaus.
- Largo, Wilson S. (2011). "Una gramática Yukpa". ms. Bogotá: Fundación para el Desarrollo de Pueblos Marginados.
- 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-Eñepa*. 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.etnolingustica.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 (Karib)". In: *Ergatividade na Amazônia*. Vol. 2. Francesc Queixalós (ed.). Brasília, Paris: Laboratório de Línguas Indígenas, Centre d'études des langues indigènes d'Amérique: 155–163.
- (2003b). "Primeras observaciones sobre la lengua yukpa". In: *Lingua Americana* 7.12.
- (2005). "Reconstructing Pre-Bakairi Segmental Phonology". In: *Anthropological Linguistics* 47.3: 261–291.
- (2006a). "Cariban Languages". In: *Encyclopedia of language and linguistics*. Keith Brown (ed.). Oxford: Elsevier: 199–203.
- (2006b). "Syntactic reanalysis in Yukpa (Cariban)". In: *International Journal of American Linguistics* 72.1: 135–150.
- Meira, Sérgio & Bruna Franchetto (2005). "The Southern Cariban languages and the Cariban family". In: *International Journal of American Linguistics* 71.2: 127–192.
- Meira, Sérgio & Spike Gildea (2009). "Property concepts in the Cariban family: adjectives, adverbs, and/or nouns". In: *The linguistics of endangered languages: Contributions to morphology and morphosyntax*. W. Leo Wetzels (ed.). Utrecht: LOT: 95–133.
- Meira, Sérgio, Spike Gildea & Berend J. Hoff (2010). "On the Origin of Ablaut in the Cariban Family". In: *International Journal of American Linguistics* 76.4: 477–515.
- Méndez-Arocha, Alberto (1959). "Un vocabulario Yabarana con Apuntes Fonéticos". In: *Antropológica* 7: 65–84.
- Oquendo, Luis (2004). "Is Japreria a Yukpa Dialect?" In: *Lingua Americana* 8.14: 87–98.

- 745 Pachêco, Frantomé B. (1997). “Aspectos da gramática Ikpeng (Karíb)”. MA thesis. Universidade Estadual de Camp-
 746 inas.
- 747 — (2001). “Morfossintaxe do verbo Ikpeng (Karíb)”. Ph.D. dissertation. Universidade Estadual de Campinas.
- 748 Payne, Thomas E. & Doris L. Payne (2013). *A Typological Grammar of Panare: A Cariban Language of Venezuela*.
 749 Brill’s Studies in the Indigenous Languages of the Americas. Leiden: Brill.
- 750 Revell, Liam J. (2012). “phytools: An R package for phylogenetic comparative biology (and other things)”. In:
 751 *Methods in Ecology and Evolution* 3: 217–223.
- 752 Robayo Moreno, Camilo Alberto (1989). ““En faisant une rame”: texte carijona”. In: *Amerindia* 14: 189–199.
- 753 — (2000). “Avance sobre morfología carijona”. In: *Lenguas indígenas de Colombia: una visión descriptiva*. María
 754 Stella González de Pérez & María Luisa Rodríguez de Montes (eds.). Santafé de Bogotá: Instituto Caro y
 755 Cuervo: 171–180.
- 756 Sapién, Racquel-María et al. (2021). “Antipassive and semantic classes of verbs in the Cariban family”. In: *The*
 757 *Multifaceted Nature of Antipassive*. Katarzyna Janic & Alena Witzlack-Makarevich (eds.). Amsterdam: John
 758 Benjamins.
- 759 Stegeman, Ray & Rita Hunter (2014). *Akawaio-English Dictionary*. SIL International. URL: [https://akawaio.](https://akawaio.webonary.org/)
 760 [webonary.org/](https://akawaio.webonary.org/).
- 761 Stolz, Thomas, Aina Urdze & Johan van der Auwera (2012). “Introduction: Irregularity – glimpses of a ubiqui-
 762 tous phenomenon”. In: *Irregularity in Morphology (and beyond)*. Thomas Stolz et al. (eds.). Berlin: Akademie
 763 Verlag: 7–38. DOI: 10.1524/9783050059587.
- 764 Swiggers, Gisela (2010). “Gramática de Referencia del Pemón”. Ph.D. dissertation. Universidad del Zulia.
- 765 Tavares, Petronila da Silva (2005). “A grammar of Wayana”. Ph.D. dissertation. University of Houston.
- 766 von den Steinen, Karl (1892). *Die Bakairi-Sprache: Wörterverzeichnis, Sätze, Sagen, Grammatik; mit Beiträgen zu*
 767 *einer Lautlehre der karaïbischen Grundsprache*. Leipzig: Koehler.
- 768 Wheatley, James (1969). “Bakairi verb structure”. In: *Linguistics* 47.47: 80–100.

769 A Predictions for the behavior of individual verbs

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

Table 27: Predictions for Proto-Waiwaian

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

Table 28: Evaluating predictions for Proto-Waiwaian

	<i>*ka[s]</i> 'say'	<i>*[i]to[m]</i> 'go'	<i>*ah</i> 'be-1'	<i>*efi</i> 'be-2'	<i>*eephi</i> 'bathe'	Score
DETRZ	✓	✓	✓	✓	✓	100.0%
DETRZ+freq	✓	✓	✓	✓	✓	100.0%
phono+freq	✓	✓	✓	✓	✓	100.0%
infl+freq	✓	✓	✓	✓	✓	100.0%
phono	✓	✓	×	×	✓	60.0%
infl	×	×	×	×	✓	20.0%

Table 29: Predictions for Proto-Pekodian

	<i>*ke</i> 'say'	<i>*ita[n]</i> 'go'	<i>*ap</i> 'be-1'	<i>*effi</i> 'be-2'	<i>*epi</i> 'come'	<i>*ipta</i> 'go down'	<i>*ipi</i> 'bathe'
DETRZ	×	×	×	×	×	×	×
DETRZ+freq	×	×	×	×	×	×	×
phono (/ _ *a, e)	×	×	×	✓	✓	×	×
phono+freq	×	×	×	×	×	×	×
infl (*w-)	✓	✓	✓	✓	✓	✓	✓
infl+freq	×	×	×	×	×	✓	✓

Table 30: Evaluating predictions for Proto-Pekodian

	<i>*ke</i> 'say'	<i>*ita[n]</i> 'go'	<i>*ap</i> 'be-1'	<i>*effi</i> 'be-2'	<i>*epi</i> 'come'	<i>*ipta</i> 'go down'	<i>*ipi</i> 'bathe'	Score
DETRZ	✓	✓	✓	✓	✓	✓	✓	100.0%
DETRZ+freq	✓	✓	✓	✓	✓	✓	✓	100.0%
phono+freq	✓	✓	✓	✓	✓	✓	✓	100.0%
phono	✓	✓	✓	×	×	✓	✓	71.4%
infl+freq	✓	✓	✓	✓	✓	×	×	71.4%
infl	×	×	×	×	×	×	×	0.0%

Table 31: Predictions for Proto-Tiriyoan

	<i>*ka</i> 'say'	<i>*[a]ta[mi]</i> 'go'	<i>*a</i> 'be-1'	<i>*e?i</i> 'be-2'	<i>*a?epi</i> 'come'	<i>*epi</i> 'bathe'
DETRZ	×	×	×	×	✓	✓
DETRZ+freq	×	×	×	×	×	✓
phono (/ _ *a, e)	×	×	×	✓	✓	✓
phono+freq	×	×	×	×	×	✓
infl (*w-)	✓	✓	✓	✓	✓	✓
infl+freq	×	×	×	×	×	✓

Table 32: Evaluating predictions for Proto-Tiriyoan

	<i>*ka</i> 'say'	<i>*[ə]tə[mi]</i> 'go'	<i>*a</i> 'be-1'	<i>*eʔi</i> 'be-2'	<i>*əʔepi</i> 'come'	<i>*epi</i> 'bathe'	Score
DETRZ+freq	✓	✓	✓	✓	✓	✓	100.0%
phono+freq	✓	✓	✓	✓	✓	✓	100.0%
infl+freq	✓	✓	✓	✓	✓	✓	100.0%
DETRZ	✓	✓	✓	✓	×	✓	83.3%
phono	✓	✓	✓	×	×	✓	66.7%
infl	×	×	×	×	×	✓	16.7%

Table 33: Predictions for Akuriyó

	<i>ka</i> 'say'	<i>[ə]tə[mi]</i> 'go'	<i>a</i> 'be-1'	<i>eepti</i> 'come'	<i>i[h]tə</i> 'go down'	<i>epi</i> 'bathe'
DETRZ	×	×	×	✓	×	✓
DETRZ+freq	×	×	×	×	×	✓
phono (/ _ə)	×	×	×	×	×	×
phono+freq	×	×	×	×	×	×
infl (<i>k</i> -)	×	×	×	×	×	×
infl+freq	×	×	×	×	×	×

Table 34: Evaluating predictions for Akuriyó

	<i>ka</i> 'say'	<i>[ə]tə[mi]</i> 'go'	<i>a</i> 'be-1'	<i>eepti</i> 'come'	<i>i[h]tə</i> 'go down'	<i>epi</i> 'bathe'	Score
phono	✓	✓	✓	✓	✓	✓	100.0%
phono+freq	✓	✓	✓	✓	✓	✓	100.0%
infl	✓	✓	✓	✓	✓	✓	100.0%
infl+freq	✓	✓	✓	✓	✓	✓	100.0%
DETRZ+freq	✓	✓	✓	✓	✓	×	83.3%
DETRZ	✓	✓	✓	×	✓	×	66.7%

Table 35: Predictions for Carijona

	<i>ka</i> 'say'	<i>tə[mə]</i> 'go'	<i>a</i> 'be-1'	<i>effi</i> 'be-2'	<i>eh[i]</i> 'come'
DETRZ	×	×	×	×	×
DETRZ+freq	×	×	×	×	×
phono (/ _ə, e)	×	×	×	✓	✓
phono+freq	×	×	×	×	×
infl (<i>*w</i> -)	✓	✓	✓	✓	✓
infl+freq	×	×	×	×	×

Table 36: Evaluating predictions for Carijona

	<i>ka</i> 'say'	<i>tə[mə]</i> 'go'	<i>a</i> 'be-1'	<i>effi</i> 'be-2'	<i>eh[i]</i> 'come'	Score
phono	✓	✓	✓	✓	✓	100.0%
DETRZ	✓	✓	✓	×	×	60.0%
DETRZ+freq	✓	✓	✓	×	×	60.0%
phono+freq	✓	✓	✓	×	×	60.0%
infl+freq	✓	✓	✓	×	×	60.0%
infl	×	×	×	✓	✓	40.0%

Table 37: Predictions for Yukpa

	<i>to</i> 'go'	<i>a</i> 'be-1'	<i>e</i> 'be-2'
DETRZ	×	×	×
DETRZ+freq	×	×	×
phono (/ _V)	×	✓	✓
phono+freq	×	×	×
infl (*w-)	✓	✓	✓
infl+freq	×	×	×

Table 38: Evaluating predictions for Yukpa

	<i>to</i> 'go'	<i>a</i> 'be-1'	<i>e</i> 'be-2'	Score
phono	✓	✓	✓	100.0%
infl	×	✓	✓	66.7%
DETRZ	✓	×	×	33.3%
DETRZ+freq	✓	×	×	33.3%
phono+freq	✓	×	×	33.3%
infl+freq	✓	×	×	33.3%