The evolution of subject-verb agreement in Eastern Tukanoan

Thiago Costa Chacon
University of Brasilia

Lev Michael
University of California, Berkeley

Abstract

This paper describes the evolution of past/perfective subject-verb agreement morphology in the Tukanoan family, reconstructing relevant aspects of Proto-Tukanoan verbal morphology and delineating the subsequent diachronic development of verbal subject agreement morphology in the Eastern branch of the family. We argue that suffixes that cumulatively expone past/perfective and person, number and gender (PNG) subject agreement resulted from the fusion of post-verbal demonstratives/pronouns expressing PNG information with suffixes expressing past/perfective TAM information. We propose that different PNG agreement categories developed at successive stages in the diversification of the family, with third person masculine singular subject agreement vs. all other PNG categories emerging first, followed animate plural agreement, and finally, the development of third person feminine agreement. The result in Eastern Tukanoan was a cross-linguistically unusual agreement system that contrasts four agreement categories: 1) 1st and 2nd person singular and 3rd person inanimate (singular and plural); 2) 3rd person animate masculine singular; 3) 3rd person animate feminine singular; and 4) 3rd animate plural.

KEYWORDS: Tukanoan languages; agreement; morphological change; morphologization; diachronic morphosyntax.

1. Introduction

The languages of the Tukanoan family exhibit verbal agreement paradigms with cross-linguistically unusual patterns of syncretism, where 1st and 2nd person singular, 3rd person inanimate singular and plural, and zero-valence verbs (e.g. weather verbs), trigger the same "agreement", as illustrated in (1) for Kubeo.[1] 3rd person singular masculine and feminine, and third person plural animate are distinct from the first

[1] List of abbreviations used in this paper: 1/2/3IN '1st and 2nd person animate and 3rd person inanimate', 3.AN.PL '3rd person animate plural', 3.FEM '3rd person feminine', 3.MSC '3rd person masculine', 3.PL '3rd person plural', ANAPH 'anaphora', CL 'classifier', CL.AN.COL 'classifier collective animate', CV 'generic syllable', DEM 'demonstrative', DIST.PST 'distant past', ET 'Eastern Tukanoan', INCL 'inclusive', LOC 'locative', LOC.SP 'locative specific', O 'patient of a transitive verb', PET 'Proto-Eastern-Tukanoan', PL 'plural', PNG 'person, number, gender', PRCS 'precisely', PRES 'present', PRO 'pronoun', PT 'Proto-Tukanoa', PWT 'Proto-Western-Tukanoan', REC.PST 'recent past', REG.PST 'regular past', RESUMPT 'presumptive', S 'subject of a transitive or intransitive verb', SG 'singular', ST 'stative', TAM 'tense, aspect, modality', TOP 'topic', V 'verb', WT 'Western Tukanoan'

syncretic set, and from each other, as in the Kubeo examples in (2). These agreement patterns are summarized in Table 1.

(1a) hawe ã-**wɨ** yɨ
already eat-1/2/3IN I
'I have eaten already.'

(1b) mɨ ke a-**wɨ**-ra
you thus say-1/2/3IN-PRCS
'You said precisely that.'

(1c) kawabɨtɨ-i-na hápura te-**wɨ** hoa-i
branch fall-ST-LOC.ESP be.heard do-1/2/IN far-LOC
'A falling branch was heard from afar.'

(1d) hawe oka-**wɨ**
already rain-1/2/3.IN
'It has rained already.'

(2a) hi-páko eda-**biko**
my-mother arrive-3.FEM
'My mother arrived.'

(2b) hi-pákɨ **eda-bi**
my-father arrive-3.MSC
'My father arrived.'

(2c) mahe = wɨ eda-mã
our.incl = CL.AN.COL arrive-3.AN.PL
'Our relatives arrived.'

Agreement feature	1/2/3.INAN	3.MSC.SG	3.FEM.SG	3.AN.PL
morpheme	-wɨ	-bi	-biko	-mã

Table 1: Kubeo past tense agreement

The goal of this paper is to provide a diachronic account of the emergence of remarkable agreement paradigms like these in Eastern Tukanoan (ET), focusing specifically on the evolution of the verbal subject agreement paradigm in finite declarative past tense clauses with unmarked evidentiality.[2] The principal argument advanced in this paper is that the development of the unusual syncretic pattern exemplified by Kubeo was triggered by the fusion, in Proto-Tukanoan, of formerly free post-verbal singular masculine pro-forms with verb-final past/perfective suffixes, yielding morphemes that exponed both past/perfective and 3rd person singular masculine features. This initial stage, where 3rd person singular masculine past/perfective forms were marked differently from all other persons, in turn triggered the development of animate plural and feminine agreement categories in different daughter languages. The result was morphologically distinct 3rd person masculine singular, 3rd person feminine singular, and 3rd person plural animate agreement suffixes, while all other PNG combinations retained reflexes of the original past/perfective marking, which effectively became a syncretic default agreement category. This process is schematized in Figure 1.

3.MSC vs. All other PNG values	>	3.MSC 3.AN.PL vs. All other PNG values	>	3.MSC 3.AN.PL 3.FEM vs. All other PNG values	=	3.MSC 3.FEM 3.AN.PL vs. 1 /2/3.IN

Figure 1: Development of agreement categories in ET

It is worth clarifying one point the above overview, namely, that it is unclear whether the Proto-Tukanoan tense/aspect suffix that played a central role in this account expressed past tense or perfective aspect. Reflexes of this morpheme seem to appear in

2 Verbs lacking overt evidential marking are construed as expressing visual, direct or first-hand evidentiality in the majority of Tukanoan languages.

both past-tense and non-past-tense paradigms of modern languages, raising the possibility that it expressed perfective aspect, and not past tense. Nevertheless, the paradigms analyzed in this paper express some type or another of past tense, with the exception of Kubeo, where the paradigm in question expressed either present or past temporal reference, depending on the stativity of the verb stem (Chacon 2012). Resolving the semantics of the Proto-Tukanoan TAM morphology in question is beyond the scope of this paper and we non-commitally refer to the relevant morpheme as a past/perfective suffix. Similarly, for purposes of convenience we refer to the modern suffixes that we discuss as simply 'past tense' agreement suffixes.

Although the focus of this paper is the diachrony of subject agreement suffixes in ET, we will reconstruct aspects of the agreement system to Proto-Tukanoan (PT), since this is immediately relevant to an account of ET subject agreement diachrony. To that end, we make strategic use of Western Tukanoan (WT) data in combination with ET data, but we do not pursue the development of subject agreement in WT languages (see Bruil, this issue, for an account of subject agreement in one WT language, Siona). The agreement systems of WT languages underwent additional diachronic developments, such as leveling with subordinate clause agreement suffixes, which merit separate treatment.

The remainder of this paper is organized as follows. In Section 2 we provide an overview of Tukanoan classification, and of verbal subject agreement in the family. In Section 3 we sketch the evolutionary trajectory that we propose for the development of past/perfective verbal subject agreement in ET, and summarize our arguments for reconstructing particular distinctions to Proto-Tukanoan (PT) and its higher-level daughter languages. In Section 4 we reconstruct the initial segment of the past/perfective agreement suffix, while in Section 5 we reconstruct the vowels found in the various agreement suffixes, and demonstrate that these suffixal vowels correspond to the vowels of demonstratives and pronouns that we reconstruct to PT as having the same PNG values in PT as the agreement suffixes in modern languages. In Section 6 we bring together the results of Sections 4 and 5, and flesh out the arguments for the development of subject agreement in PT and its subsequent development in ET, first sketched in Section 3. We conclude the paper with Section 7.

The representations given in this paper are IPA-based, and depart from the practical orthographies developed for the languages. We have generally opted for broad phonetic transcriptions that allow us to transparently represent major allophones, rather than more abstract phonological representations. One convention we have adopted that diverges somewhat from these principles is the use of glottal stop *ʔ* to represent a

laryngeal feature that tends to be realized as creaky voice (cf. Chacon 2014a, Chacon 2015). Thus <*p?*>, a sound we refer to often in this paper, is to be understood as a creaky voiceless bilabial stop, and not an ejective.

2. The Tukanoan Family

2.1. Internal Classification

The Tukanoan family includes some 29 languages, found in a large area extending from: 1) the northern part of the Brazilian-Colombian border region, in the northern and eastern extreme of their distribution; to 2) the Ecuadorean-Colombian upper Putumayo River basin in their western extreme; and to 3) the Peruvian Napo River Basin, in their southern extreme. Two major branches of the family, Western and Eastern branches, are recognized, with their members separated by hundreds of kilometers. The most up-to-date classification of the family, due to Chacon and List (2016), is given in Figure 2, and the distribution of the languages in Figure 3.[3] The reconstruction of PT consonants assumed in this paper, fundamental for the reconstruction of Tukanoan subject-verb agreement morphology, was originally proposed by Chacon (2014a).

[3] Language names abbreviations used in this paper: BAR 'Bará', BAS 'Barasana', DES 'Desano', KAR 'Karapana', KOR 'Koreguahe', KUB 'Kubeo', KUE 'Kueretu', MAI 'Maihĩki', MAK 'Makuna', PIR 'Waikhana (Pira-Tapuyo)', PIS 'Pisamira', SEK 'Sekoya', SIO 'Siona', SIR 'Siriano', TAN 'Tanimuka', TAT 'Tatuyo', TUK 'Tukano', TUY 'Tuyuka', WAN 'Kotiria (Wanano)', YUP 'Yupua', YUR 'Yuruti'

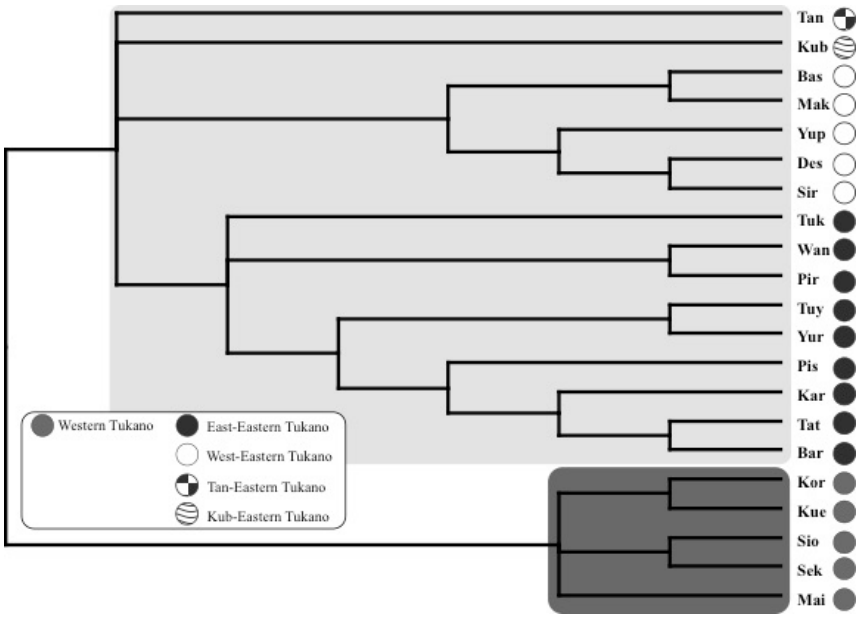

Figure 2: Classification of the Tukanoan family with abbreviations used in this paper

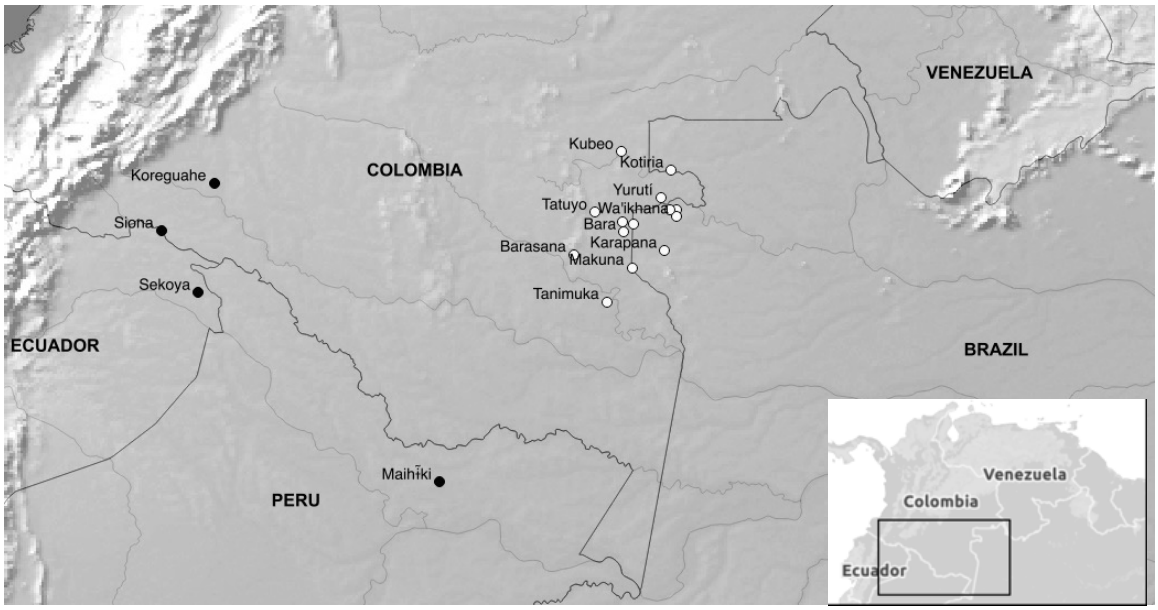

Figure 3: Location of the Tukanoan languages: ET in white and WT in black

Internal subgroups within the ET branch will be important for the discussion of the development of verbal agreement. The set of all ET languages to the exclusion of Kubeo and Tanimuka is here classified as *Nuclear ET*. This clade is further subdivided in two subgroups of languages: the *Western-ET* subgroup (Desano, Barasano, Makuna, Siriano

and Yupua) and the *Eastern-ET* one (the remaining languages). Similarities between Kubeo and WT languages have been noted since Waltz and Wheeler (1972), but shared innovations in phonology and verbal morphology between Kubeo, Tanimuka, and Nuclear ET languages indicate that it is a member of the ET branch (Chacon 2014a for phonology, and this paper on verbal morphology).

2.2 Subject agreement in ET languages

Tukanoan languages are mainly agglutinative, dependent-marking, suffixing languages. Some members of the family have recently developed subject agreement prefixes (see Chacon 2014b), in addition to retaining the subject agreement suffixes that are the focus of this paper. In general, subject agreement suffixes tend to cumulatively expone a variety of categories, including tense, aspect, evidentiality, and sentential mood, in addition to person, number and gender subject agreement features.

Eastern Tukanoan languages vary in the TAM categories marked on the verb, with some, like Kubeo, exhibiting a two-way present-recent past vs. generic-distant past contrast, and others, like Karapana, distinguishing present tense and three past tenses (recent, regular, historical). Nevertheless, the quantity of evident cognacy, and the regularity of correspondences across paradigms of most languages of the family speak to age of many components of the TAM system.

Tukanoan languages are famous for their evidential systems, but this category exhibits considerably less regularity across the family, and is more elaborated in ET languages than in WT ones (cf. Bruil 2014, Stenzel and Gomez-Imbert 2018). Depending on the language, evidentiality may be expressed by periphrastic constructions, by dedicated suffixes preceding tense and agreement suffixes, or by portmanteaux forms expressing evidentiality, tense and subject agreement. For the purposes of this paper, however, we restrict our attention to verbal forms with unmarked visual/direct/first-hand evidentiality, and we do not consider the diachrony of evidentiality further.

In Table 2 we present the TAM paradigms that are the basis of our analysis. In Appendix 1, we present 45 paradigms from 17 Tukanoan languages, constituting an exhaustive overview of all subject-verb paradigms in present and past tenses with the morphologically unmarked visual/direct/first-hand evidentiality in declarative clauses.

Pattern	Language	TAM	1/2/3.IN	3.MSC	3.FEM	3.AN.PL
1.1	Kubeo	pres/rec.pst	-wɨ	-bi	-biko	-mã
1.2	Desano	dist.pst	ˊ-bɨ	ˊ-mĩ	ˊ-mõ	ˊ-mã
	Makuna	rec.pst	-bɨ	-mĩ	-mõ	-mã
1.3	Tatuyo	rec.pst	-wɨ	-wĩ	-wõ	-wã
	Tuyuka	dist.pst	-wɨ	-wi	-wo	-wa
	Yurutí	dist.pst	-wɨ	-wi	-go	-wa
2	Karapana	reg.pst	-wɨ ~ -pɨ	-wĩ ~ -mĩ	-wõ ~ -mõ	-wã ~ -mã
	Tukano	dist.pst	-wɨ ~ -pɨ	-wĩ ~ -mĩ	-wõ ~ -mõ	-wã ~ -mã
3	Barasano	rec.pst	-bɨ ~ -hɨ	-mĩ	-mõ	-mã

Table 2: Cognate past tense subject agreement suffixes in Eastern Tukanoan

Turning to the organization of the paradigms in Table 1, a number of generalizations emerge from inspecting them. First, the general pattern of paradigmatic organization is the same in all languages, with 1/2/3.IN distinct from 3.MSC, and the latter distinct from 3.FEM and 3.AN.PL[4]. For reasons that will become clear, we often refer to the 1/2/3.IN category as the 'default' category. Second, we see that there are languages that exhibit fortis-lenis alternations within a given agreement category, i.e. Karapana, Tukano and Barasano (Patterns 2 and 3), and those that do not (Pattern 1).[5] Third, many languages share the same or very similar morphemes across many cells, with a greater diversity of forms in 3.FEM. On the basis of these observations, we can group the agreement patterns into a small number of schematized major types, given in the leftmost column of Table 1.

Significantly, the general organization of the paradigms given in Table 1 is mirrored in paradigms associated with other TAM values for the languages in question. As we suggest in the conclusion to this paper in section 6, the development of the paradigm structure given in Table 2 had ramifications for other subject verb agreement paradigms.

4 Note that here and elsewhere, we use the 3.IN category to indicate both '3rd Inanimate Singular' predicates without arguments. 1st and 2nd person are pragmatically animate by default. 3.MSC and 3.FEM are singular and animate by default.

5 In Tukano and Barasano the alternations are found across different TAM paradigms or when verbs are combined with motion suffixes (see section 4).

Finally, although our goal in this paper is to address the development of subject agreement suffixes in ET, and not WT, we present the counterparts of the paradigms in Table 3 for the WT languages (with the exception of Koreguahe, whose subject agreement has diverged radically from the general Tukanoan profile), since we will make reference to them below in making certain analytical decisions.

Language	TAM	1/2/PL	3.MSC.SG + 3.IN.SG	3.FEM.SG
Máíhĩki	pst	-bɨ/gɨ ~ -hɨ/-kɨ/-ko[6]	-gɨ~-kɨ	-go~-ko
Siona	rec.pst	-wɨ~-ɨʔɨ	-pʔi~-hiʔi (-hVʔi)	-kʔo(o)~-koʔɨ
Sekoya	rec.pst	-wɨ~-ɨʔɨ	-pi~-hiʔi	-ko~-koʔɨ

Table 3: Cognate past tense subject agreement suffixes in Western Tukanoan

Note that WT languages exhibit a similar, but not identical, paradigmatic structure as ET in the corresponding subject agreement paradigms. In particular, in WT the 1/2/PL category embraces 1.SG and 1.PL, 2.SG and 2.PL, and 3.PL, conflating animate and inanimate plurals in the same category, which constitutes distinct agreement categories in ET. In addition the 3.INAN.SG category is conflated with the 3.MSC.SG category, rather than with the first and second person singular categories, as in ET languages.

Nonetheless, the ET 1/2/3.INAN agreement suffixes are cognate with the WT 1/2/3.PL agreement suffixes, and the ET 3.MSC suffixes in ET are cognate with the WT 3.MSC + 3.IN.SG ones, with the exception of the Máíhĩki suffixes. The morphemes of the 3.FEM.SG category are not in general cognate, however, and the 3.AN.PL agreement category is found only in ET.

Note that while the paradigmatic organization of the past tense subject agreement paradigm in WT languages for lexical verbs is somewhat different from that of the corresponding paradigms in ET languages, there is a sub-domain of WT agreement system that shows greater similarity to ET agreement in its paradigmatic organizations, namely, copula agreement. As can be seen in Table 4, in this domain, 1st, 2nd and 3rd inanimate pattern with the default category, as is the case with ET lexical verb paradigms, rather than with 3.MSC, as it does in WT lexical verb agreement paradigms. Notice, also, that 3.PL is also coded by the same syncretic forms. As we discuss below,

6 In contrast to other Tukanoan languages, 2nd person verbal agreement in Máíhĩki is conditioned by gender in the same way as third person forms.

this parallel between WT and ET allows us to reconstruct with confidence that the basic distinction that emerged in PT was between 3.MSC.SG animate and a default category.

Language	1/2/3.INAN/3.PL	3.MSC	3.FEM
Máíhĩki	-hã	agɨ	ago
Siona	ʔɨ	ʔpʔi	ʔko
Sekoya	a-ʔɨ ~ -ʔɨ	a-pi	a(-ø)
Koreguahe	a-me	-ʔ-mɨ	-ʔ-mo

Table 4: Copula Agreement in WT

3. Overview of Evolution of Tukanoan Subject Agreement Morphemes

Inspection of the paradigmatic organization in Table 2 reveals a number of patterns. First, with the exception of the Yuruti 3.FEM *-go*, all ET subject agreement suffixes consist of a bilabial consonant followed by as single vowel (Kubeo 3.FEM exhibits an additional syllable) that varies in quality, suggesting that they share a common base. Similarly, within each PNG category, the quality of the vowel in the suffixes is identical. The relationship between PNG categories and suffix vowels in turn suggests that PNG categories are expressed by the suffix vowel in some way (as noted by Malone [1988]). As we shall discuss in further detail in section 5, the vowels in question are in fact those commonly found expressing those same PNG values in pronouns and demonstratives. Specifically, *ĩ* is associated with the masculine, *o* with feminine, and *ã* with animate plural. These facts lead us to conclude that ET subject agreement suffixes arose from the fusion of a base expressing the relevant TAM value with nominal elements expressing PNG values.

The evolutionary trajectory that we propose for the development of these paradigms is as follows. We posit that in Pre-Proto-Tukanoan (Pre-PT), verbs did not exhibit subject agreement, but instead bore TAM suffixes that did not cumulatively expone any PNG features. We reconstruct this suffix as exhibiting a fortis-lenis alternation like the one we see in some languages in Table 2, with the forms **pɨ* (fortis) ~ **pʔɨ* (lenis), where in the latter case, the glottalization was realized as creaky voice.[7] The choice of the final *ɨ* vowel will be discussed in section 5.

7 Note that similar alternations are attested in WT languages (see Table 3).

The 3.MSC agreement suffix is the (non-default) PNG category with cognates present in the most languages, in both ET and WT, with other categories, such as 3.FEM, exhibiting less widely-shared cognacy. The suffixes that express the 3.MSC category thus exhibit two important characteristics: 1) they are clearly cognate across the family, unlike the 3.FEM suffixes, which split into several distinct cognate sets; and 2) the quality of the final vowel on these suffixes is similar to the vowel quality of nominal roots found in pronominal and demonstrative elements that denote 3rd person masculine animate referents.

This leads us to hypothesize that 3.MSC was the first category to morphologize, and that it did so in PT, prior to further diversification in the family. We take the variation in 3.FEM to be evidence that this category morphologized subsequent to the split of ET and WT, and even subsequent to the highest-level internal diversification in ET (see, e.g., the divergent 3.FEM form in KUB and YUR; cf. section 6). Likewise, we take the difference in how 3.AN.PL is expressed between ET and WT – in particular, that in ET, this category is expressed by a distinct morpheme, while in WT, it is part of the default category – to be evidence that this category morphologized subsequent to the ET-WT split. See Figure 1 for a schematization of this trajectory.

We conclude this overview with a brief discussion of the source construction which allowed for the fusion of the PNG-expressing nominal elements with the verbal TAM suffix. Tukanoan languages exhibit two common basic constituent orders: S(O)V and (O)VS. The fusional process we sketch above requires that the nominal element expressing the PNG values of the subject appeared post-verbally in PT, suggesting that PT exhibited (O)VS order. There are two constructions that could have served as the environment for TAM-PNG fusion: 1) a simple clause in which the subject is realized as pronominal or demonstrative element, as in (3); or 2) a topicalization construction in which the subject appears clause-initially, with a resumptive pronominal element post-verbally, as in (4). Both constructions are commonly attested in modern Tukanoan languages.

(3) (O) V S_{PRO}

(4) S_{TOP} (O) V $PRO_{RESUMPT}$

For our present purposes, it is immaterial whether we identify (3) or (4) as the source construction. In either case, the nominal element bearing the 3.MSC feature fused with the TAM suffix. As we discuss in detail in section 6, similar processes led to the

emergence of the 3.FEM and 3.AN.PL agreement categories in ET, leaving the 1/2/3.IN category as the default agreement category, which never experienced fusion with a PNG-bearing nominal element. Subsequent sound changes, and the loss of prosodically-conditioned phonological alternations, discussed in section 6, resulted in the agreement systems found in modern languages of the family.

4. The PT agreement suffix: initial consonant

The purpose of this section is to show that the initial consonant of the past tense agreement suffixes in ET are reflexes of a pair of bilabial consonants related to each other by a fortis-lenis alternation. To this end we reconstruct the two alternants and their conditioning factor. In particular we reconstruct the alternation *p ~*pʔ, with the conditioning factor being the moraic quantity of the suffix group in which the past tense suffix appeared, with the fortis alternant conditioned by suffix groups with an even number of moras, and the lenis one by suffix groups with an odd number of moras.

We begin by re-presenting in Table 5 the default and the 3.MSC subject morphemes given in Tables 2 and 3 in a manner that makes it easier to examine the relevant consonant correspondences. To facilitate comparison, we have stripped out non-essential information about prosody, eliminated the non-cognate morphemes in the Máíhɨ̃ki paradigm, and highlighted WT languages in grey.[8]

	KUB	TUY, YUR	TAT	KAR, TUK	DES, MAK	BAS	MAI	SIO, SEK
Default	-wɨ	-wɨ	-wɨ	-wɨ ~ -pɨ	-bɨ	-bɨ ~ -hɨ	-bɨ ~ -hɨ	-wɨ~ -ɨʔɨ
3.MSC	-bi	-wi	-wĩ	-wĩ ~ -mĩ	-mĩ	-mĩ	–	-pʔi ~ -hiʔi

Table 5: Comparison of default and 3.MSC agreement morphemes

[8] We here offer some comments about the phonetic realization of sufixes: in ET languages, voiced sufixes can become nasalized, as in the case of the Kubeo 1/2/3.IN agreement *-wɨ ~ - w̃ɨ̃*, and the 3.MSC agreement suffix *-bi ~ mĩ*. The phonetic realization of Sekoya *pʔ* in agreement suffixes is unclear (compare Piaguaje et al. 1992; Vallejos 2013:85). In Siona *pʔ* in these contexts appears to surface as either [b] or [β] in suffix initial position (see Bruil 2014; and Bruil this issue.

We next make two observations about paradigm structure. First, both ET (Karapana, Tukano and Barasana) and WT (Maihɨ̃ki, Siona, and Sekoya) languages exhibit fortis-lenis alternations, leading us to reconstruct the fortis-lenis alternation to PT. We assume that in the ET languages that do not exhibit these alternations, the alternation was lost,[9] and that only one member of the alternating pair was retained.

We now reconstruct the segmental content of the fortis and lenis alternation by examining the segmental correspondences – both fortis and lenis – for the default and 3.MSC suffixes, given in Tables 6 and 7, respectively.

Default	PT	KUB	TUY, YUR, TAT	KAR, TUK	DES, MAK	BAS	MAI	SIO, SEK
Fortis	*p	-	-	p	-	h	h	Ø
Lenis	*pʔ	w	w	w	b	b	b	w

Table 6: Segmental correspondences for fortis and lenis forms of the default agreement suffix

For the default agreement suffix in Table 6, we reconstruct a *p ~*b alternation, relying on Chacon's (2014a) reconstruction, which shows that *$p > h > Ø$* and *$pʔ > b > w$* are widely attested processes in the family. As Chacon argues in that work, however, the *p-b* segmental contrast did not, in fact, exist in this precise form until PET, with the original PT contrast being *p-pʔ*. We thus reconstruct the alternation found in the initial segment of the default agreement suffix in PT as **p (fortis) ~ **pʔ (lenis).

Turning to the 3.MSC agreement suffix, the alternations present in Karapana, Tukano, Siona and Sekoya forms lead us to reconstruct the same alternation for this suffix, with the Siona and Sekoya forms being especially informative, preserving the lenis *pʔ* and exhibiting a debuccalized reflex of the fortis *p*. Note that ET languages do not exhibit a voiceless reflex of the fortis form in these paradigms (although such reflexes are found in other paradigms that we do not examine in this paper for reasons of space). As will be discussed in section 6, nasalization brought about by the vowel *ĩ* in the 3.MSC form triggered further leniting sound changes, resulting in voiced reflex of the fortis consonant.

9 Note that for Kubeo, Desano, Makuna, Tuyuka, Yurutí, Tatuyo, and Yukuna, the alternation was lost for both default and 3.MSC agreement categories, while in BAS, it was lost only for 3.MSC.

3.MSC	PT	KUB	TUY, YUR, TAT	KAR, TUK	DES, MAK	BAS	MAI	SIO	SEK
Fortis	*p	b	-	m	m	m	-	h	h
Lenis	*pʔ	-	w	w	-	-	-	pʔ	p

Table 7: Segmental correspondences in fortis and lenis forms of the 3.MSC agreement suffix

We conclude this section with a brief discussion of the factors conditioning the fortis and lenis alternation, reconstructing the conditioning factor to be the moraic quantity of the suffix group. We begin by looking at the modern languages which exhibit the alternation productively. Turning first to Karapana, we find that this language exhibits the *–pɨ* ~ *–wɨ* alternation in the distant past, and that it is conditioned by the moraic quantity of the suffix group, with the fortis *–pɨ* occuring when moraic quantity of the suffix group is even (5a&b), and the lenis *–wɨ* when the count is odd (Metzger 1981: 61, 73), as in (5c&d).

(5) a. pa-e[10]-**pɨ** (Karapana)
work-NEG-PST.DIST
'I have not worked.'

b. eta-a-**pɨ**
arrive-IMMED-PST
'I just arrived.'

c. pa-**wɨ**
work-PST.DIST
'I have worked.'

d. wati-koa-***wɨ***
break-COMPL-PST.DIST
'It completely broke.'

10 In Karapana the negative suffix has two allomorphs: *–eti,* when the number of moras in the suffix group is even, and –e when it is odd (see evidence in Metzger 1981:62-73).

In Barasano and Tukano, evidence of alternations conditioned in similar ways are also to be found, although this evidence requires comparing different TAM values. For example, in Tukano, the fortis-lenis alternation is attested in the alternation between the recent past *baa-a-pɨ* 'I/you/it swam (recently)', with a fortis form in an even mora suffix group, and the distant past *baa-wɨ* 'I/you/it swam (long ago)', with a lenis form in an odd mora suffix group. Similarly, in Barasano (Jones 1991:85-6), in the default agreement category, the distant past is expressed by the even mora suffix sequence *-ka-hɨ,* where agreement suffix surfaces with the fortis alternant as predicted (*p > h in Barasano), while the immediate past is expressed by the single monomoraic the lenis alternant *-bɨ*, as would be predicted of an odd mora suffix group. In both the Tukano and Barasano cases, the additional suffix present in the even mora suffix groups serve to make the relevant tense distinctions (recent vs. distant past tense in Tukano, and the converse in Barasano; similar tense distinction associated with fortis-lenis alternations are described for Siona by Bruil [2014], and this issue).

The WT languages Máíhĩki, Siona, and Sekoya also exhibit traces of the fortis-lenis alternation, although synchronically, the remnants of this alternation are conditioned by verb class. In particular, a class of irregular but high-frequency verbs surface as monomoraic in the past tense, and in the default agreement category take a reflex of the fortis alternant, while regular bimoraic roots take a reflex of the lenis alternant, as exemplified in (6a&b), respectively, for Máíhĩki[11] (Farmer 2015, Michael 2012); see also Bruil (2014) for parallel patterns in Siona.

(6) a. sá-hɨ̀ (Máíhĩki)
go-1/2/PL.PST
'I/you/we/you.PL/they went.'

b. ábɨ́-bɨ́
bathe-1/2/PL.PST
'I/you/it/we/you.PL/they bathed.'

Although the fortis-lenis alternation is now lexically conditioned, i.e. by whether the verb on which the agreement suffix appears belongs to the irregular or regular verb class, we argue that this lexical conditioning arose from the same suffix group moraic quantity rule found in ET languages. In particular, we argue that since regular verbs are

[11] Note *p > h in Máíhĩki.

obligatorily bimoraic in WT languages like Máíhĩki, this led to a reanalysis of the original criterion conditioning the fortis-lenis alternation being suffix group size, as a word size criterion. This occurred, we suggest, because the moraic oddness or evenness of a word is identical to that of the suffix group if the root is bimoraic (assuming, as is the case, that verbs rarely take prefixes), as exemplified in (7).

(7) a. $\sigma\sigma_{ROOT}$□σ□ σ even suffix group and phonological word FORTIS
b. $\sigma\sigma_{ROOT}$□□σ odd suffix group and phonological word LENIS

Once this reanalysis took place, the irregular monomoraic roots bearing the monomoraic default agreement suffix, as in (6a), formed parts of even-mora phonological words, conditioning the fortis alternant, while regular bimoraic roots bearing the same suffix formed part of odd-mora phonological words, conditioning the lenis alternant.

We thus have compelling evidence that in both ET and, historically, in WT, the fortis and lenis alternations for the past tense suffix was conditioned by the moraic quantity of the suffix group, leading us to reconstruct the same conditioning factor to PT.[12]

5. The PT agreement suffix: final vowel

This section has two main goals: 1) to reconstruct the final vowels of the agreement suffixes; and 2) to demonstrate that PNG features of the agreement suffixes in which these vowels appear are associated with the same vowels in demonstratives and pronouns across the family.

5.1 Reconstructing agreement suffix final vowels

In order to identify the major patterns necessary to reconstruct the final vowels in the agreement suffixes, we reproduce in Table 8 the ET agreement suffixes originally presented in Table 2.

[12] As a result, the function of the fortis-lenis alternation in the tense systems of Tukano, Barasana and Siona, as discussed above, are the result of morphologization after the conditioning phonological environment was lost in the languages

Pattern	Language	TAM	1/2/3.I	3.MSC	3.FEM	3.AN.PL
1.1	Kubeo	pres/rec.pst	-wɨ	-bi	-biko	-mã
1.2	Desano	dist.pst	ˊ-bɨ	ˊ-mĩ	ˊ-mõ	ˊ-mã
	Makuna	rec.pst	-bɨ	-mĩ	-mõ	-mã
1.3	Tatuyo	rec.pst	-wɨ	-wĩ	-wõ	-wã
	Tuyuka	dist.pst	-wɨ	-wi	-wo	-wa
	Yurutí	dist.pst	-wɨ	-wi	-go	-wa
2	Karapana	reg.pst	-wɨ ~ -pɨ	-wĩ ~ -mĩ	-wõ ~ -mõ	-wã ~ -mã
	Tukano	dist.pst	-wɨ ~ -pɨ	-wĩ ~ -mĩ	-wõ ~ -mõ	-wã ~ -mã
3	Barasano	rec.pst	-bɨ ~ -hɨ	-mĩ	-mõ	-mã

Table 8: Major patterns in ET past tense subject agreement paradigms

Inspection of Table 8 shows that reconstructing **ɨ* is unproblematic for the vowel in the PT 1/2/3.IN agreement suffix. Similarly, we see that for the 3.MSC category we can reconstruct **ĩ*.

Note that three languages do not exhibit evidence of nasalization in the 3.MSC suffix, Kubeo (1.1), Tuyuka (1.3), and Yurutí (1.3). For Tuyuka and Yurutí, this is part of a general loss of nasalization in the past tense agreement suffixes, as evident in the wholly oral 3.FEM and 3.AN.PL forms. We assume this to be a relatively late process, possibly due to analogy with the default category, and do not discuss it further here. The orality of Kubeo cognate *-bi* and the 3.FEM form *-biko* present a different situation, however, since Kubeo did not lose nasality across the board, as evidenced by the nasal 3.AN.PL form*, -mã*. We return to the oral nature of these Kubeo forms below, explaining it as the result of the details of the fusional process.

There is a greater diversity in the 3.FEM category, reflecting the fact that the 3.FEM category morphologized after the earliest stages of diversification of ET, resulting in at least three distinct grammaticalization events, as evidenced by the existence of three distinct cognate sets: 1) the Kubeo singleton set *-biko*; 2) the Yurutí singleton set *-go*; and 3) the cognate set consisting of the morphemes from the remaining languages. In this latter set, the vowel quality is *o* in all cases, and all but the Tuyuka form is associated with a nasal feature, leading us to reconstruct **õ* for this vowel.

Finally, in the 3.AN.PL category, all the vowels have the same quality, and the vast majority are associated with a nasal feature, leading us to reconstruct **ã*. Table 9 summarizes the preceding results.

PNG feature	reconstructed V	reconstructed to
1.SG/2.SG/3.INAN	*$\dot{\imath}$	PT
3.MSC	*$\tilde{\imath}$	early post-PET
3.FEM	*$\tilde{o}$	later post-PET
3.AN.PL	*$\tilde{a}$	PET

Table 9: Reconstructed final vowels of ET past/perfective subject agreement suffixes

It is useful to note that Table 9 foreshadows an important point that will be addressed in the next section, namely the node in the Tukanoan tree to which we reconstruct the vowels in question. As we shall argue in the next section, although we reconstruct the fusion of the perfective/past with the 3.MSC nominal element to PT, the vowel quality that we reconstruct for this morpheme in Table 9 is the result of simplification processes that took place subsequent to the initial fusional process.

5.2 Final vowels and PNG features

Having reconstructed the final vowels of the agreement suffixes, we now observe that each of the vowels listed in Table 9 are associated with particular PNG features in the nominal domain, with the exception of vowel *i*. In this section we furthermore demonstrate that the final vowels in the 3.MSC, 3.FEM and 3.AN.PL verb agreement morphemes correspond to the vowels of particular PT demonstrative and anaphoric forms, while the *i* of the default category does not.

Our first step is to partially reconstruct the aspects of PT demonstrative and anaphora system. Although a comprehensive comparison of these elements in the Tukanoan family, and the complete reconstruction of these systems in PT, is beyond the scope of this paper, we can reconstruct a number of key features of the pro-forms that served as the basis for the development verbal agreement in PT and PET. The internal reconstructions and tentative morphological analyses discussed here are partially supported by Chacon's (2014a) phonological reconstruction.

We begin by considering a sample of demonstratives and pronouns for a range of Tukanoan languages, presented in Table 10. The languages were selected opportunistically, based on the availability of relevant information, with gray cells indicating WT languages.

	BASE	MASCULINE	FEMININE	AN.PL	PRO-FORM
Tuyuka	*i ~ ĩnĩ̃*	*ĩ́nĩ̃*	*i-go*	*ĩ́nĩ̃-ã̃*	DEM (THAT)
Makuna	*i ~ ĩ̃*	*ĩ̃*	*i-so (*i[y]-o)*	*õa (*ĩ̃-o-wa)*	DEM (THAT)
	i ~ ĩ̃	*ĩ̃*	*i-so (*i[y]-o)*	*i-rã̃*	PRONOUN
Desano	*i ~ ĩ̃*	*ĩ̃-gɨ̃*	*i-go*	*ĩ̃-rã̃*	PRONOUN
	i ~ ĩ̃	*ĩʔĩ̃*	*i-go*	*õa (*ĩ̃-o-wa)*	DEM (THIS)
Kubeo	*i ~ ĩ̃*	*ỹãã́ (*i-ʔĩ̃)*	*i-kó*	*i-ná*	DEM (THIS)
	ĩ̃	*ɨ́ (*ĩ̃-ɨ)*	*ő (*ĩ̃-o)*	*ná*	PRONOUN
Tanimuka	*iʔĩ̃*	*ĩʔí*	*íʔko*	*í-rã̃*	PRONOUN/ DEM (THAT)
	iʔ	*iʔ-kí*	*iʔ-kó*	*iʔ-rã́*	DEM (THIS)
Sekoya	*i*	*i-kɨ*	*i-ko*	*i-ko-wa'i*	DEM (THIS)
	ĩ̃	*ĩ̃*	*ĩ̃-o*	*ĩ̃-o-wa'i*	PRONOUN
Siona-COL	*ʔɨ̃*	*ʔɨ̃-kɨ*	*ʔɨ̃-ko*	*ʔɨ̃-kwa*	DEM (THIS)
	ĩ̃	*ĩ̃*	*ĩ̃-o*	*ĩ̃-o-wa'i*	PRONOUN
Maihɨ̃ki	*ɨ́*	*ɨ́-kɨ̀*	*ɨ́-kò*	*ɨ́ge*	DEM (THIS)
	ĩ̃	*ĩ̃̂*	*ĩ́ɔ̃̀*	*ĩ́ti huna*	PRONOUN

Table 10: Pronominal and demonstrative forms in a set of Tukanoan languages.

First, we see that the vowels of the demonstratives and pronouns and those of the verbal agreement suffixes in Table 9 largely overlap: *i/ĩ̃* for 3.MSC, *o/õ* for 3.FEM and *ã̃* for 3.AN.PL.

Second, we find evidence that both pronouns and demonstratives are formed by adding PNG-expressing morphemes to a base, and that there is evidence for two bases: *i* and *ĩ̃*. The former appears to have served as the base for demonstratives, and the latter, for pronouns. This is clearest from ET languages like Tanimuka, and especially in the WT language Sekoya, where the distinction between demonstratives and pronouns is retained without any kind of paradigm mixing or other forms of analogical change. However, this strict separation is not found in other languages, where we see both that the demonstrative base *i-* is found in the pronominal paradigm (mostly commonly for

3.FEM and 3.AN.PL) and, conversely, that the pronominal base *ĩ-* occurs in the demonstrative paradigm (usually with 3.MSC).

The most remarkable feature in pronominal systems of ET languages is the emergence of the pro-form *i-ʔĩ, analyzed here as consisting of the compound between demonstrative base *i* and the pronominal base *ĩ,* with an epenthetic glottal stop.[13] Combining demonstrative and the pronominal bases like this was probably motivated by the need for an information structurally marked pronoun.[14] Reflexes of *i-ʔĩ are found in Desano and Tanimuka, but can be reliably reconstructed to Kubeo, Barasano and Tuyuka as these languages historically lost the glottal stop, obscuring the existence of reflexes of this pronoun.[15]

The 3.AN.PL elements are more diverse across ET languages, but generally consist of an alveolar consonant followed by *ã.* The segment *ã* is commonly found in plural morphemes in Tukanoan languages, and is likely the source of this segment in the 3.AN.PL elements in Table 10. The alveolar consonant, in contrast, likely comes from the anaphoric base *ti-* or *di-*, which is found in many ET languages, for example Kubeo *di-jãmi* : Tukano *ti-wi* 'that house' (ANAPH-CL.HOUSE).

---

[13] Note that evidence for the epenthetic glottal stop is absent in WT languages, although we do find some evidence for compound pro-forms. In particular, in Maihĩki the 3rd masculine pronoun is *ĩ̂ĩ̂*, suggesting a compound form. If there had been a glottal stop in the predecessor to this form, however, we would expect an LL tonal pattern, instead of the attested HL pattern, since *ʔ > low tone in this language (Wheeler 1990, Farmer 2012, Chacon 2014a). The form *i-ʔĩ can thus be reconstructed to PET, but not to PWT.

[14] Note that processes like this are cross-linguistically common. For instance, Ramirez (2000:380-3) analyses demonstratives in Baniwa, a neighboring Northern Arawakan language, as composed of a pronominal base plus a deictic base, e.g. *ɽhía 'he, it' + hṽ 'spatial deictic' > ɽhíhẽ 'this one'*. Likewise, in Romance languages, the evolution of deictic forms shows an even more intricate relation between different deictic bases. For instance, in French, the demonstrative pronouns have evolved from emphatic constructions combining more than one pro-form, e.g. *ce-lui-ci "this one (here)"* a compound from Latin **ecce* 'here'+ **illi* 'he (anaphoric)'+ **ecce* 'here'+ **hic 'this (demonstrative)'* (Guillot 2015).

[15] In Tukano, a reflex of the *i?ĩ 3.MSC pronoun can be marginally observed in the synchronic demonstrative system, which combined older pronouns with a new base, such as the proximal demonstrative in Tukano *aʔti-,* as evident in the form *aʔti-go* 'this woman'. The masculine form *ãʔrĩ* in particular resulted from the fusion of the base *aʔti-* with the pronoun *iʔĩ: *aʔti-ĩʔĩ > aʔtĩ > ãʔrĩ.*

Finally, for 3.FEM, although we cannot reconstruct a pro-form to PT and the reconstruction to PET is yet unclear, the overall regularity of verb agreement forms in ET suggest a single common source in Proto-Nuclear-ET.

Table 11 summarizes the reconstruction of PT and PET pro-forms, organized by PNG feature, and the vowels of the corresponding post-PET verbal agreement categories.

PNG	**PT pro-forms**	**PET pro-forms**	**Post-PET verb agreement endings**	**ET clade**
3.MSC	***ĩ-* ***i-*	**i-ʔĩ*	*-ĩ* *-i*	Nuclear ET Kubeo
3.FEM	***ĩ-o* ***i-ko*	**ĩ-o* **i-ko*	*-õ* *-ko*	Nuclear ET Kubeo
3.AN.PL	*?*	**di-ã̀* **i-di-ã̀* **ĩ-o-ã̀*	*-ã̀*	All ET

Table 11: From pro-forms to the forms expressing PNG in verb agreement

We illustrate the relationship between the reconstructed proto-forms and their reflexes in ET with tables summarizing proto-forms and their reflexes in Kubeo, Tanimuka and Desano, as given in Tables 12, 13, and 14, respectively.

PET	PRE-KUBEO	MODERN KUBEO FORMS
**ĩ-i̵* **ĩ-õ* **di-ã̀*	*ĩ-i̵* *ĩ-õ* *nã̀*	*ĩ́* ‘he’ *ṍ* ‘she’ *nã́* ‘they’
**iʔĩ*	*iʔĩ* > *ĩ-ĩ* > *ĩ-[a]ĩ*	*ỹãĩ́* ‘this masculine’
**i-ko* **i-di-ã̀*	*i-ko* *i-nã̀*	ikó ‘this feminine’ iná ‘these animate plural’

Table 12: Evolution of 3rd person pronominal and demonstrative forms in Kubeo

PET BASE FORMS	PRE-TANIMUKA	MODERN TANIMUKA FORMS
**ĩ*		
**iʔĩ*	*iʔĩ́* *iʔĩ́-ko > iʔ̲́-ko > íʔ-ko* *iʔĩ́-nã > iʔ̲́-rã > íʔ-rã*	*ĩʔĩ́* *íʔko* *íʔrã*
**i-kɨ* **i-ko* **i-di-ã*	*i-kí > iʔ-kí* *i-kó > iʔ-kó* *i-nã > iʔ-nã́*	*iʔkí* *iʔkó* *iʔnã́*

Table 13: Evolution of 3rd person pronominal and demonstrative forms in Tanimuka

PET BASE FORMS	PRE-DESANO	MODERN DESANO FORMS
**ĩ*	*ĩ-gɨ* *ĩ-o-ã*	*ĩgɨ̃* *õa*
**iʔĩ*	*iʔĩ*	*ĩʔĩ*
**i*	*i-go* *i-nã*	*igo* *irã*

Table 14: Evolution of 3rd person pronominal and demonstrative forms Desano

Kubeo and Desano show reflexes of the three base forms, while Tanimuka of only two base forms. In pre-Kubeo and pre-Desano there is evidence that the pronominal base *ĩ was inflected by gender classifiers, as it is actually the case with the WT languages in Table 10. All forms derived from *ĩ are nasal and those derived by *i keep an oral base in Desano and Kubeo. The forms derived from *iʔĩ are the sole bimoraic forms before affixation (they cannot be inflected in modern Kubeo and Desano). The Kubeo *ỹãĩ́* 'this masculine' form has en epenthetic [a] after the language merged **ʔ* with zero, creating an environment for the sequences of two /i/ vowels, which is forbidden in the language.[16]

As for Tanimuka, there is an interesting contrast based on stress placement, as analyzed by Eraso (2014), and also supported by Strom's (1992) analysis of the closely related language Retuarã. The proximal demonstrative forms have stress on the suffix, whereas the pronominal and distal demonstrative forms have stress on the base. Such a

16 A sequence of *i* vowels is resolved by an epenthetic *a* in Kubeo verbal morphology as well, e.g., *da-ima* 'they came (distant past)', *da-i-ima > da-i[a]-ima* 'they customarily come (generic)'.

pattern can be explained if we recognize a bimoraic origin of the rhizotonic forms, namely *i?ĩ, where stress would fall within the bimoraic foot contained in the base, and a monomoraic origin of the arhizotonic forms, namely *i, where a bimoraic foot could only be satisfied after affixation. Denasalization of feminine and plural forms derived from *i?ĩ were motivated by the deletion of the nasal vowel in order to keep words to a prosodic minimality constraint of two mora (cf. Eraso 2014). The nasal vowel was lost, but stress was relocated to the left syllable, preserving the stress within a clear iambic foot template. The glottal epenthesis in the forms derived from *i are interpreted as the result of analogy with the paradigm derived from *i?ĩ: speakers used stress as the main contrastive features of the two paradigms, and thus leveled all forms initiating by the same vowel *i* with the epenthesis of glottal stop.

Finally, the agreement suffix vowel *ɨ* has a distinct history, in that this vowel quality does not correspond to a coherent PNG category, but seems a residual or default form, which corresponds to the remaining 1.SG/2.SG/3.IN features, after the morphologization of forms referring to 3.MSC, 3.FEM and 3.AN.PL. This suggests that default category morpheme was not formed from the fusion of the TAM base with a PNG-indexing vowel, but is simply the reflex of the original TAM suffix. In the nominal domain, *ɨ* is found in 'masculine (non-feminine) classifiers' as well as in the 1st and 2nd person singular pronouns reconstructed as **yɨʔɨ* and **mɨʔɨ* to PT, which do not exhibit gender distinctions, contrary to 3rd person pronouns. At this point, we will keep with the assumption that *ɨ* had no PNG value in the development of verb agreement in Tukanoan languages, leaving for the future a more precise analysis of a hypothetical semantic or functional value of this vowel in the proto-language.

6. Diachronic development of subject agreement suffixes

In this section we offer a unified account of the evolution of subject-verb agreement in ET as depicted in Figure 1, drawing on the reconstruction of the initial consonants and vowels of the agreement suffixes, and of the pro-forms in the previous two sections. We begin with the default and 3.MSC categories, the latter being the first category to develop, and then to turn to the 3.FEM and 3.AN.PL categories.

6.1 The evolution of the default and 3.MSC morphemes

The development of complex verb agreement in Tukanoan languages was triggered by the morphologization of 3.MSC verbal agreement, which resulted in a basic distinction between 3.MSC agreement and a residual, default, agreement category. As discussed briefly in Section 3, there are two plausible source constructions for the

morphologization process in question. Assuming that PT exhibited (O)VS word order, as do several of its daughter languages, the two candidate source constructions for the development of 3.MSC agreement are given in schematic form in (8) (ignoring for the moment the fortis-lenis alternations in the suffixes). The first construction, (8a), is a simple main clause with a pronominal/demonstrative *i?ĩ S argument, while the second, (8b), is a construction in which a 3.MSC NP S argument has been topicalized, leaving a resumptive pronoun with the same PNG features in post-verbal position. In both cases, the verb bears the *-*p?ɨ* past/perfective suffix.

(8) a. (O) V-*p?ɨ $i?\tilde{i}_{S}$

b. S_{TOPIC} (O) V-*p?ɨ $i?\tilde{i}_{RESUMPT.}$

Regardless of which construction is selected as the source construction, we posit that the post-verbal pronominal/demonstrative element cliticized to the verb. This led to vowel hiatus between the final vowel of the past/perfective suffix and the initial vowel of the pronominal/demonstrative development, which was eventually resolved by deletion of the suffix vowel, as depicted in (9a&b), for the fortis and lenis forms of the past/perfective suffix, respectively ('σ' indicates a generic piece of suffixal material that gives the suffix group an odd number of morae).

(9) 3.MSC

	pre-PT		PT		PET	
a.	σ-pɨ i?ĩ	→	σ-pɨ=i?ĩ	→	σ-pi?ĩ	FORTIS
b.	-p?ɨ i?ĩ	→	-p?ɨ=i?ĩ	→	-bi?ĩ	LENIS

The default agreement morphology is simply the verb-final past/perfective morphology that was retained in contexts without the post-verbal pronominal/demonstrative element, i.e. with non 3rd masculine singular subject arguments. In (10a&b) we show the development of the default form, in contexts lacking a post-verbal pro-form. The second stage of these processes reflect the **p? > b* change in PET, with *p remaining unchanged.

(10) DEFAULT

	PT		PET	
a.	σ-pɨ	→	σ-pɨ	FORTIS
b.	-pʔɨ	→	-bɨ	LENIS

6.1.1 Kubeo

We now examine the sound changes which indicate a high-level split between Kubeo and Nuclear ET languages, and account for the divergent nature of the Kubeo agreement suffixes. Starting with the default category in Kubeo, we see that there were two changes: 1) a leniting sound change **b > w* in suffixes, leading to *-bɨ > -wɨ* change, schematized in (11a); and 2) the loss of the fortis/lenis alternation, with preservation of the lenis form, as schematized in (11b). Had the fortis form been retained, we would predict, based on regular sound changes, that it would be *-bi*.

The reason for the loss of the fortis/lenis alternation is unclear at this point, but we find its loss, to greater of lesser degree, scattered across both ET and WT languages. In ET languages, it was always the lenis forms that were retained whenever the loss of the alternation was unconditioned, suggesting that the lenis forms were the underlying forms, and the fortis forms a conditioned alternant (cf. Chacon 2014a).

(11) DEFAULT (Kubeo)

a.	σ-pɨ FORTIS	→	σ-bɨ	→	-wɨ
b.	-bɨ LENIS	→	-bɨ	→	-wɨ

The changes in the 3.MSC category in Kubeo are more complex. The ultimate outcome was that both the PET fortis alternant *-piʔĩ and lenis alternant *-biʔĩ reduced and merged to *-bi*. Focusing first on everything but the initial suffix, one possible route for these changes is the following: first, the glottal stop was lost, followed by denasalization of the final vowel, and reduction of the two identical vowels to a single one: *iʔĩ > iĩ > ii > i*. It is also certainly possible that denasalization preceded the loss of glottal stop, but this does not materially affect our account, and at this point we do not have evidence that would allow us to distinguish between the two hypotheses.

Turning to the initial consonant, we find that the contrast between fortis and lenis forms was lost, in favor of the form *-bi*. Again, there are multiple routes by which this may have happened, but the fact the ultimate form is *–bi* (the expected fortis reflex), rather than *–wi* (the expected lenis reflex), entails that either: 1) this is the lenis form (with leveling towards the lenis form), and that the **b > w* sound change, invoked

above, was for some reason blocked (see below); 2) that this is the fortis form (with leveling towards the fortis form), with the **p > b* sound change being due to a morphologically conditioned sound change (see below); or 3) this is a purely phonological merger, in which both the fortis and lenis forms independently came to begin with *b* for different reasons (e.g. via each of the two previously mentioned reasons).

Considering Hypothesis 1, it is possible that the leniting sound change was blocked in the lenis alternant for prosodic reasons. Recall that after the loss of the glottal stop and denasalization of the vowel, the lenis form was *bii,* which would be parsed by a bimoraic foot. At the same time, recall that fortis forms only occurred in an even mora suffix groups, with the fortis form appearing rightmost, suggesting a system of right-aligned iambic bimoraic feet. Returning to the lenis form, then, under this analysis, *bii* would constitute a single foot, and would be stressed, with stress blocking the lenition process. On this view, the alternation subsequently leveled to this lenis form, as schematized in (12).

(12) 3.MSC (Kubeo)

a. σ-piʔĩ $_{(\text{FORTIS})}$ → σ-piĩ → σ-pii → σ-bii (voicing) → -bi

b. **-**biʔĩ $_{(\text{LENIS})}$ → **-**biĩ → -bii (lenition blocked) → -bii → -bi

Hypothesis 2, in contrast, accounts for the single *-bi* 3.MSC form by observing that the fact that the initial segment of the fortis form is *b* rather than *p* reflects a broader morphologically conditioned sound change whereby morpheme initial voiced stops underwent voicing (see Chacon [2012] for the systematicity of this change). On this view, the lenis alternant may indeed have undergone the leniting sound change to *-wi*, but evidence of this sound change was subsequently erased by leveling of the lenis form towards the fortis form, as schematized in (13). It should be noted, however, that this would be the only instance of leveling towards the fortis form we need to posit, and that in all other cases where the fortis-lenis distinction was lost, it is the lenis form that survives, leading us to deprecate Hypothesis 2.

(13) 3.MSC (Kubeo)

a. σ-piʔĩ $_{(\text{FORTIS})}$ → σ-piĩ → σ-pii → σ-bii (voicing) → -bii → -bi

b. **-**biʔĩ $_{(\text{LENIS})}$ → **-**biĩ → -wii → -wii → -bii (leveling) → -bi

Hypothesis 3, as sketched out above, combines the sound-change mechanisms of Hypotheses 1 and 2, and does not posit leveling. In particular, under this hypothesis, the fortis form undergoes initial voicing, as per (12a), and the lenition of the lenis form is blocked, as per (12b), in both cases resulting in the surface form *-bi*. This hypothesis has the merit of not requiring us to invoke leveling.

6.1.2 Nuclear ET

Sound changes also affected the forms of the verbal agreement suffixes in Nuclear ET. As in Kubeo, the glottal stop was lost, but in Nuclear ET languages, the nasalization of the final vowel spread to the initial consonant. This may have been due to the fact that nasalization spread across the intervening glottal stop prior to its loss, as depicted in (14), or because nasality was retained on the final vowel after the loss of the glottal stop and then, either: 1) nasality spread leftwards, or 2) it was the nasal vowel that survived vowel hiatus resolution.

In either case, vowel hiatus from the loss of the glottal stop was resolved, and then the leniting sound changes took place: **b[m] > w* and **p > b[m]*. As for the default category, the leniting sound changes took place only in the lenis form, but not in the fortis form, as depicted in (15). The reason for a leniting sound change in the 3.MSC fortis form, but not in the default fortis form, is probably related to the presence of a nasal vowel in the former, but not the later, case. Nasalization is a trigger of leniting sound changes in the history of ET languages (cf. Chacon 2014).

(14) 3.MSC

a. σ-piʔĩ $_{(\text{FORTIS})}$ → σ-pĩʔĩ → σ-pĩ → σ-mĩ

b. -biʔĩ $_{(\text{LENIS})}$ → -mĩʔĩ → -mĩ → -wĩ

(15) DEFAULT

a. σ-pɨ $_{(\text{FORTIS})}$ → σ-pɨ

b. -bɨ $_{(\text{LENIS})}$ → -wɨ

ET languages vary in the degree to which they uniformly followed the trajectories described above. Tukano and Karapana are the only languages in which the changes resulted exactly in the forms in the end of the processes schematized in (14) and (15). In Tuyuka, Yuruti and Tatuyo, for example, the lenis/fortis alternation was lost, and we only have the lenis forms (3.MSC: *-wĩ*, DEFAULT: *-wɨ*). That fortis-lenis alternation was neutralized in favor of the lenis forms is consistent with the hypothesis, proposed in the

discussion of the Kubeo facts in §6.1.1, that the lenis forms were the underlying forms, whereas the fortis ones were the conditioned alternant. An additional change further differentiates Tuyuka and Yuruti from the rest of the ET languages, namely the denasalization of all 3rd person animate morphemes, including the feminine and the plural forms.

Another major group of languages that departs from the general processes schematized in (14) and (15) consists of Desano, Makuna and Barasano, which constitute a clade within ET (cf. Section 2). In these languages, there was no nasalization-condition leniting sound change affected the lenis forms, so that the lenis form remained *-mĩ* in the counterpart to the (14a) process. Consequently, when the counterpart to the fortis form (14b) process was completed, which also resulted in *-mĩ*, the phonological distinction between fortis and lenis in 3.MSC was lost, preserving only *mĩ*. In the default category, there was leveling between the fortis and lenis forms in favor of the lenis *-bɨ* in Makuna and Desano; only Barasano kept the fortis and lenis alternation, although subsequent *p > h in this language (also found Makuna) obscures this. The leveling towards the lenis form in the default category and the fortis form in the 3.MSC category followed the same pattern as in Kubeo, paralleling other similarities among these languages in the historical phonology of ET (Chacon 2014a; Chacon and List 2016).

6.2 3.FEM and 3.AN.PL

As discussed above, the morphological heterogeneity of the 3.FEM and 3.AN.PL categories in Tukanoan suggest that they developed subsequent to the diversification of PT into its daughter languages. In this section, we discuss the development of these categories in ET, beginning with a review of the 3.MSC agreement construction and the possible source constructions from which 3.FEM and 3.AN.PL developed, prior to the diversification of ET, given in (16). The 3.MSC and default agreement constructions at this stage is given in (16a&d), respectively, while the possible source constructions for the 3.FEM and 3.AN.PL categories, prior to their morphologization, are given in (16b&c), respectively (see Table 11 for the source of post-verbal pronouns/demonstratives).

			PET
(16)	a.	3.MSC	V-*biʔĩ*
	b.	3.FEM	V-b̶i̶ i-ko*/ĩ-o*
	c.	3.AN.PL	V-b̶i̶ *ĩ-o-ã/di-ã*
	d.	1/2/3.INAN	V-b̶i̶

Although it is difficult to be certain about the relative chronology of the morphologization of the 3.FEM and 3.AN.PL agreement categories, the evidence suggests that the 3.AN.PL morphologized first, in PET, and the 3.FEM category morphologized subsequent to the diversification of ET. In particular, we find that the 3.AN.PL category is expressed in a morphologically uniform fashion in ET (and in its conflation with the default category, in WT), suggesting that this category was innovated in PET, but not in WT. The 3.FEM category, on the other hand, is expressed more heterogeneously in ET, suggesting that it may have been a parallel development in two or more ET sub-branches. For expositional purposes we will assume that 3.AN.PL agreement morphologized first in PET after 3.MSC, followed by the parallel morphologization of 3.FEM in the early post-PET daughter languages.[17]

6.2.1 3.AN.PL

The development of 3.AN.PL verbal agreement plausibly arose from either a process of fusion of the appropriate post-verbal pronoun/demonstrative to the perfective/past suffix, as we have described for 3.MSC, or a process of analogical change.

The fusional process would have involved the basis of the pronoun **ĩ-o-ã,* encliticizing to the verb and then fusing with the verbal TAM suffix (note that it could not plausibly have involved the form **di-ã,* since there is no evidence of the initial consonant in the 3.AN.PL suffix), as schematized in (17). In an intermediate stage, the resulting TAM + PNG morpheme was bimoraic, similarly to (14a), and, thus, subject to the same subsequent phonological processes as illustrated below.

17 It is worth noting that WT languages exhibit a distinct, but relatively uniform, way of expressing 3.FEM agreement *-go* or *-ko(ʔi)*, suggesting the independent but parallel innovation of this agreement category in PWT.

(17)		pre-PET		PET		Post-PET	
a.	3.MSC	-bi-ʔĩ	→	-bi-ʔĩ	→	-bi	Kubeo
b.	3.MSC	-bi-ʔĩ	→	-bi-ʔĩ	→	-mĩ	Nuclear-ET
c.	3.AN.PL	-bɨ ĩ-o-ã	→	-mĩ-ã	→	-mã	Kubeo & Nuclear-ET

Note that this explains why in Kubeo the initial segment of the 3.AN.PL agreement suffix, *–mã,* is nasal, but the initial segment of the 3.MSC agreement suffix, *-bi,* is oral: in the former, the initial consonant and the nasal vowel formed a single syllable at, say, the PET stage, which triggered nasalization in the whole syllable (a typical process of vowel harmony in the Tukanoan languages), whereas in the 3.MSC agreement suffix, the glottal stop served as an onset for the syllable containing the nasal vowel, placing it in a different syllable and preventing the nasalization of the voiced bilabial stop.

Alternatively, the 3.AN.PL agreement suffix may have arisen via analogy, where the 3.MSC PET form *-bi-ʔĩ* was taken as the base upon which speakers recognized *-ʔĩ* as 'masculine' and innovated with the form *-bi-ã (>mĩã > -mã)* with *–ã* as a marker of 'plural'.

It is possible that future work will allow us to adjudicate between these two hypotheses, but at this point it is not possible to rule out either one.

6.2.2 3.FEM

The final stage of the emergence of the agreement categories was the morphologization of the 3.FEM category. There are three cognate sets of 3.FEM agreement suffixes: 1) the singleton Kubeo *-biko* set; 2) the singleton Yurutí *-go* set; and 3) the remaining set containing the forms *-mõ, -wõ,* and *-wo*. As in the case of the 3.AN.PL agreement suffix, it is not entirely clear at this form whether these 3.FEM agreement suffixes resulted from the fusion of post-verbal pro-forms with the past/perfective suffix, or by analogy.

If the 3.FEM agreement suffix arose via a fusional process of the type already sketched above, we can assume a stage in early Post-PET, where the situation schematized in (18) held.

(18)			
	a.	3.MSC	-bi-ʔĩ
	b.	3.FEM	-bɨ i-ko/ĩ-o

In Kubeo the 3.FEM agreement suffix *–biko* appears to have been formed by adding *-ko,* a feminine classifier, nominalizer, and subordinate clause suffix found in the

majority of Tukanoan languages. On this account, *–biko* arose by fusion of the post-verbal demonstrative **i-ko* with the past/perfective suffix *-bɨ*, as illustrated in (19).

(19)		PET		KUBEO
a.	3.MSC	-bi-ʔĩ	→	-bi
b.	3.FEM	-bɨ i-ko	→	-bi-ko

Alternatively, the form *–biko* may have arisen analogically, in which case the *-ko* feminine suffix was directly concatenated to the morpheme *-bi* 3.MSC to express the feminine agreement feature.

Likewise, the Yuruti *-go* 3.FEM, which is cognate with Kubeo *-ko* and exhibits similar grammatical functions, arose via analogy or is the result of the fusional process between the past/perfective suffix and the element *i-go,* found as a demonstrative in the closely related language Tuyuka in Table 10.

The formation of morphemes *-mõ, -wõ* and *-wo* in the Nuclear-ET languages could also have resulted from fusion, but with the pronoun base **ĩ-o which explains the nasal reflexes in these languages, as illustrated in (20).

(20)		PET		Proto-Nuclear-ET		
a.	3.MSC	-bi-ʔĩ	→	-mĩ	→	-mĩ
b.	3.FEM	-bɨ ĩ-o	→	-mĩ-õ	→	-mõ

Again, these forms may alternatively have arisen via analogy, where *-o* 'feminine' was inserted directly into the paradigm, simply substituting for the final vowels of morphemes *-mĩ 3.MSC *-mã 3.AN.PL.

7. Conclusion

In this paper we proposed an account for the emergence of subject agreement in the Tukanoan family, and a description of the development of subject agreement paradigms in the Eastern branch of the family. We showed that the evolution of subject-verb agreement was triggered by the fusion of masculine post-verbal pronouns/demonstratives with a past/perfective TAM suffix, resulting in morphemes that cumulatively exponed PNG categories with tense. And significantly, we showed how agreement systems with such cross-linguistically unusual patterns of syncretism arose.

We sought to be as specific as possible in the description of the diachronic processes involved in the emergence of the agreement categories, including accounting for the phonetic details in the morphological processes and sound changes involved. Nevertheless, it is likely that some of the reconstructions we have proposed will be further refined in the future, especially in the domain of demonstratives and pronouns. Most of the phonological reconstructions relied on Chacon (2014a), which drew its data principally from free morphemes, but sound changes in bound morphemes, especially high frequent ones, are not expected to be identical to those in free morphemes. This, combined with the lack of a reconstruction of PT vowels and their subsequent evolution means that certain aspects of the reconstructions presented here remain tentative.

The processes of morphological change described in this paper are generally consistent with the internal classification given in section 2. However, some changes needed to be explained as parallel changes, and others perhaps being due to contact. For example, languages like Kotiria and Waikhana seem to have entirely lost subject-agreement forms quite recently, given their position in the Tukanoan family tree. Likewise, Koreguahe, Tanimuka and Retuarã seem to have done the same, but perhaps at an older stage, given their position in the tree.

The evolution of 3.FEM and 3.AN.PL forms seem to have occurred after the change that morphologized the exponance of 3.MSC. The details of how the former two forms evolved, whether by analogy or a similar fusion process that caused the emergence of 3.MSC is still unclear. There are a couple points in the analysis of a fusion hypothesis that must be addressed in future work, such as why fusion took place first in 3.MSC and not at the same time in all categories, and why some of the pronominal forms required to explain this process are not all found in ET languages, although they can be reconstructed to PT.

One of the next steps to be taken in developing a fuller account of the development of agreement morphology in the Tukanoan family is to extend our results here to other TAM paradigms. Significantly, we believe that the past/perfective paradigm whose development we describe in this paper served as the template for other TAM agreement paradigms, with the PNG categories distinguished in the past/perfective paradigm also serving as the categories in these other paradigms.

Another obvious area for future work is the extension of the account developed here to more fully explain the development of agreement paradigms in WT languages, which we have touched on only slightly in this paper. Siona and Sekoya exhibit the most ET-like agreement paradigms, while the Koreguahe system has been been radically restructured so that it is a pure gender and number agreement system. Máíhĩki perhaps

presents the greatest analytical challenge, since it exhibits some cognate morphemes with ET languages, but appears to have also experienced partial leveling with subordinate clause agreement paradigms, as well as what appears to be sound changes in the vowel system that further obscure cognacy judgments.

8. Acknowledgements

The authors would like to thank Rosa Vallejos and Anne Schwarz for sharing data on Sekoya and Siona-Sekoya , and for their insightful comments on the constructions discussed in this paper.

9. References cited

Bruil, Martine. 2014. *Clause-typing and evidentiality in Ecuadorian Siona*. Utrecht: LOT.

Chacon, Thiago Costa. 2015. The Reconstruction of Laryngealization in Proto-Tukanoan. *The Phonetics and Phonology of Laryngeal Features in Native American Languages* ed. by Heriberto Avelino, Matt Coler, and W. Leo Wetzels, 258-285. Leiden: Brill.

Chacon, Thiago Costa. 2014a. A revised proposal of Proto-Tukanoan consonants and Tukanoan family classification. *International Journal of American Linguistics* 80:3.275–322.

Chacon, Thiago Costa. 2014b. Indexação Verbo-Sujeito nas línguas Tukáno. *Revista Brasileira de Linguística Antropológica* 6:1.239-268.

Chacon, Thiago Costa. 2012. *The phonology and morphology of Kubeo: The documentation, theory and description of an Amazonian language*. Ph.D. Dissertation, University of Hawai'i.

Chacon, Thiago Costa and J.M. List. 2016. Improved computational models of sound change shed light on the history of the Tukanoan languages. *Journal of Language Relationships* 13.177-203.

Cook, Dorothy and Linda Criswell. 1993. *El idioma koreguaje (tucano occidental)*. Bogotá: Asociación Instituto Lingüístico de Verano.

Eraso, Natalia. 2014. *Grammaire de la langue Tanimuka, Amazonie Colombienne*. Phd Dissertation: Université Lumière Lyon 2.

Farmer, Stephanie. 2012. The effects of low tone tonogenesis on the prosodic

system of Máíjɨqui. Talk presented at the SSILA 2012 Annual Meeting.

Farmer, Stephanie. 2015. *Establishing reference in Máíhɨ̃ki.* PhD dissertation: UC Berkeley.

Gomez-Imbert, Elsa. 2007. Nominal classification in Tukanoan languages. *Language Endangerment and Endangered Languages : Linguistic and Anthropological Studies with Special Emphasis on the Languages and Cultures of the Andean-Amazonian Border Area* ed. by W. Leo Wetzels, xxx. Indigenous Languages of Latin America series (ILLA). Publications of the Research School of Asian, African and Amerindian Studies (CNWS). The Netherlands: Leiden University.

Gomez-Imbert, Elsa, and Stephen Hugh-Jones. 2000. Introducción al estudio de las lenguas del Piraparaná (Vaupes). *Lenguas indígenas de Colombia: Una visión descriptiva* ed. Maria Stella González de Pérez and Maria Luisa Rodríguez de Montes, 321–356. Bogotá: Instituto Caro y Cuervo.

Guillot, Céline. 2015. From Old French and Middle French to Contemporary French. *Manual of Deixis in Romance Languages* ed. by Konstanze Jungbluth and Federica Da Milano, 546-568. Mouton de Gruyter.

Johnson, Orville, and Stephen Levinsohn. 1990. Gramática secoya. *Cuadernos Etnolingüísticos*, v.11. Quito: Instituto Linguístico de Verano.

Jones, Wendell and Paula Jones. 1991. *Barasano Syntax*. (Studies in the Languages of Colombia, 2.) Dallas: SIL International and University of Texas at Arlington.

Kinch, Rodney and Pamela Kinch. 2000. El Yurutí. *Lenguas indígenas de Colombia: Una visión descriptiva*, ed. by Maria Stella González de Pérez and Maria Luisa Rodríguez de Montes, 321–56. Bogotá: Instituto Caro y Cuervo.

Malone, Terrell. 1988. The origin and development of Tuyuca evidentials. *International Journal of American Linguistics* 54:2.119-140.

Metzger, Ronald G. 1981. *Gramática Popular del Carapana*. Bogotá, Colombia: Instituto Lingüístico del Verano.

Michael, Lev. 2012. *Máíjɨ̱kì finite verbal inflectional paradigms.* UC Berkeley Ms.

Miller, Marion. 1999. *Desano Grammar*. Dallas: Summer Institute of Linguistics and the University of Texas at Arlington

Piaguaje, Ramón, Elías Piaguaje, Orville E. Johnson, and Mary Johnson, compilers. 1992. *Vocabulario Secoya*. Quito: Instituto Lingüístico de Verano.

Ramirez, Henri. 2001. xxx ???

Ramirez, Henri. 1997. *A fala Tukano dos Ye'pâ-mãsa: Vol. 1: Gramática.* Manaus: Inspetoria Salesiana Missionaria da Amazônia (CEDEM).

Smothermon, Josephine, Paul Frank, and Jeffrey Smothermon. 1995. *Bosquejo del Macuna: Aspectos de la cultura material de los macunas, fonología, gramática.* Bogotá: Association [sic] Instituto Lingüístico de Verano.

Stenzel, Kristine. 2015. Considerações sobre ordem de palavras, tópico e o 'efeito foco' em Kotiria. Revel 10.223-246.

Stenzel, Kristine. 2013. *A Reference Grammar of Kotiria (Wanano)*. Lincoln: University of Nebraska Press.

Stenzel, Kristine and Elsa Gomez-Imbert. 2018. Evidentiality in Tukanoan languages. *The Oxford Handbook of Evidentiality*, ed. by Alexandra Aikhenvald, 357-387. Oxford: Oxford University Press.

Vallejos, Rosa. 2013. El secoya del Putumayo: aportes fonológicos para la reconstrucción del Proto-Tucano Occidental. *Línguas Indígenas Americanas* (LIAMES) 13:1.67-100.

Waltz, Nathan, and Alva Wheeler. 1972. Proto Tucanoan. *Comparative Studies in Amerindian Languages*, ed. by Esther Matteson, 19-49. Berlin: Mouton de Gruyter.

Wheeler, Alva. 1987. Gantëya Bain (*El pueblo siona del río Putumayo, Colombia). Vol. 1: Etnología, gramática, textos.* Bogotá: Instituto Linguistico de Verano.

Wheeler, Alva. 1992. Comparaciones lingüisticas en el grupo Tucano Occidental. *Estudios Comparativos Proto Tucano*, ed. by Stephen H. Levinsohn, 17–53. Bogotá: Alberto Lleras Camargo.