



Boletim do Museu Paraense Emílio Goeldi. Ciências Humanas

ISSN: 1981-8122

ISSN: 2178-2547

MCTI/Museu Paraense Emílio Goeldi

Nikulin, Andrey; Carvalho, Fernando  
A revised reconstruction of the Proto-Tupian vowel system  
Boletim do Museu Paraense Emílio Goeldi. Ciências Humanas, vol. 17, no. 2, e20210035, 2022  
MCTI/Museu Paraense Emílio Goeldi

DOI: <https://doi.org/10.1590/2178-2547-BGOELDI-2021-0035>

Available in: <https://www.redalyc.org/articulo.oa?id=394072583012>

- How to cite
- Complete issue
- More information about this article
- Journal's webpage in redalyc.org

UABM  redalyc.org

Scientific Information System Redalyc  
Network of Scientific Journals from Latin America and the Caribbean, Spain and Portugal

Project academic non-profit, developed under the open access initiative

## A revised reconstruction of the Proto-Tupian vowel system

### O sistema vocálico do Proto-Tupi: uma nova proposta reconstrutiva

Andrey Nikulin<sup>I</sup>  | Fernando Carvalho<sup>II</sup> 

<sup>I</sup>Núcleo Takinahakỹ de Formação Superior Indígena/Universidade Federal de Goiás. Goiânia, Goiás, Brasil

<sup>II</sup>Museu Nacional/Universidade Federal do Rio de Janeiro. Rio de Janeiro, Rio de Janeiro, Brasil

**Abstract:** This contribution is concerned with the reconstruction of the vowel qualities of Proto-Tupian, the ancestral language of the Tupian language family. The study is grounded in a bottom-up application of the comparative method and seeks to offer a more balanced reconstruction that avoids an overreliance on the Tupí-Guaraní branch. It is first shown that the height opposition traditionally reconstructed for the rounded vowel series (*\*o* vs. *\*u*) is best interpreted as an opposition between an unrounded vowel and a rounded one (*\*ə* vs. *\*o*). It is also argued that multiple instances of *\*e* in the traditional reconstruction should be rather attributed to *\*ə*. Finally, it is shown that two vowels (symbolized as *\*i* and *\*u*) must be reconstructed in lieu of the traditional *\*j*. The resulting proposal has consequences for the subgrouping of the Tupian family.

**Keywords:** Tupian languages. Comparative reconstruction. Vowels. Sound change.

**Resumo:** O foco do presente trabalho é a reconstrução do sistema vocálico do Proto-Tupi, a língua ancestral da família linguística Tupi. A investigação baseia-se em uma aplicação *bottom-up* do método comparativo e busca oferecer uma reconstrução mais equilibrada, evitando uma influência desproporcional do ramo Tupi-Guarani. Mostraremos primeiro que a diferença de altura tradicionalmente reconstruída para a série de vogais arredondadas (*\*o* vs. *\*u*) pode ser melhor compreendida como uma oposição entre uma vogal não arredondada e uma vogal arredondada (*\*ə* vs. *\*o*). Argumentamos também que múltiplas instâncias de *\*e* na reconstrução tradicional devem ser atribuídas a *\*ə*. Por fim, duas vogais (simbolizadas *\*i* e *\*u*) devem ser reconstruídas em lugar do segmento *\*j* da reconstrução tradicional. Estas propostas têm implicação para a classificação interna da família Tupi.

**Palavras-chave:** Línguas Tupi. Reconstrução comparativa. Vogais. Mudança sonora.

---

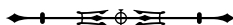
Nikulin, A., & Carvalho, F. (2022). A revised reconstruction of the Proto-Tupian vowel system. *Boletim do Museu Paraense Emílio Goeldi. Ciências Humanas*, 17(2), e20210035. doi: 10.1590/2178-2547-BGOELDI-2021-0035

Autor para correspondência: Fernando Carvalho. Museu Nacional/Universidade Federal do Rio de Janeiro. Quinta da Boa Vista, São Cristóvão, Rio de Janeiro, RJ, Brasil. CEP 20940-040 (fernaoorphao@gmail.com).

Recebido em 15/03/2021

Aprovado em 22/12/2021

Responsabilidade editorial: Ana Vilacy Galúcio



## INTRODUCTION

This article addresses the reconstruction of the vowels of Proto-Tupian (PT). Tupian is a major language family of South America, which comprises as many as ten universally recognized close-knit branches: Tupí-Guaraní, Awetí, Sateré-Mawé, Mundurukú, Juruna, Mondé, Tuparian, Arikém, Karo (= Ramarama), and Puruborá. Of these, four consist of only one language (Awetí, Sateré-Mawé, Karo, Puruborá), while Mundurukú, Juruna, and Arikém are similar in that each of them comprises one living language (Mundurukú, Yudjá, and Karitiana, respectively) and one extinct (or dormant), less well-attested language (Kuruaya, Xipaya, and Arikém, respectively); the Juruna branch also includes a third extinct language, Manitsawá, of which very little is known.

It is now universally accepted that Tupí-Guaraní, Awetí, and Sateré-Mawé are particularly closely related to each other, constituting thus a node within Tupian; furthermore, Tupí-Guaraní and Awetí share some innovations that did not affect Sateré-Mawé, suggesting an early split of the latter language off Proto-Mawé-Awetí-Tupí-Guaraní (Rodrigues, 1984-1955, p. 35; Rodrigues & Dietrich, 1997; Corrêa da Silva, 2010; Meira & Drude, 2015). In the remainder of this paper, we will label the node that comprises Tupí-Guaraní and Awetí ‘Awetí-Guaraní’, and the node that comprises Awetí-Guaraní and Sateré-Mawé will be accordingly called ‘Mawé-Guaraní’. Another proposal related to the internal classification of Tupian advances a hypothesis according to which Puruborá and Karo would form a subgroup (‘Ramarama–Puruborá’, see Rodrigues, 2007, p. 168, fn. 2; Galucio & Gabas Jr., 2002).

In this article, we critically assess the only existing reconstructive proposal that has ever been put forward for the PT segmental inventory (cf. Rodrigues, 1999, 2005, 2007), focusing on the vowel system as reconstructed in Rodrigues (2005). Rodrigues (2005) proposal is presented as established common knowledge in later reference works on the Tupian family (see e.g. Rodrigues & Cabral, 2012). However, we will argue that although the author’s cognacy judgments are, for the most part, precise, his analysis of the attested correspondence sets is problematic and, for this reason, in need of revision. Specific problems which we attempt to resolve in this contribution include the sound correspondence in the pan-Tupian word for ‘cultivated field’ (cf. Sateré-Mawé *ŋo/ko*, Awetí *ko*, Mundurukú *kə*, Yudjá *kúá*, but Makurap and Wayoró *ŋge*, Karitiana *nga*), which is considered irregular by Rodrigues (2005, p. 40), as well as the reason why the vowel *\*o* in Rodrigues’ (1999) reconstruction occurs predominantly after labial consonants (cf. Rodrigues, 1999, pp. 110–111).

After a brief presentation of the relevant aspects of Rodrigues (2005) PT reconstruction (“Earlier scholarship”), we discuss some vowel correspondences that are especially troublesome in Rodrigues’ (2005) framework, and put forward an alternative account of the evolution of the vowel system of PT (sections “PT *\*ə*” and “PT *\*u* vs. *\*i*”). The scenario we propose requires positing some non-trivial innovations shared by specific groups of Tupian, as discussed in “Implications for the subgrouping of Tupian”. After the “Conclusions”, the “List of abbreviations”, and the “References”, we include two appendices. Appendix 1 contains all the cognate sets discussed in this contribution, whereas Appendix 2 summarizes our reconstruction of Proto-Tupian consonants.

## NOTATION

Before proceeding to the bulk of the discussion, we make explicit our conventions regarding the representation of linguistic data. In order to warrant the comparability of the data, which are extracted from a variety of sources on multiple Tupian languages, we have unified the representation of all cited forms according to the following principles.

First of all, the flap *r* is represented as *r* throughout this paper, and the open-mid vowels *ɔ* and *ɛ* as *o* and *e*, respectively, as is in fact commonly done in Tupian linguistics. An exception is made for Kuruaya, where */ɔ/* and */o/*

contrast and are thus represented as **ɔ** and **o**. The vowels of the extinct language Arikém, as well as of its direct ancestor Proto-Arikém, are represented as **/i e æ ɒ ʊ/**. The choice of the less common characters **/æ ɒ ʊ/** is suggested by the variable representation of these vowels as **<a> ~ <e>**, **<o> ~ <ə> ~ <a>**, **<u> ~ <ʊ> ~ <i>** in our only sources on the language (Rondon & Faria, 1948; Nimuendajú, 1932).

In most Tupian languages, consonants in the coda position do not contrast for features other than the place of articulation. This is true for Sateré-Mawé, Awetí, Proto-Tupí-Guaraní (and most contemporary Tupí-Guaraní languages), Proto-Tuparian (and all contemporary Tuparian languages), Proto-Arikém, and Puruborá. We write such codas in small caps: **p** (labial), **ɾ** (dental/alveolar), **c** (palatal), **κ** (velar). Their precise surface realizations vary depending on the language, on the point of articulation, and on the phonological environment, and include **[p̚ m β]** for **p**, **[ɾ̚ n ɾ]** for **ɾ**, **[(i)ç̚ ɲ j]** for **c**, **[k̚ ŋ ɰ g]** for **κ**.

The phonetic inventory of many Tupian languages has two-phase stops that start out with a lower velum which raises during the occlusion. These are variably analyzed as postoralized allophones of underlying nasals or prenasalized allophones of underlying voiced stops (cf. Wetzels & Nevins, 2018), or (more rarely) as underlying prenasalized stops (cf. González, 2008). In this paper, such segments are always spelt as **mb**, **nd**, **ndʒ**, **ŋg**, regardless of their phonological status in a given language.

Nasal vowels are always explicitly marked with a tilde, even if our sources leave nasality unmarked in some environments (usually following or preceding nasal consonants).

In languages where **[s]** and **[ts]** are free, idiolectal, dialectal, or chronological variants of one phoneme, we have normalized the data from our sources in order to warrant consistency across examples (e.g., only **s** is used in Tuparí, and only **ts** in Sakurabiat), following the current transcription practices in the most recent expert source in each case.

Proto-Mawé-Awetí-Tupí-Guaraní (= Proto-Mawetí-Guaraní = PMG) reconstructions are mostly those by Meira and Drude (2015), with the following modifications. Meira and Drude (2015) **\*t̥** is rewritten as **\*c**, except when it follows an **\*i** or a **\*j** in their reconstruction, in which case we employ an *ad hoc* character **\*č** in order to capture the fact that the reflexes of **\*č** in all daughter languages are different from those of **\*c** (this also allows us to give up the reconstruction of ‘phantom’ instances of PMG **\*i** and **\*j**, which are in Meira and Drude (2015) proposal hypothesized to have been lost in all daughter languages). For example, Meira and Drude’s (2015) reconstructions such as **\*it̥et** ‘his/her name’ and **\*t̥ajt̥u** ‘armadillo’ are replaced with **\*čer** and **\*caču**, given that no daughter language preserves any segmental trace of the alleged PMG segments **\*i** and **\*j** (Sateré-Mawé *her*, *sahu*; Awetí *ter*, *tatu[peɾ]*; PTG *tset*, *tatu*). We also posit PMG **\*t̥** alongside **\*c** and **\*č** in order to account for the sound correspondence involving Sateré-Mawé **t̥** (in unstressed syllables) / **ɾ(j)** (in stressed syllables), Awetí **z̥**, and PTG **\*t̥**, found in the stem for ‘fire’ (reconstructed as **\*atia**, **\*atja** in Meira & Drude, 2015; **\*(c)at̥a** in our proposal) as well as before **\*i** (in complementary distribution with **\*t̥**)<sup>1</sup>. We also diverge from Meira and Drude (2015) in reconstructing **\*w** instead of their **\*kʷ**, whereas the segment they reconstruct as **\*w** is considered to have been independently epenthesised in the environment **\*u\_V** after the split of PMG (that way, Meira & Drude’s (2015) reconstructions such as **\*t̥uwaj** ‘tail’ and **\*t̥uwi(k)** ‘blood’ are replaced here

<sup>1</sup> The reconstruction of **\*t̥** as one segment in the word for ‘fire’ (as opposed to **\*ti ~ \*tj**) is supported by the fact that not only PTG **\*tata** / **\*-rata**, Awetí **taza** / **-aza**, but also Mundurukú **dafá**, Kuruya **láfa**, Yudjá **afí**, Xipaya **afí** have precisely one segment as its correspondence. Only Sateré-Mawé would have unfolded **\*t̥** to **ɾj** in *arja* ‘fire’.

with *\*cuac*, *\*cui* > Sateré-Mawé *suwac-po*, *su*; Awetí *-uwac*, *-uwi[k]*; PTG *\*tuwac/\*-ruwac*, *\*tuwi/\*-ruwi*)<sup>2</sup>. We accept Schleicher's (1998, pp. 18–24) suggestion, reinforced by Meira and Drude (2015, pp. 278–279), whereby only one PTG affricate is reconstructed instead of the traditional *\*c* (*\*ts*) and *\*č* (*\*tʃ*); we symbolize it as *\*ts*<sup>3</sup>.

Proto-Mundurukú reconstructions are taken from Picanço (2019). For relational stems, whose leftmost consonant often alternates depending on the left context, Picanço (2019) lists all possible allomorphs. In this article, only the allomorphs with *\*p-*, *\*ð-*, and *\*tʃ-* (rather than *\*b-*, *\*t-*, *\*j-*) are given for such stems. We also omit the hyphen, used by Picanço (2019) in order to indicate that the relational stems are bound.

For Proto-Juruna and Proto-Tuparian, a variety of proposals exist; the reconstructions in this article are mostly extracted from the more recent ones (Carvalho, 2019 for Proto-Juruna; Nikulin & Andrade, 2020 for Proto-Tuparian) or adapted from earlier proposals (Fargetti & Rodrigues, 2008 for Proto-Juruna; Moore & Galucio, 1994; Nogueira et al., 2019 for Proto-Tuparian) so as to match the phonological reconstruction of the most recent works.

The reconstruction of the Proto-Tupian consonants adopted here differs from previous proposals and is based on the correspondence sets summarized in Appendix 2. For reasons of space, it is impossible to discuss this problem in detail in this contribution. Note, however, that nothing in our reconstruction of the Proto-Tupian vowels hinges on our interpretation of the PT consonants, and the validity of our proposal would remain intact if one adopted a different interpretation (such as that of Rodrigues, 2007).

The acute accent symbolizes the high tone in tonal languages (including the languages of the Mundurukú, Juruna, and Mondé branches, Karo, and maybe also Makurap); the low tone is left unmarked. Tones other than high and low are found only in the Mondé languages, and the transcription of our sources is retained in such cases. In the Tuparian languages Tuparí and Akuntsú, contrastive stress has been described, which is also symbolized by means of an acute accent (when its position is known).

Data quoted from premodern sources, which are not expected to faithfully represent all the relevant phonological oppositions, are given 'verbatim' enclosed in chevrons. Subscript letters after such forms indicate the ultimate source of the data:  $\diamond_s$  and  $\diamond_L$  refer to Emilie Snethlage's and Lopes' data on Kuruaya (Snethlage, 1932);  $\diamond_B$  and  $\diamond_N$  refer to Barbosa's and Nimuendajú's data on Arikém (Rondon & Faria, 1948; Nimuendajú, 1932). Forms followed by  $\diamond_{KG}$  come from Koch-Grünberg (1932) on Puruborá, and those with  $\diamond_{ES}$  from Snethlage (1934), again on Puruborá. Forms in chevrons without subscript letters are from Steinen (1886) for Manitsawá, Nimuendajú (1923–1924, 1928, 1929) for Xipayá, and Sekelj (1948) for Aruá and Makurap.

Much of the discussion in this paper is based on analyzing cognate sets. In some cases, a given form is not synchronically segmentable, but only a part of it is cognate with the material of other languages. The part which is deemed non-cognate is then given in brackets. In premodern attestations (enclosed in chevrons), the cognate part is given in boldface.

<sup>2</sup> The epenthetic nature of the *w* in these stems is likewise confirmed by the fact that no corresponding consonant is found in branches such as Mundurukú (*\*ðoj* 'blood', *\*t-oaj-bi* 'tail'; see Picanço, 2019) or Tuparian (*\*jei* 'blood', *\*joac* 'tail'; see Nikulin & Andrade, 2020). Our amendment to Meira and Drude's (2015) reconstruction spares us from the necessity of positing a typologically improbable 'zigzag' development in Sateré-Mawé, whereby PT *\*w* > PMG *\*kʷ* > Sateré-Mawé *w* (*h* before *u* in stressed syllables;  $\emptyset$  before *u* in unstressed syllables). In our account, Sateré-Mawé *w* (*h*/ $\emptyset$  before *u*) simply continues PT *\*w* > PMG *\*w*. Only the ancestral language of Awetí and PTG would thus have innovated by transforming PT *\*w* > PMG *\*w* into a stop.

<sup>3</sup> The diverging reflexes in the Guaraní varieties that were earlier seen as warranting the reconstruction of two affricates *\*c* and *\*č* for PTG are now explained as late developments involving diffusion or dialect borrowing among Guaraní dialects.

## EARLIER SCHOLARSHIP

The vocalic system of Proto-Tupian has been reconstructed by Rodrigues and Dietrich (1997, p. 268) and Rodrigues (1999, p. 110, 2005) as comprising six vowel qualities (\*a, \*i, \*o, \*u, \*e, \*ɨ), each of which would also have a nasal counterpart (\*ã, \*ĩ, \*õ, \*ũ, \*ẽ, \*ĩ̃; see Rodrigues & Cabral, 2012, p. 502). In Table 1, we list the reflexes of the Proto-Tupian oral vowels according to the proposal by Rodrigues (2005).

Table 1. The reflexes of PT oral vowels (after Rodrigues, 2005, p. 37).

PT	PTG	Awetí	S.-Mawé	Yudjá	Mundurukú	Karitiana	Tuparí	Suruí-Paiter	Karo	Puruborá
*a	*a	a	a	a	a	o	a	a	a	a
*i	*i	ɨ	ɨ	ɨ, i	i	i, e	ɨ, ʉ	i	ɨ, i	ɨ, i
*o	*o	o	o	a	ə	ɨ	o	a	a	a
*u	*u	u	u	u	o	ɨ	o	o	o	o
*e	*e, *o	e, o	e, o	e, a	e, ə	a	e	e	e	e
*ɨ	*ɨ	i	i	i	i	i	i	i	i	i

It can be easily seen from the table above that, according to Rodrigues (2005), the evolution of Proto-Tupian \*e and \*ɨ in the constituent families involved some phoneme splits. In the case of PT \*e, Rodrigues (2005) posits a split in the so called ‘Eastern’ Tupian languages (Mawé-Awetí-Tupí-Guaraní, Mundurukú, and Juruna) allegedly conditioned by an adjacent consonant. The examples 1-2 illustrate the default development of PT \*e, whereas 3 instantiates the development of PT \*e affected, according to Rodrigues’ (2005, p. 40) proposal, by a following labialized stop (the labialization is reconstructed here exclusively in order to account for what are thought to be the divergent reflexes of PT \*e)<sup>4</sup>.

- (1) PT \*k<sup>h</sup>et ‘to sleep’ (Rodrigues, 2005, p. 40) > PTG \*k<sup>h</sup>er, Aw ket, Mw ket || Mu fet || Tu ?et || Kt kat || Kr ket || Pu ket-a || Pa ker
- (2) PT \*men ‘husband’ (Rodrigues & Cabral, 2012, p. 507) > PTG \*men, Aw men || Yu méná, Xi mén-a || Ma/Wy/Sk/Ak men || Kt mǎn, Ari man || Kr mén || Pu men || Ar men, Gv mǎt [mǎnt], Zo met
- (3) PT \*ep<sup>w</sup> ‘leaf’ (Rodrigues, 2005, p. 40; Rodrigues & Cabral, 2012, p. 505) > PTG \*(-)oβ, Aw (-)op, Mw -op || Mu -əp || Yu -up-á ~ úp-a, Xi s-up-á || Tu/Ak -ep || Kt s-ap || Pu tap || Gv s-ep, Cl s-ép

For \*ɨ, Rodrigues (2005, pp. 40–41) posits a split in Juruna and in the so called ‘Western’ Tupian languages of the Arikém, Tuparian, Karo (Ramarama), and Puruborá groups. This time, however, he does not identify a phonological environment which could have conditioned the alleged split (beyond a generic reference to the ‘immediate consonantal context’), nor is he explicit about whether the alleged split proceeded in the same way in all the aforementioned languages. The examples 4-6 illustrate.

- (4) PT \*k<sup>h</sup>ip ‘tree, wood’ (Rodrigues, 2005, p. 41; Rodrigues & Cabral, 2012, p. 506) > PTG \*ʔiβ, Aw/Mw ʔip || Mu/Ku ʔip || Yu ʔip-á || PTpr \*k<sup>h</sup>ip > Ma kip, Tu kup, Sk/Ak/Kp kip || Kt ʔep || Kr ma-ʔip || Pu mamka-ʔip ‘castanha tree’ || Pa [ʔ]í:b, Gv [ʔ]iip
- (5) PT \*pi ‘foot’ (Rodrigues, 2005, p. 41; Rodrigues & Cabral, 2012, pp. 503, 504) > PTG \*pi, Aw/Mw pi || Mu/Ku i || Ma mi, Tu si-to, Sk/Ak pi || Kt pi || Kr pi, pi:-be || Pu ši-be || Pa pí:-pe, Gv pi

<sup>4</sup> When citing comparanda after Rodrigues (2005) and Rodrigues and Cabral (2012), we leave unchanged their transcription conventions, morphological segmentations, cognation judgments, and reconstructions. Our analysis, which may differ significantly, is available in the Appendix 1. We have been unable to confirm the existence of some of the forms given by these authors (such as Makurap men ‘husband’).

- (6) PT *\*počij* ~ *\*pocij* 'heavy' (Rodrigues, 2005, p. 41; Rodrigues & Cabral, 2012, p. 504) > PTG *\*pocij*, Aw/Mw *potij* || Mu *poši*  
 || Yu *i-pade-tú* ~ *pade-*, Xi *pade-* || Ma/Wa *poti*, Tu *poci* ~ *posi* || Kt *piti* || Kr *piʔtí* ~ *piʔti* || Pa *pati-ga*,  
 Ar <patí ~ pati-í> 'heavy, thick', Gv *patí* ~ *patî*, Zo *pati*, Cl *pattí*

In the subsequent sections, we will argue against the proposal by Rodrigues (2005), suggesting instead that the observed sound correspondences are best accounted for by reconstructing a phonemic inventory of seven (rather than six) vowel qualities for Proto-Tupian and positing a number of mergers in the daughter languages, in addition to one conditioned split. That way, the examples in 1-6 are reconstructed in our proposal as *\*k'et* 'to sleep', *\*mẽt* 'husband', *\*jəp* 'leaf', *\*kʷp* 'tree; stick-like', *\*mbi/\*pi* 'foot', *\*pətic* 'heavy'. Tables 2 and 3 show the oral vowel inventories of Proto-Tupian in Rodrigues' (2005) and our proposals, respectively.

Table 2. PT vowel inventory in Rodrigues' (2005) proposal.

<i>*i</i>	<i>*ĩ</i>	<i>*u</i>
<i>*e</i>		<i>*o</i>
	<i>*a</i>	

Table 3. PT vowel inventory in our proposal.

<i>*i</i>	<i>*ĩ</i>	<i>*ʷ</i>
<i>*e</i>	<i>*ə</i>	<i>*o</i>
	<i>*a</i>	

## PT *\*ə*

This section deals with the reconstruction of a vowel we chose to represent as *\*ə*. We start by stating its proposed reflexes in the daughter branches and listing the relevant cognate sets. In subsequent sections, we discuss how our findings relate to Rodrigues' (2005) reconstruction of the PT vowels and consonants. We conclude that the recognition of *\*ə* as a contrastive unit allows reducing the phonological inventory of Proto-Tupian by three phonemes (*\*pʷ*, *\*kʷ*, *\*kʷʷ*), to account for the sound correspondences in a number of cognate sets which are unexplainable in Rodrigues' (2005) proposal, and to account for the limited distribution of the sound correspondence which underlies Rodrigues' (2005) reconstruction of PT *\*o* (which occurs exclusively following labial consonants). Our proposal also entails that the vowels traditionally reconstructed as *\*o* and *\*u* should be reinterpreted as PT *\*ə*, *\*o*.

## PROPOSAL

The vowel we reconstruct as PT *\*ə* has evolved in the following way in the daughter languages<sup>5</sup>. In the Mawé-Guaraní branch, it has acquired rounding and changed to PMG *\*o* (in fact, in our proposal PT *\*ə* is the only source of PMG *\*o*). In addition, it has been raised to *\*u* before a vowel (as in 'blood', 'sun') or before a glottal stop and a vowel (as in 'arrow'). In the Mundurukú branch, PT *\*ə* is regularly reflected as Mu *ə*, Ku *ĩ* (< PMu *\*ĩ* in Picanço's (2019) reconstruction) and

<sup>5</sup> Note that in this article we take nasality to be an autosegmental feature in Proto-Tupian and most Tupian languages. It rarely interacts with the development of the vocalic segments (we note it explicitly when it does).

is the only source of this PMu vowel. Before a pre-PMu vowel, however, PT \*ə is reflected as PMu \*o ('blood', 'arrow', 'sun', provided that PMu \*đoj and \*op/\*đop go back to pre-PMu \*đoi, \*oip/\*đoip). The sequence \*mã is reflected as Mu mã, Ku mã (< PMu \*mã in Picanço's (2019) reconstruction). In the Juruna branch, one finds PJu \*a except next to a labial consonant in the final syllable of the PT stem ('leaf', 'snake', 'hand', 'wing'; exception: 'what') or before a vowel ('sun'), where the regular reflex is \*u. In the Tuparian branch, PT \*ə is fronted to PTpr \*e except after labial consonants, in which case the regular reflex is PTpr \*o ('to return', 'hand', 'wing/feather', 'heavy', 'causative'). Likewise, in the Arikém branch PT \*ə is reflected as a low vowel (Kt a, Ari æ) except after labial consonants, in which case one finds Kt i, Ari u ('hand', 'vine', 'wing/feather', 'heavy'). Puruborá retains PT \*ə as ə (or ǣ, if nasality is present). In the Mondé languages, PT \*ə is reflected as a except after coronal consonants and word-initially, in which case e is found ('leaf', 'larva', 'house/village'). Only in Karo do the reflexes seem chaotic at present: one finds i in two examples ('cylindrical and small', 'to hold'), i in two examples ('heavy', 'to go', in both cases next to a \*t), o in two examples ('leaf', 'third person coreferential'), a in two prefixes ('causative', 'sociative causative'), as well as the following vowels in one example each: ə ('to go, to walk'), ǣ ('snake'), and u ('blood', a likely result of contraction of the unique vowel sequence \*əu).

- (7) PT \*apə 'interrogative word' > PMG \*apo (only Mw apo 'yes/no question particle') || PMu \*abi 'who' || (?) PJu \*apá 'what' || PTpr \*apo 'who; what (subject)'
- (8) PT \*đəp 'bitter' > PMG \*đop || PMu \*tjip || PTpr \*tep || Kt ta:p
- (9) PT \*etə- 'sociative causative' > PMG \*ero- || (?) PMu \*[d]iji- (only Mu dədʒə) || PTpr \*ete- || Kt at-ot 'to take away' || Kr ta-
- (10) PT \*ək/\*jək 'house' > PMG \*ok/\*cok || PMu \*ik-ʔa/\*đik-ʔa || PJu \*ak-á || PTpr \*ek/\*jek || Ari æk-ə 'house' || Pu ʔək-a || Ar <eg> 'house, village'  
derived: PT \*ək-at 'plaza' > PMG \*ok-at (only PTG \*ok-at) || Kt ak-ot 'together', Ari ækəð-ə 'plaza'  
derived: PT \*ək-ēt/\*jək-ēt 'plaza'/'door' > PMG \*ok-ēt 'door' || PMu \*ik-ēn/\*đik-ēn (only Ku iʃēn 'door') || PTpr \*ek-ēt (only Tu ek-ēt 'plaza') || Kt ak-ān 'village'
- (11) PT \*əkʷip/\*jəkʷip 'arrow' > PMG \*uʔip/\*cuʔip || PMu \*op/\*đop || PTpr \*ekip/\*jekip || (?) Pa jáb, Ar <ndjap>, Gv dʒáp/áɖʒáp ~ íɖʒáp, Zo dʒap
- (12) PT \*jakək 'army ant' > PMG \*caʔok (only PTG \*taʔok) || PMu \*đaʔik (only Mu daʔək) || PTpr \*jakek
- (13) PT \*jəβa 'forehead' > PMG \*coβa 'face' || PTpr \*jeβa
- (14) PT \*jəp 'leaf' > PMG \*-cop || PMu \*đip || PJu \*súp-á || PTpr \*jep || Kt sap, Ari səβ-ə || Kr naʔ-jop || Pu təp 'leaf, hair' || Ar/Gv tsep, Zo sep, Sl bá-sep
- (15) PT \*jəu 'blood' > PMG \*cui || PMu \*đoj || PTpr \*jei || Ari nǣe || Kr ju
- (16) PT \*kə 'to go, to walk' > (?) PMG \*[e]ko 'to be, to live' || PMu \*[ki]ki || Kr kə || Pa ka:, Gv kaà, Zo ka
- (17) PT \*kəc 'to plant' > PMG \*koc || PJu \*kat-ú 'to bury, to plant' || Gv kaja
- (18) PT \*kək 'to hold' > PMG \*kok (only PTG \*kok 'to support') || PMu \*kik (only Mu kək) || Kt kak || Kr kik
- (19) PT \*mbə (absolute)\*pə (relational) 'hand, finger, vine-like' > PMG \*mbo/\*po || PMu \*pi || PJu \*bu-á || PTpr \*mbo 'hand' || Kt pi, Ari pə 'hand' || Kr piʔ 'cylindrical and small' || Pu bə 'long, vine-shaped' || Pa/Ar/Gv/Zo/Sl pá-be  
derived: PT \*utu-pə 'vine' > PMG \*itipo || PMu \*tji-bi || PTpr \*itimbo (only Ma <ötömbó>) || Kt tepí, Ari tepu  
derived: PT \*ja-pə 'root' > PMG \*capo || PMu \*tabi(?) (fossilized 3<sup>rd</sup> person)
- (20) PT \*mbəc 'snake' > PMG \*mboc || PMu \*pij || PJu \*put-á || (?) Kr mǎj[gǎra] || Pu mǎñ[ũp]⁶ || Pa mǎj[kir], maj[kó:ra:] (spp.), Gv/Zo baj
- (21) PT \*mǎ- 'causative' > PMG \*mō- || PMu \*ma- || PJu \*mǎ- || PTpr \*mō- || Kt m- (irregular vowel loss), Ari mǎ- || Kr ma- || Gv mǎ-
- (22) PT \*mǎk 'to stick, to glue' > PMG \*mōk || PMu \*mǎñ 'to lean'
- (23) PT \*ngə 'cultivated field' > PMG \*ngo || PMu \*ki 'women's field' || PJu \*kú-á || PTpr \*nge || Kt nga, Ar ngæ || Pu <á><sub>KG</sub> (if Koch-Grünberg's <á> stands for ə) || Pa ŋa, Ar <ngá>, Gv/Zo ga
- (24) PT \*ngəat 'sun' > PMG \*nguāt || (?) PMu \*koat[o] 'summer' || PJu \*kuaɓ[adɪ] || PTpr \*ngear 'sun, sky' || Pa nǎd, Ar/Gv/Zo/Sl gát
- (25) PT \*paβək 'to be visible' > PMu \*-pápik (only Mu papák) || PTpr \*[e]paβok- (only Tu epapók- 'to return')
- (26) PT \*pepʔə 'wing, flight feather' > PMG \*pepo || PJu \*pebu-a || PTpr \*perʔo || Kt papí 'arrow feather, fin', Ari pæpə 'wing'

<sup>6</sup> Galucio (2005) gives the phonetic form as [mǎ'ñũm] and analyzes [ǎ] as a realization of /ǎ/. In Galucio et al. (2015, p. 271), the form mǎjũp(m) is given instead. We assume the representation /ǎ/ for the vowel in question.



- (27) PT *\*pətic* 'heavy' > PMG *\*potic* || PMu *\*pítʃi* || PJu *\*padét-ú* || PTpr *\*poti* || Kt *piti* || Kr *piʔti* || Pa *pati-ga*, Ar *patií ~ pati-í* 'heavy, thick', Gv *patî*
- (28) PT *\*tə* 'to go (*singular*)' > PMG *\*to* || PMu *\*tʃi* || PJu *\*tʃa* || Kr *ti*
- (29) PT *\*-tə* 'pronominal formative (pl.)' > PMG *\*-to* || PMu *\*-ji* || PTpr *\*-te* || Kt *-tʃa /-ta/*, Ari *-tæ* || Kr *-tə*
- (30) PT *\*tə-* 'third person coreferential' > PMG *\*to-* (only Mw *to-*) || (?) Xi *d(u)-*<sup>7</sup> || PTpr *\*te-* || Kt *ta-*, Ari *tæ-* || Kr *to-*
- (31) PT *\*tək* 'to pound, to grind' > PMG *\*tok* || PMu *\*tʃik* (only Mu *tʃək* 'to break') || PJu *\*(pá-)dák-u* || Kt *tak*
- (32) PT *\*tək* 'larva' > PMG *\*tok* (only PTG *\*tsok*) || PJu *\*[a]dák-á* || PTpr *\*tek* (only Tu *tek*) || Kt *[ken]tak* || Pa *[ka]dé:g* 'coconut larva', (?) *[mã:m]nég* 'caterpillar sp.', (?) *nég* 'boil', Ar *mandeg* 'caterpillar', Gv *mã-dék* 'lizard'
- (33) PT *\*ʔək* 'tuber' > PMG *\*ʔok* || PMu *\*ʔik* 'belly' || PTpr *\*ek* (only Tu *ek*)  
derived: PT *\*mānī-ʔək* 'manioc' > PMG *\*mānīʔok* || PMu *\*másik* || PJu *\*maniak-á*

## ADVANTAGES WITH RESPECT TO RODRIGUES (2005)

In what follows, we discuss four correspondence sets derived by Rodrigues (2005) from three Proto-Tupian vowels. Two of them, identified by Rodrigues (2005) with PT *\*e* and *\*u*, show no overlap at all; these correspondence sets appear as (a) and (d), respectively, in Table 4 below. The remaining two correspondences, given as (b) and (c) in Table 4, are attributed in our account to PT *\*ə*, yet an entirely different account is proposed by Rodrigues (2005). The correspondence (c) has the same reflexes as (d) in Karitiana and Tuparí, but other languages show distinct reflexes, which are typically lower than those of (d); it is associated by Rodrigues (2005) with PT *\*o*. The correspondence (b) shows significant overlaps with (a) and (c): in Tupí–Guaraní, Awetí, Sateré-Mawé, Mundurukú, and Yudjá, the observed reflexes are identical to those of (c), whereas in all other branches the correspondence set in question — in Rodrigues' (2005) account — coincides completely with (a).

Table 4. Four correspondence sets for Tupian vowels (adapted from Rodrigues, 2005, p. 37).

	PTG	Awetí	S.-Mawé	Yudjá	Mundurukú	Karitiana	Tuparí	S.-Paiter	Karo	Puruborá
(a)	<i>*e</i>	e	e	e	e	a	e	e	e	e
(b)	<i>*o</i>	o	o	a	ə	a	e	e	e	e
(c)	<i>*o</i>	o	o	a	ə	ɨ	o	a	a	a
(d)	<i>*u</i>	u	u	u	o	ɨ	o	o	o	o

According to Rodrigues (2005, pp. 40, 42, 2007, pp. 175–176, 181–182, 186), the overlapping pattern that involves the correspondence (b) — which coincides with (c) in Tupí–Guaraní, Awetí, Sateré-Mawé, Mundurukú, and Yudjá, but with (a) in the remaining branches — can be explained by reconstructing a secondary rounding feature for the consonant that immediately follows the vowel (the available options in Rodrigues's, 2005 reconstruction are *\*pʷ*, *\*kʷ*, *\*kʷʔ*). In the proto-language of Mawé–Guaraní, Mundurukú, and Yudjá ('Eastern' Tupian languages in Rodrigues', 2005 terms), this contextual factor would have induced the merger of the correspondence set in (b) with the *\*o* series, Yudjá being later subject to *\*o* > *a* and Mundurukú undergoing *\*o* > *ə*. The remaining branches

<sup>7</sup> Word-initially, PT *\*t* is expected to yield PJu *\*tʃ* > Xi *t*, not *d*. Also the regular reflex of PT *\*ə* is *u* only before vowels, but not before consonants. Together, these facts cast doubt on the inclusion of the Xipaya prefix in this cognate set. Note, however, that at least the apparently irregular vowel reflex could be attributed to leveling: at an earlier stage, *\*da-* may have occurred before consonants and *\*du-* before vowels. The allomorph *du-* is still used in some pre-vocalic contexts in Xipáya (*du-ázi* 'his (own) wife'), but shows up as *d-* before unaccented vowels (*d-aká* 'his/her (own) house'; Rodrigues, 1995, p. 12).

— that is, Arikém, Tuparian, Mondé, Ramarama, and Puruborá — would have not been subject to any contextual coloring and show reflexes identical to those of PT *\*e*, which leads Rodrigues (2005, 2007) to reconstruct *\*e* for the correspondence set in question and to posit a conditioned split in his ‘Eastern’ languages. In synthesis, then, Rodrigues’ (2005) proposal for the PT segments whose reflexes appear in the correspondences in Table 4 above is as follows: *\*e* for the correspondences (a) and (b), *\*o* for (c), and *\*u* for (d). Moreover, the context-dependent merger of the series (b) and (c) in Mawé-Guaraní, Mundurukú, and Juruna is attributed to the influence of a secondary rounding feature hosted on the following consonant.

Under closer scrutiny, however, it appears that the available evidence does not support either the identification of the correspondence (b) as a context-dependent offshoot of (a) or the reconstruction of a labialized stop series for Proto-Tupian. Although we concur with Rodrigues (2005) in reconstructing PT *\*e* for the correspondence (a), we disagree with his diachronic interpretation of the remaining three correspondences in that:

- we consider (b) to be the default development of PT *\*ə* (rather than a positional development of *\*e*);
- we consider (c) to be a positional development of PT *\*ə* (rather than the default reflex of a PT phoneme of its own, symbolized as *\*o* by Rodrigues (2005));
- we derive (d) from PT *\*o*, which in our account is the only rounded vowel of PT (as opposed to Rodrigues’ (2005) reconstruction, whereby PT had both *\*u* and *\*o*).

Rodrigues’ (2005) account is seriously undermined by the following facts.

1. First of all, the consonants reconstructed as labialized by Rodrigues (2005) in etyma that instantiate the correspondence set (b) appear not to have reflexes distinct from those of non-labialized consonants.
2. Furthermore, the correspondence in (b) may also occur morpheme-finally or morpheme-internally before vowels, making it impossible to attribute the emergence of the correspondence to a following consonantal segment.
3. The correspondence set in (c) may be explained away as a conditioned offshoot of (b).
4. The reflexes listed by Rodrigues (2005) for (b) and (c) in Suruí-Paiter, Karo, and Puruborá are partially based on non-cognate material and are thus incorrect.
5. Finally, there is typological evidence that renders Rodrigues’s (2005) hypothesis implausible.

Each of these five points is discussed in the subsequent sections.

## PURPORTED LABIALIZED CONSONANTS HAVE THE SAME REFLEXES AS PLAIN CONSONANTS

Let us consider the reflexes of the PT segments that Rodrigues (2007) reconstructs as labialized consonants. As will become clear, their reflexes do not differ from those of their plain (non-labialized) counterparts, and the only reason for positing such phonemes in Rodrigues (2007) reconstruction is to account for the correspondence set (b). Once it is recognized that the (b) series does not result from a conditioned split of *\*e*, it is no longer necessary to reconstruct labialized consonants for Proto-Tupian.

We will start by examining the occurrences of *\*pʷ* that are supposed to account for the alleged rounding of PT *\*e* in ‘Eastern’ Tupian. Rodrigues (2007) reconstructs it for two roots, *\*epʷ* ‘leaf’ and *\*epʷa* ‘face’ (as well as in its derivative *\*epʷa-pokʷ* ‘to appear’). In the former case, all Tupian branches have a reflex with a plain labial stop *p* (in some languages, which lack an opposition between oral and nasal codas, it is symbolized as *p̥*). In most Tupí-Guaraní languages, as well as in the Arikém language before the suffix *-a*, the stop is lenited to *β* or a similar

sound (cf. Schleicher, 1998, pp. 29–32; Storto & Baldi, 1994); this development regularly targets word-final stops of any origin in these languages<sup>8</sup>. In our reconstruction, the vowel correspondence between PMG *\*o*, PMu *\*i*, PTpr *\*e*, Kt *a*, Ari *æ*, Pu *ə*, and Mo *e* is derived from PT *\*ə*, whereas the correspondence between the word-final consonants straightforwardly continues PT *\*-p*. That way, there is no need to reconstruct a labialized stop for ‘leaf’ in our proposal.

- (34) PT *\*jəp* ‘leaf’ (Rodrigues (2007): *\*ep<sup>w</sup>*) > PMG *\*-cop* || PMu *\*ðip* || PJu *\*súp-á* || PTpr *\*jep* || Kt *sap*, Ari *-sæβ-o* || Kr *naʔ-jop* || Pu *təp* ‘leaf, hair’ || Ar/Gv *tsep*, Zo *sep*, Sl *bá-sep*

An identical rhyme is found in the word for ‘bitter’. Rodrigues (2007, p. 196) reconstructs its PT etymon as *\*riop* and lists its reflexes in Sateré-Mawé, Awetí, PTG, and Mundurukú. Had he considered the Tuparí and Karitiana cognates, he would have likely reconstructed *\*riep<sup>w</sup>*.

- (35) PT *\*ðəp* ‘bitter’ (Rodrigues, 2007: *\*riop*) > PMG *\*ðop* || PMu *\*tʃip* || PTpr *\*tep* || Kt *ta:p*

In Rodrigues’ (2007) account, *\*p<sup>w</sup>* is claimed to have a divergent reflex between vowels in the languages of the Mawé-Guaraní (PT *\*p<sup>w</sup>* > PMG *\*β*, as opposed to PT *\*p* > PMG *\*p*) and Mundurukú (PT *\*p<sup>w</sup>* > PMG *\*p*, as opposed to PT *\*p* > PMG *\*b*) branches. Rodrigues (2007) gives only two cognate sets that instantiate the sequence *\*ep<sup>w</sup>*: PT *\*ep<sup>w</sup>a* ‘face’ and *\*ep<sup>w</sup>apok<sup>?</sup>* ‘to appear’; Corrêa da Silva (2010, p. 128) claims the latter to be a derivative of the former. Rodrigues (2007, p. 186) lists the following reconstructions and reflexes (quoted *verbatim*).

- (36) PT *\*ep<sup>w</sup>a* ‘face’ (Rodrigues, 2007 reconstruction) > Mw *-ewa*, Aw *-owa*, PTG *\*-oβa* || Mu *dopa*, Ku *t-upa* || Ma *č-epa* ‘forehead’, Tu *epa* ‘eye, light’, *epa-psi* ‘face’, Sk/Ak *ebapi* ‘face’, *eba-opap* ‘eye’ || Kt *s-ipo* ‘eye’  
 (37) PT *\*ep<sup>w</sup>apok<sup>?</sup>* ‘to appear’ (Rodrigues, 2007 reconstruction) > PTG *\*oβapo* || Mu *j-ebapək* || Tu *epapok* ‘to arrive’

Note, however, that even within Rodrigues’ own framework the proposed etymology for ‘face’ presents serious irregularities. In his account, PT *\*e* before a labialized consonant would be expected to yield Mw *o* (rather than *e*), Mu *ə* (rather than *o*), Ku *i* (rather than *o* — note that *u* in the datum cited by Rodrigues (2007) is a phonetic variant of /*o*/), and Kt *a* (rather than *i*). We surmise that in this case Rodrigues (2007) has failed to distinguish between two unrelated cognate sets, which we derive from PT *\*jəβa* ‘forehead’ and *\*jopʔa* ‘face’ (in addition, Mw *-ewa*, or *sewa* in our notation, appears to be unrelated to either etymon).

- (38) PT *\*jəβa* ‘forehead’ (Rodrigues, 2007: *\*ep<sup>w</sup>a* ‘face’) > PMG *\*coβa* || PTpr *\*jeβa*  
 (39) PT *\*jopʔa* ‘face’ (Rodrigues, 2007: *\*ep<sup>w</sup>a*) > PMu *\*ðópa* || Kt *sí:po* ‘eye’, *sipo* ‘seed’

PT *\*β* in *\*jəβa* ‘forehead’ is reconstructed based on the correspondence PMG *\*β* ~ PTpr *\*β*, otherwise found in the cognate set for ‘wind’, PMG *\*iβiçu* ~ PTpr *\*iβijo* (cf. Nikulin & Andrade, 2020, p. 292). In *\*jopʔa*, the consonant cluster is reconstructed in order to account for the voiceless intervocalic stop *\*p* in PMu, also found in etymologies

<sup>8</sup> At the PTG level, at least, only *\*-p* and *\*-ɾ* were subject to lenition, but the dorsal stop *\*-k* was not (Schleicher, 1998, p. 32; Meira & Drude, 2015, p. 281). A few languages like Kayabí and members of the Kagwahiva cluster extended the process to reflexes of PTG *\*-k* as well.

such as PMu *\*óropo* ~ PTpr *\*oropʔo* ~ PMG *\*uruβu* 'vulture'<sup>9</sup>. In our account, the correspondence between PMG *\*o* and PTpr *\*e* need not be conditioned by any feature hosted on the following consonant.

As for the cognate set for 'to appear', we believe that PTG *\*oβapo* should be excluded from it (no other examples are known where a velar coda in Mundurukú or Tuparian would correspond to zero in PTG)<sup>10</sup>. Moreover, the Mundurukú cognate does not actually contain an initial vowel (the root is *papák* 'to be visible', with the allomorph *bapák* occurring after vowels; cf, Picanço, 2005, p. 17). That way, this cognate set does not instantiate the vowel correspondence which interests us in this section, nor does it back up the reconstruction of *\*pʷ*.

- (40) PT *\*paβæk* 'to be visible' > PMu *\*-párik* (only Mu *papák*) || PTpr *\*[e]paβok-* (only Tu *epapók-* 'to return')

Now we turn to the purported PT *\*kʷ*. According to Rodrigues (2007), this phoneme induces contextual coloring in the preceding vowel in the following cognate sets.

- (41) PT *\*ək/\*jək* 'house' (Rodrigues, 2007: *\*ekʷ*) > PMG *\*ok/\*cok* || PMu *\*ík-ʔa/\*ðík-ʔa* || PJu *\*ak-á* || PTpr *\*ek/\*jek* || Ari *æk-o* 'house' || Pu *ʔək-a* || Ar <eg> 'house, village'  
derived: PT *\*ək-at* 'plaza' (Rodrigues, 2007: *\*ekʷat*) > PMG *\*ok-at* (only PTG *\*ok-at*) || Kt *ək-ot* 'together', Ari *ækod-o* 'plaza'  
derived: PT *\*ək-ēt/\*jək-ēt* 'plaza'/'door' (Rodrigues, 2007: *\*ekʷen*) > PMG *\*ok-ēt* 'door' || PMu *\*ik-ēn/\*ðik-ēn* (only Ku *ifēn*, 3<sup>rd</sup> *t-ifēn* 'door') || PTpr *\*ek-ēt* (only Tu *ek-ēt* 'plaza') || Kt *ək-ān* 'village'
- (42) PT *\*tək* 'to pound, to grind' (Rodrigues, 2007: *\*čekʷ*) > PMG *\*tok* || PMu *\*tʃik* (only Mu *tʃək* 'to break') || PJu *\*(pá)-dák-u* || Kt *tak*
- (43) PT *\*tək* 'larva' (Rodrigues, 2007: *\*čekʷ*) > PMG *\*tok* (only PTG *\*tsok*) || PJu *\*[a]dák-á* || PTpr *\*tek* (only Tu *tek*) || Kt *[ken]tak* || Pa *[ka]dé:g* 'coconut larva', (?) *[mā:m]nég* 'caterpillar sp.', (?) *nég* 'boil', Ar <mandeg> 'caterpillar', Gv *mā-dék* 'lizard'
- (44) PT *\*ʔək* 'tuber' (Rodrigues, 2007: *\*ʔekʷ*) > PMG *\*ʔok* || PMu *\*ʔik* 'belly' > Mu *ʔək*, Ku *ʔik* || PTpr *\*ek* (only Tu *ek*)  
derived: PT *\*mānī-ʔək* 'manioc' (Rodrigues, 2007: *\*mani* + *\*ʔekʷ*) > PMG *\*mānīʔok* || PMu *\*māsik* || PJu *\*maniak-á*

To the best of our knowledge, no PT reconstructions for the following cognate sets are proposed in the works by Rodrigues (2007). He would probably reconstruct *\*tʔakʔekʷ* 'army ant' and *\*kekʷ* 'to hold'.

- (45) PT *\*jaʔək* 'army ant' > PMG *\*caʔok* (only PTG *\*taʔok*) || PMu *\*ðaʔik* (only Mu *daʔək*) || PTpr *\*jakek*
- (46) PT *\*kək* 'to hold' > PMG *\*kok* (only PTG *\*kok* 'to support') || PMu *\*kik* (only Mu *kək*) || Kt *kak* || Kr *kik*

As can be seen in these examples, the putative PT *\*kʷ* does not have labialized reflexes in any single Tupian branch. In fact, in all cases it is reflected precisely in the same way as PT *\*k* (represented as *\*k* in the coda position).

<sup>9</sup> An anonymous reviewer notes that positing PMu *\*p* as a regular reflex of PT *\*pʔ* could be problematic if one accepts that PMu *\*eba* 'wing' (Picanço, 2019, p. 136) is related to PT *\*pepʔə* 'wing'. Although we find the comparison in question tempting, the mismatch between PMu *\*a* and PT *\*ə* precludes us from considering PMu *\*eba* a reflex of PT *\*pepʔə*. Instead, it is possible that the Mundurukú form continues a different derivative of a hypothetical PT root *\*pep-*, whose erstwhile existence is suggested by Tuparí *pépʔe* 'fin' and *pépʔa* 'butterfly' (Alves, 2004, p. 236) alongside *pépʔo* 'feather' (< PT *\*pepʔə*).

<sup>10</sup> Of relevance to this issue, note that the lack of an explicit bottom-up lexical reconstruction of PTG is a significant drawback of the Tupian comparative literature. Although much of the work by Rodrigues often uses Old Tupi (or some other conservative/well-attested language) as a kind of proxy for PTG, we were unable to find either PTG *\*oβapo* 'to appear' (Rodrigues, 2007, p. 186), or any presumed reflex of this etymon, after searching the PTG vocabulary of Mello (2000) and the extensive lexical documentation of conservative languages such as Old Guaraní (Restivo, 1893 [1722]) and Old Tupi (Drummond, 1952–1953). These are reasons strong enough to believe that Rodrigues' (2007) PTG *\*oβapo* is a lexical ghost.

The only example which, according to Rodrigues (2007, p. 182; 2008, p. 6), instantiates a labialized reflex of PT *\*kʷ* in a Tupian language is PT *\*ekʷ-at* ‘plaza’ > Xi *koað-á*, Sk *ekʷat*, Tu *ekoat-pe* ‘area around the house’<sup>11</sup>. We reconstruct the etymon in question as PT *\*ək-ər* and reject the appurtenance of the cited words to this cognate set. First of all, the ultimate sources of the Xipaya and Sakurabiat words provide glosses which are quite distant from ‘(village) plaza’: <ɰʷazá> ‘village of foreigners’ (Nimuendajú, 1928, p. 827), <hekʷat> ‘field’ (Hanke et al., 1958, p. 205). Second, Sakurabiat *kʷ* does not regularly correspond to Tuparí *ko* (cf. Nikulin & Andrade, 2020).

The third purported labialized stop of Proto-Tupian is reconstructed by Rodrigues (2007) as *\*kʷʔ* for one single etymon.

- (47) PT *\*əkʷʔ/\*jəkʷʔ* ‘arrow’ (Rodrigues, 2007: *\*ekʷʔip*) > PMG *\*uʔip/\*cuʔip* || PMu *\*op/\*ðop* || PTpr *\*ekip/\*jekip* || (?) Pa *jáb*, Ar <ndjap>, Gv *dʒáp/ádʒáp* ~ *ídʒáp*, Zo *dʒap*

Rodrigues (2007, p. 186) states explicitly that *\*kʷʔ* is reflected in the daughter languages precisely in the same way as *\*kʔ* (we prefer to symbolize the segment in question with the *ad hoc* character *\*k̠*), and the labialization in the etymology for ‘arrow’ is reconstructed only in order to account for the correspondence between a front vowel in the Tuparian languages and rounded vowels in the ‘Eastern’ Tupian languages. Also note that the PT stem for ‘arrow’ almost certainly contains the formative for tree- or stick-like objects *\*-kʷʔip* in Rodrigues’, 2007 reconstruction), with reflexes in all Tupian languages, none of which shows any trace of labialization. In this sense, our proposal is superior to Rodrigues’ (2007) in that no need arises to reconstruct an extra consonant found in only one stem.

## THE CORRESPONDENCE SET (B) BEFORE VOWELS OR PAUSE

In this subsection, we examine the instances of the correspondence set (b) in environments in which no consonant follows the vowel in question. These receive no explanation in the framework of Rodrigues (2007): his proposal attributes the emergence of the series (b) to the rounding conditioned by a following labialized consonant, whereas in the etyma under consideration there is no consonant which could have triggered the alleged conditional rounding. The relevant data are provided below.

- (48) PT *\*etə-* ‘sociative causative’ > PMG *\*ero-* || (?) PMu *\*[ð]iji-* (only Mu *dəðʒə-*) || PTpr *\*ete-* || Kt *at-ot* ‘to take away’ || Kr *ta-*  
 (49) PT *\*ŋgə* ‘cultivated field’ > PMG *\*ŋgo* || PMu *\*kɪ* ‘women’s field’ || PJu *\*kú-á* || PTpr *\*ŋge* || Kt *ŋga*, Ar *ŋgæ* || Pu <tá><sub>KG</sub> || Pa *ŋa*, Ar <ngá>, Gv/Zo *ga*  
 (50) PT *\*ŋgəər* ‘sun’ > PMG *\*ŋguar* || (?) PMu *\*koət[o]* ‘summer’ || PJu *\*kuaɓ[adɪ]* || PTpr *\*ŋgear* ‘sun, sky’ || Pa *ŋád*, Ar/Gv/Zo/SI *gát*  
 (51) PT *\*jəw* ‘blood’ > PMG *\*cui* || PMu *\*ðoj* || PTpr *\*jei* || Ari *ŋæe* || Kr *ju*  
 (52) PT *\*-tə* ‘pronominal formative (pl.)’ > PMG *\*-to* || PMu *\*-ji* || PTpr *\*-te* || Kt *-tʃa /-tal*, Ari *-tæ* || Kr *-tə*  
 (53) PT *\*tə-* ‘third person coreferential’ > PMG *\*to-* (only Mw *to-*) || (?) Xi *d(u)-* || PTpr *\*te-* || Kt *ta-*, Ari *tæ-* || Kr *to-*

Regarding the etymology for ‘field’, Rodrigues (2005, p. 40) explicitly notes that the vowel of the Karitiana reflex (Kt *ŋga /ŋa/*) corresponds to that of the Makurap (Tuparian) reflex (Ma *ŋge /ŋe/*), but is not the regular outcome of PT *\*o*,

<sup>11</sup> Rodrigues (2007, p. 182) and Rodrigues and Cabral (2012, p. 508) also reconstruct word-initial instances of PT *\*kʷ* in some cognate sets, which are allegedly preserved in the Awetí–Guaraní languages as Aw *kw*, PTG *\*kʷ*. Meira and Drude (2015) show that the structure *(\*)kʷV* in these languages continues earlier sequences of the type *\*koV*, which may or may not result from elision of an erstwhile intervocalic consonant. Some examples are PMG *\*kocap* ‘to pass’ > Aw *kwap*, PTG *\*kʷap* (Meira & Drude, 2015, p. 294); PT *\*ŋgəər* ‘sun’ > Aw *kwat*, PTG *\*kʷat* (this paper). It is, therefore, doubtful that PT had *\*kʷ* word-initially.

reconstructed by Rodrigues (2005) based on the reflexes in other branches of Tupian<sup>12</sup>. Rodrigues (2005, pp. 39-40) also notes that his proposal fails to account for the Yudjá reflex (Yu **kú-á**), as PT **\*o** in his proposal is supposed to be reflected as **a** in the Juruna branch; he tentatively suggests that the Yudjá word is a loan from Língua Geral Amazônica, a Tupí-Guaraní language. In our account, none of these problems arises, as we hypothesize that PT **\*ə** regularly yields Kt **a**, Ma **e**, and (before vowels) Yu **u**.

## ROUNDING OF PT **\*ə** IN TUPARÍ AND ARIKÉM

Rodrigues (1999, pp. 110-111) notes that PT **\*o** — in his own reconstruction — occurs almost exclusively after labial consonants, and tentatively suggests that this vowel arose as a positional variant of PT **\*u**. We argue that the correspondence set which Rodrigues (1999, 2005) equates with PT **\*o** is indeed secondary: it may be indeed explained away as an offshoot of another PT vowel in the environment  $C_{[+labial]}(?)$ . However, our account differs from Rodrigues' (1999) suggestion in that the correspondence set in question is derived from PT **\*ə** (rather than from another rounded vowel) and in that the proposed conditioned split is attributed to a shallower level than Proto-Tupian: its effect is clearly visible in only two Tupian branches, Tuparian and Arikém. Instead of yielding PTpr **\*e**, Kt **a**, Ari **æ**, one finds PTpr **\*o**, Kt **i**, Ari **u** in the aforementioned context. The following examples illustrate.

- (54) PT **\*apə** 'interrogative word' > PTpr **\*apo** 'who; what (subject)'
- (55) PT **\*papək** 'to be visible' > PTpr **\*[e]papok** (only Tu **epapók** 'to return')
- (56) PT **\*mbə** (abs.) **\*pə** (rel.) 'hand, finger, vine-like' > PTpr **\*mbo** || Kt **pí**, Ari **pɯ**  
derived: PT **\*utu-pə** 'vine' > PTpr **\*itimbo** (only Ma **ötömbó**) || Kt **tepí**, Ari **tepu**
- (57) PT **\*mā-** 'causative' > PTpr **\*mō-** || Ari **mɯ-** (but Kt **m-**)
- (58) PT **\*pətic** 'heavy' > PTpr **\*poti** || Kt **piti**
- (59) PT **\*pepə** 'wing, flight feather' > PTpr **\*pepo** || Kt **papí** 'arrow feather, fin', Ari **pæpu** 'wing'

After non-labial consonants or word-initially, the reflexes are PTpr **\*e**, Kt **a**, Ari **æ**.

- (60) PT **\*ək/\*jək** 'house' > PTpr **\*ek/\*jek** || Ari **æk-o** 'house'  
derived: PT **\*ək-at** 'plaza' > Kt **ak-ot** 'together', Ari **ækvō-o** 'plaza'  
derived: PT **\*ək-ēt/\*jək-ēt** 'plaza'/'door' > PTpr **\*ek-ēt** (only Tu **ek-ēt** 'plaza') || Kt **akān** 'village'
- (61) PT **\*əkɯp/\*jəkɯp** 'arrow' > PTpr **\*ekip/\*jekip**
- (62) PT **\*jakək** 'army ant' > PTpr **\*jakek**
- (63) PT **\*jəβa** 'forehead' > PTpr **\*jeβa**
- (64) PT **\*jəp** 'leaf' > PTpr **\*jep** || Kt **sap**, Ari **-sæβ-o**
- (65) PT **\*jəw** 'blood' > PTpr **\*jei** || Ari **ɲæe**
- (66) PT **\*kək** 'to hold' > Kt **kak**
- (67) PT **\*dəp** 'bitter' > PTpr **\*tep** || Kt **ta:p**
- (68) PT **\*ŋgə** 'cultivated field' > PTpr **\*ŋge** || Kt **ŋga**, Ar **ŋgæ**
- (69) PT **\*ŋgəat** 'sun' > PTpr **\*ŋgeat** 'sun, sky'
- (70) PT **\*paβək** 'to be visible' > PTpr **\*[e]paβok** (only Tu **epapók** 'to return')
- (71) PT **\*tək** 'to pound, to grind' > Kt **tak**
- (72) PT **\*tək** 'larva' > PTpr **\*tek** (only Tu **tek**) || Kt **[ken]tak**
- (73) PT **\*ʔək** 'tuber' > PTpr **\*ek** (only Tu **ek**)

<sup>12</sup> cf. the original: ". . . AR [= Arikém branch] Karitiána **ɲa** (a vogal desta forma corresponde à do Makuráp, mas não à do étimo **\*\*ɲko**, TP [= Tuparian branch] Kepkiriwat **go**, mas Makuráp **ke** (e não é o reflexo regular de **\*\*o** nesta língua) . . ." (Rodrigues, 2005, p. 40). Note that Rodrigues (2005) takes Kepkiriwat to be a Tuparian language, in which we do not concur.

Not a single exception has been identified.

## REFLEXES OF (B) AND (C) IN KARO, PURUBORÁ, AND MONDÉ

According to Rodrigues (2005), the correspondence sets (b) and (c) have different reflexes not only in Tuparí and Karitiana (as we have shown in the previous section, they are in fact in a complementary distribution in these languages), but also in Karo, Puruborá, and Mondé (represented by Suruí-Paiter in Rodrigues', 2005 study). The (b) series (derived from PT \**e* before a labialized consonant in Rodrigues', 2005 proposal) is supposed to be reflected as *e* in Karo, Puruborá, and Suruí-Paiter, whereas the correspondence set (c) (< PT \**o* according to Rodrigues, 2005) is expected to yield *a* in the three languages.

In reality, however, the Karo and Puruborá reflexes listed by Rodrigues (2005, pp. 39-40) for the correspondence sets (b) and (c) turn out to be nonexistent, and the Mondé reflexes may be equally accounted for if one accepts our reconstruction of PT \**ə*. Consider the following etymologies (R = Rodrigues, 2007):

- (74) PT \**etə*- 'sociative causative' (R \**erie*-/ \**erio*-) > Kr *ta*-
- (75) PT \**ək*/\**jək* 'house' (R \**ek*<sup>w</sup>) > Pu *ʔək-a* || Ar <*eg*> 'house, village'
- (76) PT \**jəp* 'leaf' (R \**ep*<sup>w</sup>) > Kr *naʔ-jop* || Pu *təp* 'leaf, hair' || Ar/Gv *tsep*, Zo *sep*, Sl *básep*
- (77) PT \**kə* 'to go, to walk' (R \**eko*) > Kr *kə* || Pa *ka*·, Gv *kaà*, Zo *ka*
- (78) PT \**kək* 'to hold' (would-be reconstruction in Rodrigues' terms: \**kek*<sup>w</sup>) > Kr *kik*
- (79) PT \**mbə* (absolute)/\**pə* (relational) 'hand, finger, vine-like' (R \**po*/\**mpo*) > Kr *pɪʔ* 'cylindrical and small' || Pu *bə* 'long, vine-shaped' || Pa/Ar/Gv/Zo/Sl *pá-be*
- (80) PT \**mbəc* 'snake' (R \**mpoj*) > (?) Kr *mǎj[gǎra]* || Pu *mǎŋ[ũp]* || Pa *máj[kir]*, *maj[kó:ra:]* (spp.), Gv/Zo *baj*
- (81) PT \**mǎ*- 'causative' (R \**mǎ*-) > PMG \**mǎ*- || PMu \**ma*- || PJu \**mǎ*- || PTpr \**mǎ*- || Kt *m*- (irregular vowel loss), Ari *mú*- || Kr *ma*- || Gv *mā*-
- (82) PT \**ŋgə* 'cultivated field' (R \**ŋko*/\**ŋke*) > Pu <*tá*><sub>KG</sub> || Pa *ŋa*, Ar <*ngá*>, Gv/Zo *ga*
- (83) PT \**ŋgəar* 'sun' (R \**ŋat*\*) > Pa *ŋád*, Ar/Gv/Zo/Sl *gát*
- (84) PT \**pətic* 'heavy' (R \**pocij*) > Kr *piʔti* || Pa *pati-ga*, Ar <*patíí* ~ *pati-í*> 'heavy, thick', Gv *patíi*
- (85) PT \**tə* 'to go (singular)' (R \**co*) > Kr *ti*
- (86) PT \**tə*- 'third person coreferential' > Kr *to*-
- (87) PT \**tək* 'larva' (R \**ček*<sup>w</sup>) > Pa [*ka*]*dé:g* 'coconut larva', (?) [*mā:m*]*nég* 'caterpillar sp.', (?) *nég* 'boil', Ar <*mandeg*> 'caterpillar', Gv *mā-dék* 'lizard'
- (88) PT \**jəu* 'blood' > Kr *ju*

The reflex *e*, listed by Rodrigues (2005) for Karo and Puruborá, simply does not occur in the available data. In Karo, one finds *o*, *i*, *ə*, *ǎ*, *a*, and *u* (with no obvious distribution), and in Puruborá, *ə* is found in all cases. Currently we have no explanation for the reflexes in Karo (but note that Rodrigues' proposal also fails to account for them).

In the Mondé languages, one does indeed find *a* and *e* in accordance with Rodrigues' (2005) predictions (PT \**e* > *e*; PT \**o* > *a*). However, it is also possible to account for the Mondé reflexes if one recognizes that the etyma of all the aforementioned cognate sets contained one and the same vowel, PT \**ə*. Note that all instances of *e* occur following a coronal consonant ('leaf', 'larva') or word-initially ('house/village'), whereas all instances of *a* ('hand', 'snake', 'cultivated field', 'sun', 'heavy') occur following a peripheral (labial or velar) consonant. We propose, therefore, that PT \**ə* was fronted to Proto-Mondé \**e* following coronal consonants or word-initially and yielded Proto-Mondé \**a* elsewhere, and that the Mondé languages lend no support to the Proto-Tupian age of the distinction between the correspondence sets (b) and (c). We parenthetically note that the fronting of the type \**ə* > *e* following coronal consonants is also known from the history of Djeoromitxí, a Macro-Jê language of the Jabutian branch (Voort, 2007, p. 147), which is, like the Mondé languages, spoken in the Rondonian East. Typologically, the



functioning of labial and velar consonants as a natural class in processes triggering vowel backing (as in  $*\text{ə} > \text{a}$ ) is amply documented (Hyman, 1973; Vago, 1976)<sup>13</sup>.

## GENERAL PHONETIC CONSIDERATIONS

The development PT  $*\text{eC}^w > *oC$ , posited by Rodrigues (2005), inter alia for his 'Eastern' Tupian languages, conjoined with the reconstruction of such consonants with labial off-glides — the factor that accounts for these environmentally restricted vocalic outcomes — raise two issues of phonetic plausibility of historical reconstructions: one related to directionality considerations of the presumed coloring effect of the PT consonants, the other related to the distribution of the labialized consonants. First of all, a baffling aspect of the labializing effect exerted by these consonants is that it always affects the preceding, not the following vowel: a pre-vocalic labialized stop has no effect on a following vowel. From a phonetic point of view this is extremely counterintuitive. If labialization, or, more precisely, a labial release feature, is to play the role of contrastive feature distinguishing between these consonants (i.e.  $*p^w$  and  $*k^w$ ) and their plain counterparts (i.e.  $*p$  and  $*k$ ), one would expect its 'coloring' influence upon adjacent vowels to be realized more strongly (if not exclusively) on a following rather than a preceding vowel (that is, in a  $C^w$ -to-V transition, as opposed to the V-to- $C^w$  boundary).

As to their distribution, Rodrigues' (2007) PT labialized stops tend to occur, or are found quite frequently, in word-final position. In fact, the most significant phonotactic gap in their distribution in Rodrigues' (2007) proposal is the absence of  $*p^w$  from word-initial position. The expectation, commented on above, that a consonant with a secondary articulation will exert a stronger coarticulatory effect on a following rather than a preceding vowel derives from the fact that such segments depend, for their realization, on a following resonant element. As a consequence, we also expect such consonants to be less-optimally realized (*qua* contrastive segments) in word-final or pre-consonantal position, with no vocoid to work as a base for its contrastive release features to be imposed. As a matter of fact, plenty of evidence suggests that this is the case (see, e.g., Blevins, 2004, p. 116). In the words of Ladefoged and Maddieson (1996, p. 357):

Thus we can say that labialization is typically concentrated on the release phase of the primary articulation it accompanies. This observation has both phonetic and phonological significance. Many more languages have a restriction between the presence of labialization and the choice of the following vowel, than between its presence and the choice of the preceding vowel, and in many languages with labialized consonants the set of syllable-final consonants, if any, does not include labialized ones.

Aside from general considerations stemming from principles of acoustics and perceptual phonetics, it is not hard to find cross-linguistic evidence supporting the contention that such secondary articulations found in stop consonants behave phonologically as if 'looking for' a supporting vowel. Thus, in Khwarshi, an Eastern Caucasian language, labialization is found as a secondary articulation feature, mostly in velar and uvular consonants (Khalilova, 2009, pp. 17-18). The contrast is restricted basically to word-initial and word-medial position preceding a vowel, as in *et<sup>w</sup>a* 'fly' vs. *eta* 'touch', *lak<sup>w</sup>a* 'see' vs. *laka* 'lick'. The dynamic phonology of the language also demonstrates a preference for such labialized release

<sup>13</sup> In principle, it is still possible that the regular Mondé reflex of PT  $*\text{ə}$  is *a* even after *t/d*. This would allow us to propose two new Tupian etymologies for Mondé roots at the expense of the etymology for 'larva' shown above: PT  $*\text{təx}$  'to pound, to grind' > Pa *-tagá* 'to smash' (as in *to-dagá* 'to pound'), Gv *tágá* 'to beat'; PT  $*\text{dəp}$  'bitter' > Pa *[pe]táp*, Ar *<petab>*, Gv *[pe]típ* (note that Gv *i* is usually derived from /a/ in diminutives). We thank an anonymous reviewer for bringing the Paiter form *[pe]táp* 'bitter' to our attention.



consonants to occur preceding a vowel. Labialization is either lost (89a) or transferred to another consonant, one that precedes a vowel (89b), whenever a  $-C(V)$  suffix is added to a root containing a final labialized stop.

- (89) a. *l-ak<sup>w</sup>-a* → *l-ak-fe*                      b. *l-ak<sup>w</sup>-a* → *l-ak-x<sup>w</sup>-a*  
           IV-see-INF                      IV-see-PRS                      IV-see-INF                      IV-see-CAUS-INF

The joint effect of these generalizations, both static phonotactic patterns and processes in the dynamic phonology of Khwarshi, is to suggest that having consonants with a labial release preceding something other than a vowel is a highly undesirable or marked configuration in this language. Similar regularities are found in the phonologies of many unrelated languages, and can be understood more broadly in terms of the acoustic and phonetic constraints mentioned above, the same that make Rodrigues' (2007) proposal of stop consonants with secondary offglides that are almost always realized in contexts other than that of a following vowel very implausible.

## INTERIM SUMMARY

Above we have presented evidence against positing a sound change whereby PT  $*e$  would have acquired rounding ( $> *o$ ) preceding labialized consonants as well as against reconstructing these labialized consonants for PT. Instead, we have proposed that the suspect correspondence should be derived from PT  $*ə$ . Moreover, in our reconstruction, PT  $*ə$  accounts for some sound correspondences deemed irregular in Rodrigues' (2005) proposal as well as for the correspondences which underlie his reconstruction of  $*o$ . Our proposal is summarized in Table 5, where we list the reflexes of PT  $*ə$  as well of two other vowels which do not present split reflexes: PT  $*e$  and  $*o$  (in Rodrigues', 2005, 2007 interpretation,  $*e$  and  $*u$ ).

Table 5. PT  $*ə$ ,  $*e$ , and  $*o$  and their reflexes. <sup>A</sup> = before PT  $*i$  or  $*u$  in the next syllable;<sup>14</sup> <sup>B</sup> = before a vowel; <sup>C</sup> = next to a labial in a stem-final syllable; <sup>D</sup> = after a labial; <sup>E</sup> = after a coronal.

PT	PMG	PMu	PJu	PTpr	Kt/Ari	Kr	Pu	Mo
$*e$	$*e$	$*e$	$*e$ , $*a^A$	$*e$	$a/a\epsilon$	$e$	$e$	$e$ , $i^A$
$*ə$	$*o$ , $*u^B$	$*i$ , $*o^B$	$*a$ , $*u^{BC}$	$*e$ , $*o^D$	$a/a\epsilon$ , $i/u^D$	$o$ , $i$ , $i$ , $ə$ , $ě$ , $a$ , $u$	$ə$	$a$ , (?) $e^E$
$*o$	$*u$	$*o$	$*u$	$*o$	$i/u$	$o$	$o$	$o$

The notational change in the reconstruction of the sole rounded vowel of Proto-Tupian ( $*o$ ) as opposed to ( $*u$ ) does not affect the correctness of the sound correspondences identified by Rodrigues (2005) in any way. It is suggested by the fact that the typical realization of its reflex is a mid vowel in most Tupian branches (Mundurukú, Tuparian, Karo, Puruborá, and Mondé). Since Rodrigues' (2005)  $*o$  is reinterpreted as an unrounded vowel  $*ə$  in our account, it is now unproblematic to reconstruct  $*o$  in PT stems such as  $*aměko$  'jaguar',  $*jacjo$  'armadillo',  $*jako$  'lizard',  $*jeko$  'monkey',  $*jōk$  'flea',  $*jopi(-?a)$  'egg',  $*ko$  'to ingest',  $*ndo$  'hill, rock',  $*ndok$  'to eat (intr.)',  $*ōp$  'to give',  $*ōt$  'I' (and the first person prefix  $*o-$ ),  $*toko$  'to bite',  $*top$  'to see',  $*waco$  'alligator', among many others (we do not list these well-established etymologies in Appendix 1 for reasons of space).

<sup>14</sup> It has not been previously suggested that the combination PT  $*e...i/u$  has special reflexes in Juruna ( $*a...i$ ) and Mondé ( $i...i$ ). Examples include PT  $*ewit$  'bee, honey',  $*ejw$  'marico bag',  $*nēcuk$  'horsefly' and, in Juruna only,  $*mēpit$  'son (female ego)', which are reflected as PJu  $*awitš-á$ , —,  $*nātík-á$ ,  $*mābi-a$ ; Aruá  $<ivirej \sim ividei>$ ,  $<itji>$ ,  $<digá>$  (but  $*mēpit > <mambid>$ ).

## PT \**u* VS. \**i*

In this section, we will argue that it is necessary to reconstruct two distinct vowel phonemes in place of Rodrigues' (2005) Proto-Tupian \**i*. According to Rodrigues (2005), Proto-Tupian \**i* would have undergone a number of splits in the daughter languages depending on the immediate consonantal environment, yielding *i/i* in Yudjá, *e/i* in Karitiana, *i/ɛ* in Tuparí, *i/i* in Karo, and *i/i* in Puruborá. Unfortunately, Rodrigues (2005) does not specify the consonantal environments which would have triggered the putative fronting of \**i* in the daughter languages, nor is he explicit on whether these environments were identical for each constituent branch (Juruna, Arikém, Tuparian, Karo, and Puruborá). In what follows, we show that Rodrigues' (2005) reconstruction collapsed two correspondence sets into one and that two distinct vowels must therefore be reconstructed for Proto-Tupian. We symbolize them as PT \**u* and \**i*<sup>15</sup>, as in the minimal pair PT \**ju* 'liquid' vs. PT \**ji* 'urine', still retained in Karitiana as *se* 'liquid' and *si* 'urine'. Their reflexes are identical in some branches (PMG \**i*, PMu \**i*, PJu \**i*, Mo *i*); for this reason, in what follows we are concerned only with the remaining branches (that is, Tuparian, Arikém, Karo, and Puruborá). At the end of the section, however, we will see that the distinction between the PT vowels in question is indirectly preserved in the Mundurukú branch as well.

PT \**u* is reconstructed for the correspondence set which involves the following reflexes: PTpr \**i* (> Ma/Sk/ Ak *i*, Wy/Tu *ɛ*), Kt/Ari *e*, Pu *i*. In Karo, one usually finds *i* in the word-initial position (*i-ci* 'water', *iti* 'deer'), *ə* if the syllable is nasal (*nəp* 'louse', *wakəja* 'agouti'), and *i* elsewhere (*i-ci* 'water', *ma-ɣip* 'tree', *tɪt* 'to cook', *itɪ* 'deer', *jaɪ* 'howler monkey'); we are as of now unable to account for the apparently aberrant reflex *pək* 'to burn'. In Karo *ju* 'blood', the vowel *u* continues the PT sequence \**əu* and is thus not necessarily irregular. Below we list the PT etyma where evidence from multiple branches converges to the reconstruction of PT \**u* as opposed to PT \**i* (reflexes in branches which do not distinguish between them are omitted).

- (90) PT \**jau* 'howler monkey' > PTpr \**jaɪ* || Kr *jaɪ* || (?) Pu *aɜɪ*
- (91) PT \**jəu* 'blood' > PTpr \**jeɪ* || Ari *næe* || Kr *ju*
- (92) PT \**ju* 'liquid' > Kt *se*, Ari *-se* || Kr *i-ci* 'water'<sup>16</sup>
- (93) PT \**kwnĩṅã(ɾ)* 'scorpion' > PTpr \**kɪnĩṅã* || Kt *kennõn*, Ari *kednõð-o*
- (94) PT \**k'wɪɾ* 'green' > PTpr \**ɪɾ* 'unripe, young' || Kt *ket* 'green, blue', Ari *keð-* 'blue'
- (95) PT \**kɯc* 'earth' > PTpr \**kɪc* 'earth' || Kt *ɣej*, Ari (?) *ej-o* || Pu *ɪc*  
derived: PT \**kɯc-pi* 'earth' > Kt *ejə-pi*
- (96) PT \**kɯp* 'tree; stick-like' > PTpr \**kɪp* || Kt *ɣep*, Ari (?) *ɛp-o* 'tree; bone' || Kr *ma-ɣip* || Pu *ɪp*
- (97) PT \**kɯɾ* 'son' > PTpr \**kɪɾ* 'sperm; child; youngster' || Kt *ɣet* 'son; paternal aunt (*female ego*)', Ari (?) *ɛð-o* 'son; daughter (*female ego*)'
- (98) PT \**ṅgw* 'liquid' > PTpr \**ṅgɪ* 'liquid, saliva' || Kt *ṅge* 'blood'
- (99) PT \**ṅgwɾ* 'louse' > PTpr \**ṅgɪp* || Kt *ṅgep*, Ari *ṅgep-o* || Kr *nəp* || Pu *tɪp*
- (100) PT \**ṅgwɯɾ* 'salt' > PTpr \**ṅgɪɣɪɾ* || Pu *tɪɾ[a]*
- (101) PT \**pɯk* 'to burn' > PTpr \**pɪk-* || Kr *pək*
- (102) PT \**tɯɾ* 'to cook' > PTpr \**tɪɾ* || Kt *[o]tet* || Kr *tɪt*
- (103) PT \**ɯp* 'dirty, black' > PTpr \**ɪp* || Kt *ɛ:m*, Ari *ɛm-*
- (104) PT \**ɯtɯɯ* 'deer' > PTpr \**ɪtɪ:* || (?) Kt/Ari *nde* || Kr *itɪ* || Pu *idɪ*
- (105) PT \**ɯtɯ-pə* 'vine' > PTpr \**ɪtɪmbo* (only Ma <*õtömbó*>) || Kt *tepi*, Ari *tepu*
- (106) PT \**wăkuĩṅã* 'agouti' > PTpr \**wăkɪṅã* || Kr *wakəja* || Pu *wakɪṅã*
- (107) PT \**wɯ* 'to blow' > PTpr \**wɪ* || Kt *he:*
- (108) PT \**ɯ* 'water' > PTpr \**ɣɪ(:)* || Kt/Ari *e* 'rain', *e-se* 'water, river' || Kr *i-ci*  
derived: PT \**ɯ-ɣa* 'gourd' > PTpr \**ɣɪɣa*

<sup>15</sup> Note that we do not claim that the phonetic values of these segments were necessarily a back and a central unrounded vowel, respectively. At this time, other interpretations (such as \**i/i* vs. \**i/ɪ*) cannot be discarded.

<sup>16</sup> If Kr *ci* 'liquid' and Pu *fi* 'chicha; blood, menstruation' turn out to be related, they would constitute an exception. However, it appears possible to derive them from PT \**ji* 'urine' (> 'liquid').

The following example can be considered regular if it turns out that Lemos Barbosa (1951) transcription <mixon><sub>B</sub> of the Arikém cognate (see Rondon & Faria, 1948, p. 199) stands for *mũsõ* (for the development PT \**w* > Kt/Ari *m* in nasal environments, compare Kt *mĩnõ* 'Brazil nut').

- (109) PT \**wuɕjã* 'shell' > PTpr \**wĩnã* (only Tu *unã*) || Ari *mũsõ*

In the following examples, the PT etymon is preserved in Tuparian, but not in other branches which distinguish between \**u* and \**i*. Based on the Tuparian evidence, we reconstruct \**u*.

- (110) PT \**wpeɕ* 'duck', \**əɕwɐ/\*jəɕwɐ* 'arrow', \**twt* 'digging stick', \**uβicjo* 'wind', \**kupi(-kwt ~ -ɔwt)* 'younger sister (female ego)',<sup>17</sup> \**nēcɯk* 'horsefly', \**ejw* 'marico bag' > PTpr \**ipek*, \**ekip/\*jekip*, \**tɪt*, \**ibijo* (only Tu *upsió*), \**kɪpi*, \**nēcik*, \**ejɪ*

There are also two PT etyma which have reflexes in Karitiana, but not in other languages which distinguish between \**u* and \**i*: PT \**mbVɔwt ~ \*mbVɕwt* 'necklace',<sup>18</sup> \**tɪɔwt ~ \*tɪɕwt* 'maternal aunt' > Kt *mboɔet ~ mōet*, *teɔet*. Based on the Karitiana evidence, we reconstruct PT \**u* (we assume that the latter stem is derived from PT \**tɪ* 'mother' and thus reconstruct \**i* in the first syllable, translaryngeal assimilation of vowels is not unheard of in Karitiana; alternatively, the Karitiana word could be related to Ari <utaířã><sub>N</sub> 'id.').

PT \**i* is reconstructed for the correspondence set which involves the following reflexes: PTpr \**i* (preserved as *i* in all daughter languages), Kt/Ari *i*, Pu *i*. Below we list the PT etyma where evidence from multiple branches converges to the reconstruction of PT \**i* as opposed to PT \**u* (reflexes in branches which do not distinguish between them are omitted).

- (111) PT \**apɪ* 'ear' > PTpr \**api-jep*, \**api-* || Kt *opi* 'earring', Ari *opi* 'earring', *opikurumæɔæβ-o* 'earhole'  
 (112) PT \**aɔɪ* 'sloth' > PTpr \**aɔi-ato* (only Sk *aj-atso*) || Kt *oɔi*, Ari *o(?)i* || Kr *aɔi* || Pu *aɔi*  
 (113) PT \**(j)āpɪ* 'nose' > PTpr \**nāpi(-ɔa)* || Kt *ndɔpi-ɔop*, Ari *nōpi* 'nose', *nōpi-mæɔæβ-o* 'nostril', *nōpi-mbɐ* 'nasal adornment'  
 (114) PT \**(j)atɪ* 'pain' (in some languages also 'sour') > PTpr \**jati* || Kt *oti*  
 (115) PT \**jɪ* 'urine' > PTpr \**ji-ngi* (only Wy *ndi-gu*) || Kt/Ari *si* || Kr *ci* 'liquid' || Pu <*žídoka*><sub>KG</sub> 'urine', *ji* 'chicha; blood, menstruation'<sup>19</sup>  
 (116) PT \**jiɪt* 'flower' > PTpr \**ji:ɪ(-ɔa)* || Kt *ew-o-si:t*, Ari *eβ-o-si:ð-o*  
 (117) PT \**mbɪ/\*pɪ* 'foot' > PTpr \**mbi(-to)* || Kt/Ari *pi* || Kr *pi* || Pu *ji-be*  
 (118) PT \**mbɪɔa/\*pɪɔa* 'liver' > PTpr \**pɪɔa* || Kr *pía* || Pu *bía*  
 (119) PT \**mɪcō* 'curassow' > PTpr \**mícō* || Kt *mbisí*, Ari *misũ*  
 (120) PT \**pətɪc* 'heavy' > PTpr \**poti* || Kt *piti* || Kr *piɔti*  
 (121) PT \**pɪ* 'inner part' > PTpr \**-pi* || Kt *-pi*, as in *ejɐ-pi* 'earth'  
 (122) PT \**pɪcja* 'heel' > Kt *pi:so*, Ari *piso* || Pu *bija[ka]*  
 (123) PT \**pɪcjo* 'breath' > PTpr \**pijo* || Kt *[nō]pisi*  
 (124) PT \**tɪ* 'mother' > PTpr \**ti* || Kt/Ari *ti*  
 (125) PT \**wɪ* 'ax' > PTpr \**wi* || Kt *hi* 'knife (dated)', Ari *hi* || Pu *wi-a*  
 (126) PT \**wica* 'stone' > Kt *i(t)so* 'whetstone', Ari *iso(-o)* || Kr *ijá*  
 (127) PT \**wɪp* 'cooked' > Kt *hip* 'to roast', (?) Ari *mbuð-o-hip-op-o* 'frying pan' || Kr *wip*

To these one may add the verbs Tu *[e]pik-*, Kt *(a-)mbik* and possibly Pu *[t]api-a* 'to sit', if these are related to PT \**apɪk* 'to sit'.

<sup>17</sup> This root has a plausible cognate in Karitiana, *kipeet*, but the vowel of the first syllable is in any case irregular.

<sup>18</sup> Rodrigues (2007, p. 190) lists Tu *oir-pot* as a cognate, which could be a counterexample to our proposal. However, PT \**mb* is expected to be preserved as *p* in Tuparí, casting doubt on the validity of the etymology.

<sup>19</sup> Semantically, Kr *ci* 'liquid' and Pu *ji* 'chicha; blood, menstruation' are closer to PT \**jw* 'liquid', but the vowel \**u* would be expected to yield Karo and Puruborá *i* in this position.

In the following examples the apparent mismatch between the reflexes in Tuparian and Arikém is explained by vowel fusion: PT *\*ewi* > *\*æhi* > Kt *e:* (instead of *\*ai*), Ari *e* (instead of *\*æi*); PT *\*i(p)kw* or *\*i(p)ʔw* > *\*iʔe* > Kt *e:* (instead of *\*ie*).

- (128) PT *\*ewiʔ* 'bee, honey' > PTpr *\*ewiʔ* || Kt *e:t*, Ari *eð-o* 'bee', Ari *eð-o-se* 'honey' || (?) Kr *[p]ewiʔ* || Pu *iwiʔ*  
 (129) PT *\*kɪp(-kwɪ ~ -ʔwɪ)* 'brother' > PTpr *\*kɪp* 'younger brother (*male ego*)' || Kt *ke:t* 'younger brother (*male ego*)', Ari *keð-o*  
 (130) PT *\*kupi(-kwɪ ~ -ʔwɪ)* 'younger sister (*female ego*)' > PTpr *\*kɪpi* || Kt *kipe:t*

In the following examples, the reconstruction of PT *\*i* as opposed to PT *\*u* is based on non-converging evidence from only one branch, as no cognates in other diagnostic branches have been identified.

- (131) PT *\*tɪk* 'resin', *\*jepi* 'payment', *\*kupi(-kwɪ ~ -ʔwɪ)* 'younger sister (*female ego*)', *\*(?)ike(-kwɪ ~ -ʔwɪ)* 'elder sibling', *\*uβiɕjo* 'wind' > PTpr *\*tɪk* (only Tu *sik*), *\*jepi* (only Tu *epsi*), *\*kɪpi*, *\*ike* 'elder brother (*male ego*)', *\*iβijo* (only Tu *uɾsió*)  
 (132) PT *\*pɪtɪk* 'to take, to grab', *\*wati* 'moon' > Kt *pɪtik* 'to take out, to empty', *oti* (Ari *otɪ*)  
 (133) PT *\*wetɪk* 'sweet potato' > Pu *<witiká>*<sub>KG</sub>, *<widiká>*<sub>ES</sub>

Finally, one etymology has been found with a mismatch between Tuparian and Karo regarding the reconstruction of *\*i* or *\*u*. It is possible that the Karo word is ultimately unrelated to PT *\*mēpiʔ*, as the vowel correspondence in the first syllable and the nasal reflex of PT *\*p* appear to be irregular.

- (134) PT *\*mēpiʔ* 'son (*female ego*)' > PTpr *\*mēpiʔ* 'child, sororal nephew/niece (*female ego*)' || (?) Kr *mēmət*

We now turn to the Juruna branch, for which Rodrigues (2005) also posited a split of PT *\*i* into Yu *i*, *i*. However, this alleged split does not have anything to do with the reconstruction of PT *\*u* and *\*i* proposed in this section: in our account, both vowels yield PJu *\*i*, usually preserved in the daughter languages (PT *\*kiwɪ* 'green', *\*ŋgɪwɪ* 'louse', *\*nēcɪwɪ* 'horsefly', *\*mbiʔa/\*piʔa* 'liver', *\*(j)atɪ* 'pain, sour', *\*ewiʔ* 'bee, honey', *\*tɪk* 'resin', *\*pɪtɪk* 'to take, to grab', *\*mēpiʔ* 'son (*female ego*)', *\*apɪk* 'to sit', *\*wɪp* 'cooked' > PJu *\*[a]kɪʒ-ú*, *\*kɪpá*, *\*nātɪk-á*, *\*bɪʔá*, *\*fadɪ* 'sour', *\*awɪʒ-á*, *\*dɪk-á*, *\*pidɪk-ú*, (?) *\*mābi-a* 'daughter', *\*abɪk-ú*, *\*[u]wɪp-u* 'to cook'). A complication arises from the fact that PJu *\*i* has a special reflex in the word-initial position in Yudjá (*e-*) and Xipaya (*i-*): PT *\*kwɪc* 'earth' > PJu *\*it-á* > Yu *etá* 'sand, beach'; PT *\*kwɪc-pi* 'earth' > PJu *\*ipí-á* > Yu *epíá*, Xi *ipia*; PT *\*kwɪp* 'tree; stick-like' > PJu *\*ipá* 'stick' > Yu *epá*, Xi *ipa*. Note that the extinct Manitsawá appears to have preserved a non-front vowel in this case, as in *<upá>* 'Holz' (Steinen, 1886, p. 361). In the following example, Yudjá appears to have raised the vowel in the prevocalic position, feeding glide epenthesis: PT *\*ʔwɪ* 'water' > PJu *\*i-á* > *\*e-á* > *\*i-á* > Yu *ijá* (cf. Xi *ija*). Finally, in the following two etymologies PT *\*i* is reflected in unexpected ways in the Juruna languages: PT *\*pətiɕ* 'heavy', *\*ti* 'mother' > PJu *\*padét-ú*, *\*di-á*. At present, we are unable to account for the development of these words.

We conclude this section with presenting a piece of indirect evidence for the opposition between PT *\*u* and *\*i* from the Mundurukú branch. In Proto-Mundurukú, PT *\*p* is lost before front vowels, including *\*i* < *\*i*, as in PT *\*pe* 'path', *\*peo* 'pus', *\*pe* 'tobacco', *\*pi* 'foot', *\*picja* 'heel', *\*pitik* 'to take, to grab', *\*jepi* 'payment', *\*w-epik* 'to revenge' > PMu *\*e*, *\*ei* 'to swell', *\*e*, *\*i*, *\*(?)iða*, *\*itɪk*, *\*ðéi*, *\*w-eik* (cf. Rodrigues, 2007, p. 173). However, in all available examples *\*p* is retained before a PT *\*u*: PT *\*puw* 'to burn', *\*kɪp-kwɪ ~ \*kɪp-ʔwɪ* 'brother' > PMu *\*pik*, *\*kipit*, suggesting that PT *\*i* and *\*u* were still distinct at the time when PT *\*p* was lost before front vowels (i.e., PT *\*i* had already become front, but PT *\*u* had not). This argument is unfortunately not as strong as it could have been in light of the existence of cognate sets in which PT *\*p* was not lost even before reflexes of PT *\*i* or *\*i* (PT *\*jupi-ʔa* 'egg',

\**pi*jo ‘breath’, \**kupi*(-*ɰuɪ* ~ -*ʔuɪ*) ‘younger sister (*female ego*)’, \**pi* ‘inner part’ > PMu \**ðopi*ə, \**pi*ðo, \**kibj*t, \**pi*),<sup>20</sup> but the fact that there is not a single example of loss of PT \**p* preceding a \**i* < PT \**u* is hardly due to chance.

## IMPLICATIONS FOR THE SUBGROUPING OF TUPIAN

Rodrigues and Cabral (2002, 2012) and Rodrigues (2007, p. 170) put forward a hypothesis whereby the Tupian family is considered to be split, in a binary manner, into two large branches: ‘Eastern’ (comprising Sateré-Mawé, Awetí, Tupí-Guaraní, Mundurukú, and Juruna) and ‘Western’ (comprising Mondé, Tuparian, Arikém, Ramarama, and Puruborá). These two proposals, however, have remained insufficiently demonstrated in the sense that they have not yet been supported with bunches of identified shared innovations.

Figure 1 summarizes the distribution of the innovations identified in this paper. Only phonologically significant innovations shared by more than one low-level group are shown, namely: (i) the merger of PT \**i* and \**j* (Puruborá, Arikém, Tuparí, Karo, Mondé, Mundurukú), (ii) the merger of PT \**ə* and \**e* in the default environment (Arikém, Tuparí), (iii) the merger of PT \**ə* and \**o* after labials (Arikém, Tuparí), (iv) the merger of PT \**u* and \**i* (Mondé, Mundurukú, Juruna, Sateré-Mawé, Awetí, Tupí-Guaraní), and (v) the merger of PT \**ə* and \**o* before vowels (Mundurukú, Juruna, Sateré-Mawé, Awetí, Tupí-Guaraní).

The following groups share more than one innovation related to the evolution of PT \**ə* and \**i*: Arikém and Tuparí (3 innovations), Mundurukú, Juruna, Sateré-Mawé, Awetí, and Tupí-Guaraní (2 innovations), and Mundurukú and Mondé (2 innovations). Of these, the former two sets are strong candidates for valid clades: both include non-trivial,

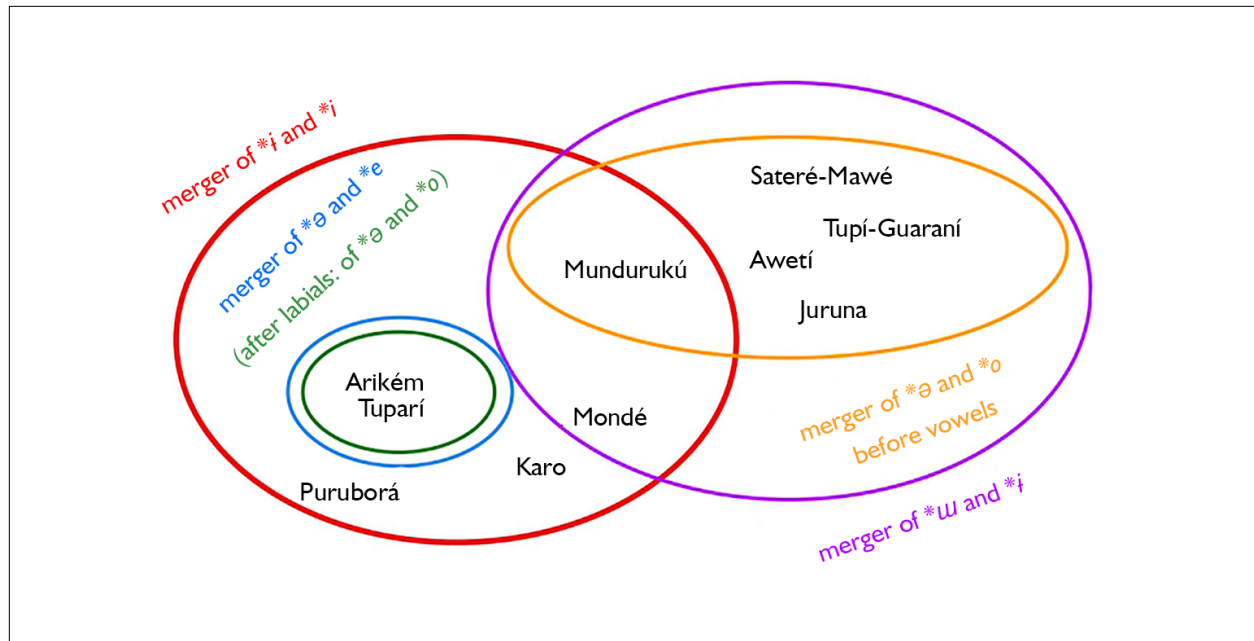


Figure 1. Distribution of innovations affecting PT \**ə* and \**i*.

<sup>20</sup> PT \**mbi*ʔa/\**pi*ʔa ‘liver’ > PMu \**pi*ə is not necessarily an exception, as in this case one might suspect that the PT absolute form with \**mb*- was generalized in PMu.

positionally conditioned innovations (merger of *\*ə* and *\*o* following labials in Tuparí and Arikém, preceding vowels in Mundurukú, Juruna, Sateré-Mawé, Awetí, and Tupí-Guaraní). In contrast, the set comprising Mondé and Mundurukú is in all likelihood spurious (or paraphyletic): in addition to being incompatible with the proposal which links Mundurukú to Juruna and Mawé-Guaraní, there is indirect evidence which suggests that the fronting of PT *\*u* in Mundurukú counterfed the loss of *\*p* before front vowels, an innovation specific to that branch. Therefore, the triple merger of PT *\*i*, *\*ɨ*, and *\*u* as *\*i* has probably occurred independently in the phonological history of Mondé and Mundurukú.

That way, evidence from the development of the PT vowels supports the identification of two mid-level clades within Tupian. The node comprising Tuparí and Arikém is defined by the sound change *\*ə* > *\*e* (default) / *\*o* (after labials); we suggest the label Tuparikém for this subgrouping hypothesis.<sup>21</sup> The vowel inventory of Proto-Tuparikém (*\*i e ə o*) is preserved without changes in Proto-Tuparian, whereas in Proto-Arikém these vowels yielded /*i e æ ɒ ʉ*/ (> Karitiana /*i e a o i*/) by means of a vowel shift identified by Storto and Baldi (1994). The second clade includes Mundurukú, Juruna, and Mawé-Guaraní (that way, our findings partially corroborate Rodrigues' (2005) hypothesis regarding the validity of his Eastern branch) and has the merger of PT *\*ə* and *\*o* before vowels as its defining innovation. It may have proceeded in two stages: first, PT *\*ə* and *\*o* may have changed into *\*o* and *\*u* in a chain shift (this is precisely the state reconstructed by Rodrigues, 2005 for Proto-Tupian); in turn, the vowel *\*o* (from PT *\*ə*) may have been raised to *\*u* in prevocalic contexts (thus merging with *\*u* from PT *\*o*). After that, Proto-Eastern Tupian *\*o*, *\*u* yielded PMG *\*o*, *\*u*; PMu *\*i*, *\*o*; PJu *\*a* (*\*u* next to labials in stem-final syllables), *\*u*. As the Eastern Tupian languages reach their greatest diversity between the Lower Madeira and the Lower Iriri, the Proto-Eastern Tupian *Urheimat* has to be sought in that region.

## CONCLUSION

This paper has presented a reconstruction of the Proto-Tupi (PT) inventory of oral vowels alternative to that advanced by Rodrigues (2005), this being clearly the accepted view on the PT vocalism since its adoption in reference works on the family such as Rodrigues (1999) and Rodrigues and Cabral (2012). Our proposal is summarized in Table 6.

Table 6. PT vowels and their reflexes (proposal). <sup>A</sup> = before PT *\*ɨ* or *\*u* in the next syllable; <sup>B</sup> = before a vowel; <sup>C</sup> = next to a labial in a stem-final syllable; <sup>D</sup> = after a labial; <sup>E</sup> = after a coronal.

PT	PMG	PMu	PJu	PTpr	Kt/Ari	Kr	Pu	Mo
<i>*i</i>	<i>*i</i>	<i>*i</i>	<i>*i</i>	<i>*i</i>	<i>i</i>	<i>i</i>	<i>i</i>	<i>i</i>
<i>*ɨ</i>	<i>*ɨ</i>	<i>*i</i>	<i>*ɨ</i>	<i>*i</i>	<i>i</i>	<i>i</i>	<i>i</i>	<i>i</i>
<i>*u</i>	<i>*ɨ</i>	<i>*i</i>	<i>*ɨ</i>	<i>*ɨ</i>	<i>e</i>	<i>ɨ</i>	<i>ɨ</i>	<i>i</i>
<i>*e</i>	<i>*e</i>	<i>*e</i>	<i>*e</i> , <i>*a</i> <sup>A</sup>	<i>*e</i>	<i>a/æ</i>	<i>e</i>	<i>e</i>	<i>e</i> , <i>i</i> <sup>A</sup>
<i>*ə</i>	<i>*o</i> , <i>*u</i> <sup>B</sup>	<i>*ɨ</i> , <i>*o</i> <sup>B</sup>	<i>*a</i> , <i>*u</i> <sup>BC</sup>	<i>*e</i> , <i>*o</i> <sup>D</sup>	<i>a/æ</i> , <i>i/ʉ</i> <sup>D</sup>	<i>o</i> , <i>ɨ</i> , <i>i</i> , <i>ə</i> , <i>ẽ</i> , <i>a</i> , <i>u</i>	<i>ə</i>	<i>a</i> , (?) <i>e</i> <sup>E</sup>
<i>*o</i>	<i>*u</i>	<i>*o</i>	<i>*u</i>	<i>*o</i>	<i>i/ʉ</i>	<i>o</i>	<i>o</i>	<i>o</i>
<i>*a</i>	<i>*a</i>	<i>*a</i>	<i>*a</i>	<i>*a</i>	<i>o/ɒ</i>	<i>a</i>	<i>a</i>	<i>a</i>

<sup>21</sup> It is interesting to note that the defining innovation of this branch did not affect Kepkiriwat, an extinct language of Rondônia sometimes classified as Tuparian (cf. Hanke et al., 1958, p. 188; Rodrigues, 1999, p. 109; Galucio, 2001, pp. 5–6; Aragon, 2008, pp. 6, 10–11, 2014, pp. 3, 15, 19–20; Rodrigues & Cabral, 2012, p. 497, *inter alia*). The default reflex of PT *\*ə* in Kepkiriwat appears to be *o* rather than *e*, as in <uóque><sub>R</sub>, <uóc><sub>B</sub> 'house, village', <óp><sub>B</sub> 'leaf', <gó><sub>B</sub> 'cultivated field' (Rondon & Faria, 1948, pp. 181, 187, 191) < PT *\*ə*k 'house' or maybe the first person form *\*o-jə*k 'my house', *\*jə*p 'leaf', *\*ɣgə* 'cultivated field'. This suggests that Kepkiriwat is not a Tuparian language but rather forms a branch of its own. The issue awaits further investigation.

We have argued that this new proposal is superior to the Rodrigues (2005) reconstruction in that it avoids the postulation of unexplained bifurcations of reflexes and the proposal of exception-ridden splits that, moreover, lack phonetic plausibility. Rodrigues' (2007) proposal of a series of labialized consonants to PT is rejected too, as the segments in question lack reflexes different from those of their plain counterparts and because the positional developments of contextual vowels were shown to be spurious.

## ACKNOWLEDGMENTS

We are grateful to two anonymous reviewers for their comments on the presentation and the substance of the paper. These have certainly improved the quality of our submission, and the authors are fully responsible for any remaining errors or shortcomings. We are especially grateful to the editors and technical staff of the Boletim for their swift and high-quality work in preparing the proofs and dealing with our observations on necessary adjustments and revisions.

## ABBREVIATIONS

1	First person	Mu	Mundurukú
3	Third person	Mw	Sateré-Mawé
3CRF	Coreferential third person	Pa	Suruí-Paiter
IV	Class IV	PJu	Proto-Juruna
abs.	Absolute	PL	Plural
Ak	Akuntsu	PMG	Proto-Mawetí-Guaraní
Ar	Aruá	PMu	Proto-Mundurukú
Ari	Arikém	PRS	Present
Aw	Awetí	PT	Proto-Tupian
CAUS	Causative	(P)TG	(Proto-)Tupí-Guaraní
Cl	Cinta-larga	PTpr	Proto-Tuparian
excl.	Exclusive	Pu	Puruborá
Gv	Gavião	rel.	Relational
incl.	Inclusive	sg	Singular
INF	Infinitive	Sk	Sakurabiat (Mekéns)
Kp	Kepkiriwat	Sl	Salamãý
Kr	Karo	Tu	Tuparí
Kt	Karitiana	Wy	Wayoró
Ku	Kuruaya	Xi	Xipaya
Ma	Makurap	Yu	Yudjá
Mn	Manitsawá	Zo	Zoró
Mo	Mondé		

## REFERENCES

- Alarcon, D. F., Millikan, B., & Torres, M. (2016). Apresentação. In D. F. Alarcon, B. Millikan & M. Torres (Orgs.), *Ocekadí: Hidrelétricas, conflitos socioambientais e resistência na Bacia do Tapajós* (pp. xv–xvii). Programa de Antropologia e Arqueologia da Universidade Federal do Oeste do Pará.



- Alves, P. M. (2004). *O léxico do Tuparí: Proposta de um dicionário bilingue* [Ph.D. dissertation, Universidade Estadual Paulista Júlio de Mesquita].
- Aragon, C. C. (2008). *Fonologia e aspectos morfológicos e sintáticos da língua Akuntsú* [MA thesis, Universidade de Brasília].
- Aragon, C. C. (2014). *A grammar of Akuntsú, a Tupian language* [Ph.D. dissertation, University of Hawai'i at Mānoa].
- Betts, L. V. (1981). *Dicionário Parintintin-português, português-parintintin*. Sociedade Internacional de Linguística.
- Blevins, J. (2004). *Evolutionary Phonology*. Cambridge University Press.
- Bontkes, W. (1978). *Dicionário preliminar: Suruí-português, português-suruí*. Summer Institute of Linguistics.
- Braga, A. O. (1992). *A fonologia segmental e aspectos morfofonológicos da língua Makurap (Tupí)* [MA thesis, Universidade Estadual de Campinas].
- Braga, A. O. (2005). *Aspects morphosyntaxiques de la langue Makurap/Tupí* [Ph.D. dissertation, Université de Toulouse–Le Mirail].
- Carvalho, F. O. (2019). Revisitando o Proto-Jurúna: A reconstrução da série de oclusivas orais. In E. S. Oliveira, E. A. Vasconcelos & R. D. Sanches (Eds.), *Estudos Linguísticos na Amazônia* (pp. 215–236). Pontes Editores.
- Caspar, F. (n.d.). *German-Tupari dictionary: A unpublished manuscript*. <http://www.etnolinguistica.org/caspar:german-tupari-a>.
- Corrêa da Silva, B. C. (2010). *Mawé/Awetí/Tupí-Guaraní: Relações linguísticas e implicações históricas* [Ph.D. dissertation, Universidade de Brasília].
- Costa, R. N. V. (2002). Fonologia segmental da língua Kuruaya. *Moara*, (17), 85–101.
- Dietrich, W. (2009). Correspondências fonológicas e lexicais entre Karitiána (Arikém, Tupí) e Tupí-Guaraní. *Revista Brasileira de Linguística Antropológica*, 1(2), 25–48. <https://doi.org/10.26512/rbla.v1i2.12365>
- Drummond, C. (1952–1953). *Vocabulário na Língua Brasília* (2 Vols.). Boletim da Faculdade de Filosofia, Ciências e Letras da Universidade de São Paulo.
- Fargetti, C. M. (2001). *Estudo fonológico e morfossintático da língua Juruna* [Ph.D. dissertation, Universidade Estadual de Campinas].
- Fargetti, C. M., & Rodrigues, C. L. R. (2008). Consoantes do Xipaya e do Juruna: Uma comparação em busca do proto-sistema. *Alfa*, 52(2), 535–63.
- Felzke, L. F., & Moore, D. (2019). Terminologias de parentesco dos grupos da família linguística Mondé. *Boletim do Museu Paraense Emílio Goeldi. Ciências Humanas*, 14 (1), 15–32. <https://doi.org/10.1590/1981.81222019000100003>
- Figueiredo, M. V. (2010). *A flecha do ciúme: O parentesco e seu avesso segundo os Aweti do Alto Xingu* [Ph.D. dissertation, Universidade Federal do Rio de Janeiro].
- Franceschini, D. (1999). *La langue sateré-mawé : Description et analyse morphosyntaxique* [Ph.D. dissertation, Université Paris VII–Denis Diderot].
- Gabas Jr., N. (1989). *Estudo fonológico da língua Karo (Arara de Rondônia)* [MA thesis, Universidade Estadual de Campinas].
- Gabas Jr., N. (1999). *A Grammar of Karo, Tupi (Brazil)* [Ph.D. dissertation, University of California, Santa Barbara].
- Galucio, A. V. (2001). *The morphosyntax of Mekens (Tupí)* [Ph.D. dissertation, University of Chicago].
- Galucio, A. V., & Gabas Jr., N. (2002). Evidências de agrupamento genético Karo-Puruborá, tronco Tupi. *XVII Encontro Nacional da ANPOLL*, Universidade Federal do Rio Grande do Sul.
- Galucio, A. V. (2005). Puruborá: Notas etnográficas e linguísticas recentes. *Boletim do Museu Paraense Emílio Goeldi. Ciências Humanas*, 1(2), 159–92.



- Galucio, A. V., Meira, S., Birchall, J., Moore, D., Gabas Jr., N., Drude, S., . . . Rodrigues, C. R. (2015). Genealogical relations and lexical distances within the Tupí linguistic family: A lexicostatistical and phylogenetic approach. *Boletim do Museu Paraense Emílio Goeldi: Ciências Humanas*, 10(2), 229–274. <https://doi.org/10.1590/1981-81222015000200004>
- Galvão, E. (1953). Cultura e sistema de parentesco das tribos do Alto Xingu. *Boletim do Museu Nacional*, 14, 1–56.
- Gavião, I. K. S. (2019). *Nomes, verbos, adjetivos, posposições e predicções na língua dos Ikólóéhj (Gavião, fam. Mondé, tronco Tupi)* [MA thesis, Universidade de Brasília].
- Gomes, D. M. (2006). *Estudo morfológico e sintático da língua Mundurukú (Tupi)* [Ph.D. dissertation, Universidade de Brasília].
- González, H. A. (2008). Una aproximación a la fonología del tapiete (Tupí–Guaraní). *LIAMES*, 8(1), 7–43. <https://doi.org/10.20396/liames.v8i1.1469>
- Hanke, W., Swadesh, M., & Rodrigues, A. D. (1958). Notas de fonologia Mekens. In J. Comas (Ed.), *Miscellanea Paul Rivet octogenario dicata* (Vol. II, pp. 187–217). Universidad Nacional Autónoma de México.
- Hyman, L. (1973). The feature [grave] in phonological theory. *Journal of Phonetics*, 1, 329–37.
- Kakumasu, J. Y., & Kakumasu, K. (2007 [1988]). *Dicionário por tópicos Kaapor–português*. Associação Internacional de Lingüística–SIL Brasil.
- Karitiana, N. (2016). *Cultura, memória e aspectos da variação linguística da língua do povo Byyyjty Osop Aky na aldeia Kyōwā da terra indígena Karitiana* [Licentiate Degree Monograph, Universidade Federal de Rondônia].
- Khalilova, Z. (2009). *A Grammar of Khwarshi* (LOT Dissertation Series). LOT, Netherlands Graduate School of Linguistics.
- Koch-Grünberg, T. (1932). Wörterlisten “tupý”, maué und purúborá. *Journal de la Société des Américanistes*, 24(1), 31–50.
- Lacerda, M. C. (2014). *Bekā Pamakube (lugar de aprender). Aprendendo com os Zoró: Análise da identidade indígena através da experiencia das escolas nas aldeias do povo indígena Zoró* [Ph.D. dissertation, Universidad de Salamanca].
- Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. Blackwell.
- Landin, D. (2005). *Dicionário e léxico Karitiana / português* (2 ed.). Sociedade Internacional de Lingüística.
- Lemos Barbosa, P. A. (1951). *Pequeno vocabulário Tupi–português. Com quatro apêndices: Perfil da língua Tupi; palavras compostas e derivadas; metaplasmos; síntese bibliográfica*. Livraria São José.
- Lima, S. O. (2008). *A estrutura argumental dos verbos na língua Juruna (Yudjá)* [MA thesis, Universidade de São Paulo].
- Meer, T. H. (1982). *Fonologia da língua Suruí* [MA thesis, Universidade Estadual de Campinas].
- Meira, S., & Drude, S. (2015). A summary reconstruction of proto-maweti–guarani segmental phonology. *Boletim do Museu Paraense Emílio Goeldi: Ciências Humanas*, 10(2), 275–296. <https://doi.org/10.1590/1981-81222015000200005>
- Mello, A. A. S. (2000). *Estudo histórico da família lingüística Tupi–Guarani: Aspectos fonológicos e lexicais* [Ph.D. dissertation, Universidade Federal de Santa Catarina].
- Mendes Jr., D. G. (2007). *Comparação fonológica do Kuruáya com o Mundurukú* [MA thesis, Universidade de Brasília].
- Monsserrat, R. M. F. (2005). Notícia sobre a língua Puruborá. In A. D. Rodrigues & A. S. A. C. Cabral (Eds.), *Novos estudos sobre línguas indígenas* (pp. 9–22). Editora UnB.
- Moore, D. (1984). *Syntax of the language of the Gavião Indians of Rondônia, Brazil* [Ph.D. dissertation, City University of New York].
- Moore, D. (2005). Classificação interna da família lingüística Mondé. *Estudos Lingüísticos*, 34, 515–20.
- Moore, D., & Galucio, A. V. (1994). Reconstruction of Proto-Tupari consonants and vowels. In M. Langdon & L. Hinton (Eds.), *Proceedings of the Meeting of the Society for the Study of the Indigenous Languages of the Americas, July 2–4, 1993, and the Hokan-Penutian Workshop, July 3, 1993, both held at the 1993 Linguistic Institute at Ohio State University in Columbus, Ohio* (pp. 119–137). Survey of California and other Indian Languages.

- Nikulin, A., & Carvalho, F. O. (2019). Estudos diacrônicos de línguas indígenas brasileiras: Um panorama. *Macabéa: Revista Eletrônica do Netlli*, 8(2), 255–305.
- Nikulin, A., & Andrade, R. (2020). The rise and fall of approximants in the Tuparian languages. *Journal of Language Relationship*, 18(4), 284–319.
- Nimuendajú, C. (1923–1924). Zur Sprache der Šipáia-Indianer. *Anthropos*, 18–9(4–6), 836–57.
- Nimuendajú, C. (1928). Wortliste der Šipáia-Sprache. *Anthropos*, 23(5–6), 821–50.
- Nimuendajú, C. (1929). Zur Sprache der Šipáia-Indianer. *Anthropos*, 24(5–6), 863–96.
- Nimuendajú, C. (1930). Zur Sprache der Kuruáya-Indianer. *Journal de la Société des Américanistes*, 22(2), 317–45.
- Nimuendajú, C. (1932). Wortlisten aus Amazonien. *Journal de la Société des américanistes*, 24(1), 93–119.
- Nogueira, A. F. S. (2011). *Wayoro êmêto: Fonologia segmental e morfossintaxe verbal* [MA thesis, Universidade de São Paulo].
- Nogueira, A. F. S. (2019). *Predicação na língua Wayoro (Tupí): Propriedades de finitude* [Ph.D. dissertation, Universidade de São Paulo].
- Nogueira, A. F. S., Galucio, A. V., Soares-Pinto, N., & Singerman, A. R. (2019). Termos de parentesco nas línguas Tuparí (família Tupí). *Boletim do Museu Paraense Emílio Goeldi. Ciências Humanas*, 14(1), 33–64. <https://doi.org/10.1590/1981.81222019000100004>
- Payne, D. L. (1991). A classification of Maipuran (Arawakan) languages based on shared lexical retentions. In D. C. Derbyshire & G. K. Pullum (Eds.), *Handbook of Amazonian languages* (Vol. 3, pp. 355–499). Mouton de Gruyter.
- Picanço, G. L. (2005). *Mundurukú: Phonetics, phonology, synchrony, diachrony* [Ph.D. dissertation, University of British Columbia].
- Picanço, G. L. (2019). *A fonologia diacrônica do Proto-Mundurukú (Tupí)*. Appris.
- Restivo, P. (1893 [1722]). *Lexicon Hispano-Guaranicum*. Wilhem Kohlhammer.
- Ribeiro, M. J. P. (2010). *Dicionário Sateré-Mawé/português* [MA thesis, Universidade Federal de Rondônia, campus de Guajará-Mirim].
- Rocha, I. (2011). *A estrutura argumental da língua Karitiana: Desafios descritivos e teóricos* [MA thesis, Universidade de São Paulo].
- Rocha, I. (2014). Processos de causativização na língua Karitiana. *Boletim do Museu Paraense Emílio Goeldi. Ciências Humanas*, 9(1), 183–97. <https://doi.org/10.1590/S1981-81222014000100012>
- Rodrigues, A. D. (1984–1985). Relações internas na família lingüística Tupí–Guaraní. *Revista de Antropologia*, 27(8), 33–53.
- Rodrigues, A. D., & Dietrich, W. (1997). On the linguistic relationship between Mawé and Tupí–Guaraní. *Diachronica*, 14(2), 265–302.
- Rodrigues, A. D. (1999). Tupí. In R. M. W. Dixon & A. Y. Aikhenvald (Eds.), *The Amazonian Languages* (pp. 107–24). Cambridge University Press.
- Rodrigues, A. D. (2002). Correspondências lexicais e fonológicas entre Tupí–Guaraní e Tuparí. In A. S. A. C. Cabral & A. D. Rodrigues (Orgs.), *Línguas indígenas brasileiras: Fonologia, gramática e história. Atas do I Encontro Internacional do Grupo de Trabalho sobre Línguas Indígenas da ANPOLL* (Vol. 1, pp. 288–297). Editora da Universidade Federal do Pará.
- Rodrigues, A. D. (2005). As vogais orais do Proto-Tupí. In A. D. Rodrigues & A. S. A. C. Cabral (Orgs.), *Novos estudos sobre línguas indígenas* (pp. 35–46). Editora UnB.
- Rodrigues, A. D. (2007). As consoantes do Proto-Tupí. In A. S. A. C. Cabral & A. D. Rodrigues (Orgs.), *Línguas e culturas Tupí* (pp. 167–203). Curt Nimuendajú.
- Rodrigues, A. D. (2008). Linguistic reconstruction of elements of prehistoric Tupí culture. In E. B. Carlin & S. Kerke (Eds.), *Linguistics and Archaeology in the Americas: The Historization of Language and Society* (Brill's Studies in the Indigenous Languages of the American, Vol. 2). Brill.
- Rodrigues, A. D., & Cabral A. S. A. C. (2002). Revendo a classificação interna da família Tupí–Guaraní. In A. D. Rodrigues & A. S. A. C. Cabral (Eds.), *Línguas Indígenas Brasileiras: Fonologia, Gramática e História* (pp. 327–37). UFPA.

- Rodrigues, A. D., & Cabral, A. S. A. C. (2012). Tupían. In L. Campbell & V. Grondona (Eds.), *The Indigenous languages of South America: A comprehensive guide* (Vol. 2, pp. 495–574). Mouton de Gruyter.
- Rodrigues, C. L. R. (1995). *Étude morphosyntaxique de la langue xipaya (Brésil)* [Ph.D. dissertation, Université Paris VII–Denis Diderot].
- Rondon, C. M. S., & Faria, J. B. (1948). *Glossário Geral das tribos silvícolas de Mato-Grosso e outras da Amazônia e do Norte do Brasil* (Tom. I). Imprensa Nacional.
- Sabino, W. K. (2016). *Awetýza tiʔíngatú: Construindo uma gramática da língua Awetý, com contribuições para o conhecimento do seu desenvolvimento histórico* [Ph.D. dissertation, Universidade de Brasília].
- Santos, C. A. B. (2013). *Aspectos da fonologia do Mundurukú do Madeira (AM)* [MA thesis, Universidade de Brasília].
- Schleicher, C. O. (1998). *Comparative and internal reconstruction of the Tupi–Guarani language family* [Ph.D. dissertation, University of Wisconsin–Madison].
- Sekelj, T. (1948). *Wordlist Aruá-Makurap-Zaboti-Arikapú-Tupari*. Unpublished manuscript. <http://www.etnolinguistica.org/sekelj:1>
- Silva, E. B. (2009). *Estruturas fonéticas e fonológicas de vogais e consoantes da língua Kuruaya* [MA thesis, Universidade Federal do Pará].
- Singerman, A. R. (2018). *The morphosyntax of Tuparí, a Tupían language of the Brazilian Amazon* [Ph.D. dissertation, University of Chicago].
- Snethlage, E. (1932). Chipaya- und Curuaya-Wörter. *Anthropos*, 27, 65–93.
- Snethlage, E. H. (1934). *Wörterverzeichnis der Borôbyrá-Sprache*. Unpublished manuscript. <http://www.etnolinguistica.org/emil:4>
- Steinen, K. (1886). *Durch Central-Brasilien: Expedition zur Erforschung des Schingú im Jahre 1884*. F. A. Brockhaus.
- Storto, L. (1999). *Aspects of a Karitiana grammar* [Ph.D. dissertation, Massachusetts Institute of Technology].
- Storto, L., & Baldi, P. (1994). The Proto-Arikém vowel shift. *Paper presented at the 68<sup>th</sup> Annual Meeting of the Linguistic Society of America*.
- Vago, R. (1976). More evidence for the feature [grave]. *Linguistic Inquiry*, 7(4), 671–4.
- Voort, H. V. (2007). Proto-Jabutí: Um primeiro passo na reconstrução da língua ancestral dos Arikapú e Djeoromitxi. *Boletim do Museu Paraense Emílio Goeldi. Ciências Humanas*, 2(2), 133–68. <https://doi.org/10.1590/S1981-81222007000200007>
- Wetzels, W. L., & Nevins, A. (2018). Prenasalized and postoralized consonants: The diverse functions of enhancement. *Language*, 94(4), 834–66. <http://dx.doi.org/10.1353/lan.2018.0055>
- Zoró, T. K., & Camargos, Q. F. (2019). Estruturas interrogativas polares e informacionais na língua Pangyĵĕj (Zoró, família Mondé, tronco Tupi). *Revista Brasileira de Linguística Antropológica*, 11(2), 111–33. <https://doi.org/10.26512/rbla.v11i02.28508>

## AUTHORS' CONTRIBUTIONS

A. Nikulin contributed to conceptualization, data curation, formal analysis, investigation, methodology, and project administration; and F. Carvalho to conceptualization, data curation, formal analysis, investigation, and methodology.

Appendix 1. Etymologies.

(Continue)

This appendix includes all etymologies that contain the relevant vowels in at least two branches of the family, as well as several other etymologies that were mentioned in the body of the text. In what follows, cognates which cannot be regularly derived from the reconstructed etyma are marked with (!), and the irregular reflexes of specific segments are underlined> (except in cases of irregular deletion of segments).

(1) PT *\*apə* 'interrogative word'

PMG *\*apo* > Mw *apo* 'yes/no question particle' (Ribeiro, 2010, p. 49; reconstruction ours) || PMu *\*abi* 'who' > Mu *ábə* ~ *abə*, Ku *abi* (Picanço, 2005, p. 209) || (?) PJu *\*apá* (!) 'what' > Yu *apá*, Xi *apa* (Fargetti & Rodrigues, 2008, p. 561; reconstruction ours) || PTpr *\*apo* 'who; what (subject)' > Wy *apo-k<sup>w</sup>a:t*, Tu *apó* 'who', Sk *apo* (Guaratira dialect) (Nogueira, 2019, p. 132; Alves, 2004, p. 150; Galucio, 2001, p. 167)

(2) PT *\*api* 'ear'

PMG *\*api[ta]* 'earhole' > Aw *?apita-kwar* 'earhole', PTG *\*apitsa* 'earhole', *\*[j]apitsa[ka]* 'to think' (Sabino, 2016, p. 97; Mello, 2000, p. 165; reconstruction ours) || PTpr *\*api-jep*, *\*api-* > Ma *api-cep*, Wy *api-tep*, Tu *apsi[rɪp]-ʔa* (cf. *apsi-kur* 'earring', *apsi-ʔe* 'to listen'), Sk/Ak *api-tep* (Braga, 1992, p. 31; Alves, 2004, pp. 150, 151; Galucio et al., 2015, p. 254; reconstruction ours) || Kt *opi* 'earring', Ari *opi* 'earring' (Storto, 1999, p. 14; Nimuendajú, 1932, p. 111, <api<sub>Nopi-kuru-mæʔæβ-o 'earhole' (Rondon & Faria, 1948, p. 194; Nimuendajú, 1932, p. 109, <opicúru maaú<sub>B</sub>>, <apikuru-maábə<sub>N</sub>>)</sub>

(3) PT *\*apik* 'to sit'

PMG *\*apik* > Mw/Aw *apik*, PTG *\*apik* (Corrêa da Silva, 2010, p. 410) || PMu *\*(?)ábik* > Mu *ábik*, Ku *ábik* (Picanço, 2019, p. 142) || PJu *\*abik-ú* > Yu *abikú*, Xi *biku* (!) (Fargetti & Rodrigues, 2008, p. 563; reconstruction ours) || (?) PTpr *\*[e]pik-* > Tu *epsik-* (Alves, 2004, p. 168; reconstruction ours) || (?) Kt *mbik* ~ *a-mbik* (!) (Storto, 1999, p. 41; Landin, 2005, p. 53) || (?) Pu *[t]api-a* (Galucio, 2005, p. 186)

(4) PT *\*aʔi* 'sloth'

PMG *\*aʔi* > PTG *\*aʔi* (Corrêa da Silva, 2010, p. 406) || PMu *\*aj* ~ *\*ai* > Mu *aj*, Ku *aj* (Picanço, 2019, p. 136) || PTpr *\*aʔi-ato* > Sk *aj-atso* (Galucio et al., 2015, p. 271; reconstruction ours) || Kt *oʔi*, Ari *o(?)i* (Landin, 2005, p. 20; Nimuendajú, 1932, p. 112, <ai<sub>N</sub>>) || Kr *aʔi* (Gabas Jr., 1999, p. 19) || Pu *aʔi* (Galucio, 2005, p. 172)

(5) PT *\*ðəp* 'bitter'

PMG *\*ðəp* > Mw *noP*, Aw *loP*, PTG *\*roP* (Nikulin & Carvalho, 2019, p. 292, endnote 11) || PMu *\*tʃip* > Mu *tʃəp*, Ku *tʃip* (Picanço, 2019, p. 138) || PTpr *\*tep* > Ma <tep>, Tu *tép-ʔa* 'bitter', *tép-ʔur* 'sour' (Sekelj, 1948; Alves, 2004, p. 258; reconstruction ours) || Kt *ta:p* (Landin, 2005, p. 29) || (?) Pa *[pe]tʃáp*, Ar <petab>, Gv *[pe]tʃip* (Sekelj, 1948; Gavião, 2019, p. 125; anonymous reviewer, personal communication, 2022)

(6) PT *\*ejw* 'marico bag'

PTpr *\*eji* > Ma *éci*, Wy *endw*, Tu *éu*, Sk *eti*, Ak *etí* (Nikulin & Andrade, 2020, p. 295) || Ar <itji> (Sekelj, 1948)<sup>22</sup>

(7) PT *\*etə-* 'sociative causative'

PMG *\*ero-* (!) > Mw *ero-*,<sup>23</sup> Aw *(e)zo-*, PTG *\*ero-* (Franceschini, 1999, pp. 231–237; Sabino, 2016, p. 67; Corrêa da Silva, 2010, p. 412; reconstruction ours) || (?) PMu *\*[ɔ]iɪi-* (!) > Mu *dəðə-* (Gomes, 2006, p. 82) || PTpr *\*ete-* > Ma *ete-*, Wy *ete-*, Tu *eté-* (3 *s-ite-*), Sk *etse-* (Braga, 2005, p. 186; Nogueira, 2019, p. 51ff.; Alves, 2004, p. 57; Singerman, 2018, pp. 125–130; Galucio, 2001, p. 98; reconstruction ours) || Kt *atot* 'to take away' (Storto, 1999, p. 106) || Kr *ta-* (!) (Gabas Jr., 1999, pp. 65–67)

<sup>22</sup> An anonymous reviewer suggests that this form could be a Tuparian loan. Note, however, that the correspondences are regular and non-trivial, a fact which is consistent with its being an inherited form.

<sup>23</sup> With the verb *to* 'to go.sg', this prefix unexpectedly takes the form *ere-* (Franceschini, 1999, p. 234). In addition, its final vowel is raised to *u* before a *w* (*war* 'to go.pl' → *eru-war*) and undergoes regressive translaryngeal harmony, thus behaving identically to the vowel of the prefix *to-* '3CRF'.

## Appendix 1

(Continue)

- (8) PT *\*ewit* 'bee, honey'  
 PMG *\*ewit* > Mw *ewit*, Aw *ekit*, PTG *\*eir* (Corrêa da Silva, 2010, p. 404; reconstruction ours) || PMu *\*eit* > Mu *eit*, Ku *eít-i* (Picanço, 2019, p. 140) || PJu *\*awitá* > Yu *awitá* 'honey', Xi *wia* (!) (Fargetti & Rodrigues, 2008, p. 560; reconstruction ours) || PTpr *\*ewit* > Wy *ēngwīt*, Tu *ewit*, Sk *ekwīt*, Ak *ekwīt* (Nikulin & Andrade, 2020, p. 300) || Kt *e:t*, Ari *eð-o* 'bee' (Storto, 1999, p. 62; Nimuendajú, 1932, p. 111, <érā><sub>N</sub>), Ari *eðose* 'honey' (Storto, 1999, p. 62; Rondon & Faria, 1948, p. 195; Nimuendajú, 1932, p. 111, <êrocê><sub>B</sub>, <érā-sê><sub>N</sub>) || (?) Kr *[p]ewit* (Gabas Jr., 1999, p. 16) || Pu *iwit* (Rodrigues, 2005, p. 41)<sup>24</sup> || Ar *ivirei* ~ *ividei* 'bee, honey', Gv/Zo *íit* 'bee', Sl *íbit-áp* 'bee' (Sekelj, 1948; Galucio et al., 2015, p. 263)
- (9) PT *\*ək* (absolute)/*\*jək* (relational) 'house'  
 PMG *\*ok/\*cok* > Mw *sok* 'nest', Aw *ok*, PTG *\*ok/\*-rok* (Meira & Drude, 2015, p. 293) || PMu *\*ík-ʔa/\*ðík-ʔa* > Mu *ək-ʔá/dək-ʔá* (Picanço, 2005, pp. 60, 118; reconstruction ours) || PJu *\*ak-á* > Yu *aká*, Xi *aka* (Fargetti & Rodrigues, 2008, p. 560; reconstruction ours) || PTpr *\*ek/\*jek* > Ma *ek/cek*, Tu *ek/hek*, Wy *ek/ndek*, Sk *ek/tjek*, Ak *ek/tek* (Nikulin & Andrade, 2020, pp. 295–296) || Ari *æk-o* (Rondon & Faria, 1948, p. 195; Nimuendajú, 1932, p. 111, <ácu><sub>B</sub>, <ákā><sub>N</sub>)<sup>25</sup> || Pu *ʔək-a* (Galucio, 2005, p. 172; Monserrat, 2005, p. 19) || Ar *eg* 'house, village' (Sekelj, 1948)<sup>26</sup>  
 derived: PT *\*ək-at*  
 PMG *\*ok-at* > PTG *\*ok-at* (Mello, 2000, p. 184; PMG reconstruction ours) || Kt *ak-ot* 'together', Ari *ækvð-o* 'plaza' (Landin, 2005, p. 6; Nimuendajú, 1932, p. 111, <uiakádō><sub>N</sub>)  
 derived: PT *\*ək-ēr* (absolute)/*\*jək-ēr* (relational) 'plaza'/'door' (cf. *\*ēr* 'mouth')  
 PMG *\*ok-ēr* 'door' > Mw *oken-ipi*, Aw *ot-ēr*, PTG *\*ok-ēr* (Meira & Drude, 2015, p. 292) || PMu *\*ík-ēn/\*ðík-ēn* > Ku *ifjēn*, 3<sup>rd</sup> *t-ifjēn* 'door' (Snethlage, 1932, p. 71, as <teschan><sub>s</sub> or <ischene><sub>L</sub>; reconstruction ours) || PTpr *\*ek-ēr* > Tu *ek-ēr* 'plaza' (Alves, 2004, p. 164; reconstruction ours) || Kt *ak-ān* 'village' (Storto, 1999, p. 112)
- (10) *\*əkup* (absolute)/*\*jəkup* (relational) 'arrow'  
 PMG *\*uʔip/\*cuʔip* > Aw *uʔwip*, PTG *\*uʔip/\*-ruʔip* (Sabino, 2016, p. 226; Mello, 2000, p. 203; reconstruction ours) || PMu *\*op/\*ðop* > Mu *op/dop*, Ku *ɔp/lɔp* (Picanço, 2019, p. 139) || PTpr *\*ekip/\*jekip* > Ma *ekip/tjekip*, Tu *ekup* (Nikulin & Andrade, 2020, pp. 296–297) || (?) Pa *jáb*, Ar *ndjap*, Gv *džáp/ádžáp* ~ *ídžáp*,<sup>27</sup> Zo *džap* (Bontkes, 1978, p. 19; Sekelj, 1948; Moore, 1984, p. 150; Zoró & Camargos, 2019, p. 122)
- (11) PT *\*(ʔ)ike(-kwit ~ -ʔwit)* 'elder sibling'  
 PMG *\*(ʔ)ikeʔit* 'elder brother (*male ego*)' > Mw *ʔikeʔet* (!), Aw *ʔitiʔit* (!), PTG *\*ʔikeʔit* (Meira & Drude, 2015, p. 294) || PTpr *\*ike* 'elder brother (*male ego*)' > Wy *ike*, Tu *-ike* ~ *-cke*, Sk/Ak *ike* (Nogueira et al., 2019, p. 42)
- (12) PT *\*jaʔək* 'army ant'  
 PMG *\*caʔok* > PTG *\*taʔok*<sup>28</sup> (reconstruction ours) || PMu *\*ðaʔik* > Mu *daʔək* (Alarcon et al., 2016, p. xvi; reconstruction ours) || PTpr *\*jakek* > (?) Wy *akek* (!), Tu *hakék[e]*, Sk *takek* (Nikulin & Andrade, 2020, p. 294)
- (13) PT *\*(j)āpi* 'nose'  
 PMG *\*ʔāpi* > Mw *ʔāpi*<sup>29</sup> (Ribeiro, 2010, p. 48; reconstruction ours) || (?) PMu *\*abi* > Mu *abi* 'tip' (Gomes, 2006, p. 225; reconstruction ours) || (?) Xi *[j]amī-kua* (Galucio et al., 2015, p. 254) || PTpr *\*nāpi(-ʔa)* > Ma *nāpi*, Wy *āpi-a*, Tu *āpsí*, Sk *āpi-tsa*, Ak *āpi-ta* (Nikulin & Andrade, 2020, p. 297) || Kt *ndžopi-ʔop*, Ari *nōpi* (Storto, 1999, p. 20; Rondon & Faria, 1948, p. 193; Nimuendajú, 1932, p. 109,

<sup>24</sup> We were unable to locate this form in any primary source on Puruborá. However, it corresponds well to the data of other Tupian languages and thus we provisionally leave it here, given that Rodrigues (2005) may have had access to unpublished Puruborá data.

<sup>25</sup> Rodrigues (2005, p. 40) also includes Kt *s-ak* 'cave' and Kr *ek* 'inside' into this cognate set. We were unable to locate such forms in the available sources on these languages.

<sup>26</sup> An anonymous reviewer suggests that this form could be a Makurap loan.

<sup>27</sup> An anonymous reviewer notes that the basic Gavião (2019) form for 'arrow' is *idžáp* or *ijáp*, and that *adžáp* ~ *ajáp* derives from a process of reanalysis of the initial *i* as an affix and its change to *a*.

<sup>28</sup> Reconstruction based on Old Tupí *taʔók-a* (Lemos Barbosa, 1951, p. 148), Ka'apor *toʔok* (Kakumasu & Kakumasu, 2007 [1988], p. 30).

<sup>29</sup> Corrêa da Silva (2010, p. 402) lists Aw *ʔāpīc*, PTG *\*apīc* as cognates, but their palatal coda does not match the absence of a coda in Sateré-Mawé.

Appendix 1

(Continue)

- ⟨nhampí⟩<sub>B</sub>, ⟨u-nyāpí⟩<sub>N</sub>; Ari *nōpi-mǣʔæβ-o* 'nostril', *nōpi-mbu* 'nasal adornment' (Nimuendajú, 1932, p. 109, 111, ⟨u-inyampi-maábā⟩<sub>N</sub>, ⟨nyāpi-bú⟩<sub>N</sub>) || (?) Pa *ámī[ā(:)]*, Ar *amīi*, Gv *ámīi*, Zo *ámīi*, Sl *aβī[ʔā]* (Bontkes, 1978, p. 2; Galucio et al., 2015, p. 254)
- (14) PT *\*(j)ati* 'pain' (in some languages also 'sour')  
 PMG *\*cati* > Mw *sati*, Aw *ati*, PTG *\*tatsi/\*-ratsi* (Meira & Drude, 2015, p. 294) || PJu *\*jadí* 'sour' > Yu *i-fadí*, Xi *fadí* (Fargetti & Rodrigues, 2008, p. 560; reconstruction ours) || PTpr *\*jati* > Ma *cáti*, Wy *ati*, Tū *así*, Sk *atsi*, Ak *aʔí* (Nikulin & Andrade, 2020, p. 297) || Kt *oti* (Storto, 1999, p. 206) || Pa *[j]ati-ga*, Ar *⟨atiká⟩* 'ill', Gv *-atí* (Bontkes, 1978, p. 19; Sekelj, 1948; Moore, 1984, p. 112)
- (15) PT *\*jau* 'howler monkey'  
 PTpr *\*jai* > Wy *ndaú*, Tū *haú*, Sk *ta:ʔi*, Ak *tai-kóp* (Nikulin & Andrade, 2020, p. 294) || Kr *jai* (Galucio et al., 2015, p. 267) || (?) Pu *aʒi* (!) (Galucio, 2005, p. 181)
- (16) PT *\*jepi* 'payment'  
 PMG *\*cepi* > Aw *tepi*, PTG *\*tepi/\*-repi* (Corrêa da Silva, 2010, p. 407; reconstruction ours) || PMu *\*ðéi* > Mu *deí* (Picanço, 2005, p. 239; reconstruction ours) || (?) Yu *bí[á]* (!) 'to be expensive, to pay' (Fargetti, 2001, p. 186) || PTpr *\*jepi* > Tū *epsi* (Singerman, 2018, p. 86; reconstruction ours)
- (17) PT *\*jəβa* 'forehead'  
 PMG *\*coβa* 'face' > Aw *towa*, PTG *\*toβa/\*-roβa* (Sabino, 2016, p. 97; Mello, 2000, p. 183; reconstruction ours) || PTpr *\*jəβa* 'forehead' > Ma *cépa*, Tū *épa* 'eye', Wy *eβa*, Sk/Ak *eba-* (in compounds) (Nikulin & Andrade, 2020, pp. 292, 297)
- (18) PT *\*jəp* 'leaf'  
 PMG *\*-cop* > Mw *[i]hop*, Aw *op*, PTG 3<sup>rd</sup> *\*ts-op* (Meira & Drude, 2015, p. 293) || PMu *\*ðip* > Mu *dəp*, Ku *lip* (Picanço, 2019, p. 139) || PJu *\*súpá* > Yu *úpá*, Xi *supa* (Fargetti & Rodrigues, 2008, p. 561; reconstruction ours) || PTpr *\*jep* > Ma *cep*, Wy *ndep*, Tū *hep*, Sk *tep*, Ak *tep* (Nikulin & Andrade, 2020, p. 295) || Kt *sap*, Ari *-sæβ-o* (Storto, 1999, p. 62; Rondon & Faria, 1948, p. 200; Nimuendajú, 1932, p. 113, ⟨goxáua⟩<sub>B</sub>, ⟨éβā-sábā⟩<sub>N</sub>) || Kr *naʔ-jop* (Gabas Jr., 1999, pp. 26, 211) || Pu *təp* 'leaf, hair' (Galucio, 2005, p. 167) || Ar/Gv *tsep*, Zo *sep*, Sl *bá-sep* (Galucio et al., 2015, p. 252)
- (19) PT *\*jəw* 'blood'  
 PMG *\*cui* > Mw *su:*, Aw *tuwí[k]*, PTG *\*tuwí/\*-ruwí* (Meira & Drude, 2015, p. 292) || PMu *\*ðoj* > Mu *doj*, Ku *loj* (Picanço, 2019, p. 142) || PTpr *\*jei* > Ma *céi*, Wy *aú*, Tū *éu*, Sk *ai*, Ak *eʔi* (Nikulin & Andrade, 2020, p. 297) || Ari *jǣe* (Rondon & Faria, 1948, p. 194; Nimuendajú, 1932, p. 109, ⟨nhaé⟩<sub>B</sub>, ⟨nyaé⟩<sub>N</sub>) || Kr *ju* (Gabas Jr., 1999, p. 30)
- (20) PT *\*ji* 'urine'  
 PMG *\*ci* > Mw *si*, Aw *tí*, PTG *\*ti* (Corrêa da Silva, 2010, p. 403; reconstruction ours) || PTpr *\*ji-ngí* > Wy *ndi-gu* (Nogueira, 2019, p. 18; reconstruction ours) || Kt/Ari *si* (Landin, 2005, p. 24; Nimuendajú, 1932, p. 109, ⟨u-sī⟩<sub>N</sub>) || Kr *ci* 'liquid' (Gabas Jr., 1999, p. 92) || (?) Pu *⟨židoka⟩<sub>KG</sub>* (Koch-Grünberg, 1932, p. 34) || (?) Ar *⟨ndjikapá⟩* (Sekelj, 1948)
- (21) PT *\*jiiɾ* 'flower'  
 (?) PMG *\*[po]cí:ɾ* > Mw *pohi:ɾ*, Aw *[a]potiɾ*, PTG *\*potiɾ* (Meira & Drude, 2015, p. 293) || PMu *\*ðit* > Mu *dit*, Ku *lit* (Picanço, 2019, p. 139) || PTpr *\*ji:ɾ(-ʔa)* > Ma *kíp-cír-et*, Wy *ndi:r-a* (cf. *kup-ndi:ɾ* 'forest'), Tū *hír-ʔa*, (cf. (?) *hi:ɾ* 'side dish'), Sk *tir-a*, Ak *tirá* (Nikulin & Andrade, 2020, p. 295) || Kt *ew-o-si:ɾ*, Ari *eβ-o-si:ð-o* (Landin, 2005, p. 9; Rondon & Faria, 1948, p. 200; Nimuendajú, 1932, p. 113, ⟨euóxíra⟩<sub>B</sub>, ⟨éβā-ʔasiíra⟩<sub>N</sub>) || Gv *ʔiit*, (?) Ar *⟨iptjiu⟩* (Moore, 1984, p. 204; Sekelj, 1948)
- (22) PT *\*jopʔa* 'face'  
 PMu *\*ðópa* > Mu *dopá*, Ku *lopa* (Picanço, 2005, p. 61; Nimuendajú, 1930, p. 319; reconstruction ours) || Kt *si:po* 'eye', *sipo* 'seed'<sup>30</sup> (Storto, 1999, p. 15)

<sup>30</sup> Arikém has a very similar root, which differs from the Karitiana one in having a labial coda: *p-supopβ-o* ⟨axupáua⟩<sub>B</sub> / ⟨u-asupáua⟩<sub>N</sub> 'eye', *s-upopβ-o* ⟨ixipáua⟩<sub>B</sub> 'seed' (Rondon & Faria, 1948, pp. 193, 200; Nimuendajú, 1932, p. 109). It is plausibly cognate with Proto-Tuparian *\*jopap* 'grain' > Proto-Core Tuparian *\*opap* 'grain', *\*eβaopap* 'eye' (Nikulin & Andrade, 2020, pp. 292, 297). It is likely that Proto-Arikém inherited both roots from Proto-Tupian (say, Proto-Arikém *\*supo* 'eye', *\*supop* 'seed'), which were later contaminated in both Arikém languages due to their phonetic and semantic similarity: Karitiana retained only *\*supo*, and Arikém only *\*supop*.



## Appendix 1

(Continue)

(23) PT \**jw* 'liquid'

PMG \**ci* > Mw *h-i* 'river' (fossilized 3<sup>rd</sup> person), PTG \**ti/\*-ri* 'water (*in natura*)' (Corrêa da Silva, 2010, p. 154) || PMu \**ði* > Mu *di*, Ku *li* (Picanço, 2019, p. 140) || Kt *se*, Ari *-se* (Landin, 2005, p. 27; Rondon & Faria, 1948, pp. 193–195, 198; Nimuendajú, 1932, p. 110–111, as in *kɥ-se* <cucê><sub>B</sub> 'saliva', *æsp-se* <axóxê><sub>B</sub> 'tear', *nõm-se* <noncê, nõancê><sub>B</sub> 'milk', *e-se* <exê, ecê><sub>B</sub>, <esê><sub>N</sub> 'water, river', *eð-d-se* <êrocê><sub>B</sub>, <êrã-sê><sub>N</sub> 'honey') || Kr *i-ci* 'water' (Gabas Jr., 1999, p. 168) || Pu *fi* 'chicha' (Koch-Grünberg, 1932, p. 44; Snethlage, 1934; Monserrat, 2005, pp. 17, 19 (also 'blood, menstruation')); <ži><sub>KG</sub>, <schi><sub>ES</sub> || Ar <endjatji> 'tear', <namdji> 'milk', Gv *tji* 'liquid, sweat' (Sekelj, 1948; Moore, 1984, p. 204)

(24) PT \**kə* 'to go, to walk'

(?) PMG \**[e]ko* 'to be, to live' > Mw/Aw *eko*, PTG \**eko* (Meira & Drude, 2015, p. 292) || PMu \**[ki]ki* > Mu *dʒe-kəkə*, Ku <adikiki> (Picanço, 2019, p. 137; Nimuendajú, 1930, p. 338) || Kr *kə* (Gabas Jr., 1999, p. 60) || Pa *ka*; Gv *kaà*, Zo *ka* (Bontkes, 1978, p. 8; Moore, 1984, p. 31; Zoró & Camargos, 2019, p. 122)

(25) PT \**kəc* 'to plant'

PMG \**koc* > Mw/Aw *koc*, PTG \**ibi-koc* 'to dig' (Franceschini, 1999, p. 66; Sabino, 2016, p. 134; Mello, 2000, p. 206; reconstruction ours) || Pju \**kat-ú* 'to bury, to plant' > Yu/Xi *katú* (Fargetti & Rodrigues, 2008, p. 561; reconstruction ours) || Pa *káj*, Gv *kaja* (Meer, 1982, p. 27; Gavião, 2019, p. 93)

(26) PT \**kək* 'to hold'

PMG \**kok* > PTG \**kok* 'to support' (Corrêa da Silva, 2010, p. 408) || PMu \**kik* > Mu *kək* (Picanço, 2005, p. 61; reconstruction ours) || Kt *kak* (Landin, 2005, p. 13) || Kr *kik* (Gabas Jr., 1999, p. 86)

(27) PT \**kíp(-kwɪt ~ -ʔwɪt)* 'brother'

PMG \**kibiɪt* 'brother (*female ego*)' > Mw/Aw *kiwiɪt*, PTG \**kibiɪt* (Corrêa da Silva, 2010, p. 405; reconstruction ours) || PMu \**kipit* > Mu *kipit* 'younger brother (*female ego*)', Ku *kopit* (Picanço, 2005, p. 213, 2019, p. 139) || PTpr \**kíp* 'younger brother (*male ego*)' > Wy/Tu/Sk *kíp* (Nogueira et al., 2019, p. 42) || Kt *ke:t* 'younger brother (*male ego*)', Ari *keð-o* (Landin, 2005, p. 13; Nimuendajú, 1932, p. 110, <u-kêrã><sub>N</sub>)

(28) PT \**kɪpi(-kwɪt ~ -ʔwɪt)* 'younger sister (*female ego*)'

PMG \**kipiɪt* > PTG \**kipiɪt* (Corrêa da Silva, 2010, p. 407; reconstruction ours) || PMu \**kibit* > Mu *kibit* (Picanço, 2019, p. 139) || PTpr \**kipi* > Wy *kupi*, Tu *kupsí-ʔi*, Sk/Ak *kipi* (Nogueira et al., 2019; Alves, 2004, p. 207) || Kt *kipe:t* (!) (Landin, 2005, p. 15)

(29) PT \**kwnĩṅã(ɾ)* 'scorpion'

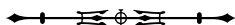
PTpr \**kũnĩṅã* > Wy *kɥ(ɥ)nĩṅã*, Tu *kwnĩká*, Sk *kinĩṅã* (Nikulin & Andrade, 2020, p. 308; reconstruction amended from \**kũnĩṅga*) || Kt *kennõn*, Ari *kednõð-o* (!) (Landin, 2005, p. 13; Nimuendajú, 1932, p. 113, <kēdnãṛã><sub>N</sub>)

(30) PT \**k'et* 'to sleep'

PMG \**ket* > Mw *ket*, Aw *teɾ*, PTG \**k'et* (Meira & Drude, 2015, p. 294) || PMu \**fet* > Mu/Ku *fet* (Picanço, 2019, p. 138) || PTpr \**ʔet-* > Ma/Wy *et-*, Tu *ʔet-*, Sk/Ak *et-* (Braga, 2005, p. 176; Nogueira, 2019, p. 50; Alves, 2004, p. 275; Galucio, 2001, p. 23; Aragon, 2014, p. 104; reconstruction ours) || Kt *kat*, Ari *kæt* (Storto, 1999, p. 14; Rondon & Faria, 1948, p. 197; Nimuendajú, 1932, pp. 115–116, <catauá><sub>B</sub>, <-kad, -katyu><sub>N</sub>) || Kr *ket* (Gabas Jr., 1999, p. 39) || Pu *ket-a* (Monserrat, 2005, p. 20; Galucio et al., 2015, p. 256) || Pa *ker*, Ar/Gv/Zo *kere*, Sl *keet* (Bontkes, 1978, p. 9; Moore, 1984, p. 60; Galucio et al., 2015, p. 256)

(31) PT \**k'wɪt* 'green'

PMG \**k'it* > Mw *k'it-ʔi*, Aw *k'it* 'green, blue', PTG \**k'it* (Corrêa da Silva, 2010, p. 411; reconstruction ours) || PMu \**kit* > Mu *kit* (Picanço, 2005, p. 213; reconstruction ours) || Pju \**[a]k'ib-ú* > Yu *ak'ibú ~ ak'ébú*, Xi *akiu* (Fargetti & Rodrigues, 2008, pp. 553, 563; reconstruction ours) || PTpr \**it* 'unripe, young' > Ma *it*, Wy/Tu *ɥɾ*, Ak *it* (Braga, 1992, p. 33; Alves, 2004, p. 271; Nogueira, 2019, p. 155; reconstruction ours) || Kt *ket* 'green, blue', Ari *keð-* 'blue' (Storto, 1999, pp. 14, 100, fn. 27; Landin, 2005, p. 13; Rondon & Faria, 1948, p. 196; Nimuendajú, 1932, p. 114, <taqueri><sub>B</sub>, <i-kerã><sub>N</sub>)



Appendix 1

(Continue)

(32) PT \**kɯc* 'earth'

PMG \**ʔic* > Mw *ʔic*, PTG \**ʔic* (Meira & Drude, 2015, p. 292) || PJu \**it-á* > Yu *etá* 'sand, beach' (Fargetti, 2001, p. 289) || PTpr \**kic* 'earth' > Ma *kic*, Wy/Tu *kuc*, Ak *kic* (Galucio et al., 2015, p. 259; reconstruction ours) || Kt *ʔej*, Ari (*ʔej*-o (Storto, 1999, p. 62; Rondon & Faria, 1948, p. 194; Nimuendajú, 1932, p. 110, <éio><sub>b</sub>, <ëya><sub>N</sub>) || Pu *ʔic* (Monserrat, 2005, p. 18; Galucio et al., 2015, p. 259)

derived: PT \**kɯc-pi* 'earth'

PMG \**ibi* > Mw *iwi-* (in *iwi-kuʔi* 'beach', *iwi-rup* 'ashes', *iwi-hĩk* 'dew'), Aw *iwi-* (in *iwi-kwat* 'hole', *iwi-lũʔũp* 'dirt', *iwi-tĩk* 'cloud'), PTG \**ibi* (Ribeiro, 2010, p. 98; Sabino, 2016, pp. 102, 149; Mello, 2000, p. 206; reconstruction ours) || PMu \**ipi* > Mu *ipi*, Ku *ipi* (Picango, 2005, p. 344; Silva, 2009, p. 89; reconstruction ours) || PJu \**ipí-á* > Yu *epíá*, Xi *ipia* (Fargetti & Rodrigues, 2008, p. 563; reconstruction ours) || Kt *ej-pi* (Landin, 2005, p. 8)

(33) PT \**kɯp* 'tree; stick-like'

PMG \**ʔip* > Mw *-ʔip*, Aw *ʔip*, PTG \**ʔip* (Meira & Drude, 2015, p. 294; Corrêa da Silva, 2010, p. 403) || PMu \**ʔip* > Mu/Ku *ʔip* (Picango, 2019, p. 141) || PJu \**ip-á* 'stick' > Yu *epá*, Xi *ipa*, Mn <upá> (Fargetti & Rodrigues, 2008, p. 562; Steinen, 1886, p. 361; reconstruction ours) || PTpr \**kíp* > Ma *kíp*, Wy/Tu *kup*, Sk *kíp-kiba* (cf. *kíp* 'club, log'), Ak *kíp* (Galucio et al., 2015, p. 252; Galucio, 2001, pp. 23, 45; reconstruction ours) || Kt *ʔep*, Ari (*ʔep*-o 'tree; bone' (Storto, 1999, p. 55; Rondon & Faria, 1948, pp. 194, 199–200; Nimuendajú, 1932, p. 109, <ébo, évo, -éuo, euó><sub>b</sub>, <ěbā, ěbā-, ěb-><sub>N</sub>) || Kr *ma-ʔip* (Gabas Jr., 1999, p. 165) || Pu *ʔip* (Galucio, 2005, p. 172; Monserrat, 2005, p. 18) || Pa *i:b*, Ar *iip*, Gv *iip*, Zo/Si *iip* (Bontkes, 1978, p. 6; Galucio et al., 2015, p. 252)

(34) PT \**kɯt* 'son'

PMG \**-ʔit* > Aw [*ta*]*ʔit* 'son (male ego)', PTG \**-ʔit* 'sororal nephew (male ego)<sup>31</sup> (Sabino, 2016, p. 87; reconstruction ours) || PMu \**ʔit* > Mu *ʔit* 'son (female ego)', Ku *ʔit* (Gomes, 2006, p. 240; Costa, 2002, p. 99; reconstruction ours) || PTpr \**kít* 'sperm; child; youngster' > Wy/Tu *kut* (Nogueira, 2019, pp. 55, 65; Alves, 2004, p. 208; reconstruction ours) || Kt *ʔet* 'son; paternal aunt (female ego)', Ari (*ʔeð*-o 'son; daughter (female ego)' (Storto, 1999, p. 14; Nimuendajú, 1932, p. 110, <u-ērā><sub>N</sub>)

(35) PT \**mbə* (absolute)\**pə* (relational) 'hand, finger, vine-like'

PMG \**mbo/po* > Mw *mo/po*, Aw *po*, PTG \**mbo/po* (Meira & Drude, 2015, p. 293)<sup>32</sup> || PMu \**pi* > Mu *pə*, Ku *pi* (Picango, 2019, p. 138) || PJu \**bu-á* > Yu *wá*, Xi *ba*, Mn <hubuá> (Fargetti & Rodrigues, 2008, p. 562; Steinen, 1886, p. 360; reconstruction ours) || PTpr \**mbo* > Ma *mbo*, Wy *mbo/-bo*, Tu *po*, Sk *po-pi*, Ak *po* (Nikulin & Andrade, 2020, p. 305) || Kt *pi*, Ari *pu* (Storto, 1999, p. 43; Rondon & Faria, 1948, p. 193; Nimuendajú, 1932, p. 109, <pú><sub>b</sub>, <u-pu><sub>N</sub>) || Kr *pʔ* 'cylindrical and small' (Gabas Jr., 1999, p. 165) || Pu *bə* 'long, vine-shaped' (Galucio, 2005, p. 167)<sup>33</sup> || Pa/Ar/Gv/Zo/Si *pá-be* (Bontkes, 1978, p. 14; Galucio et al., 2015, p. 255)

derived: PT \**utuw-pə* 'vine'

PMG \**itipo* > Mw *iripo* (!), Aw *itipo*, PTG \**itsipo* (Corrêa da Silva, 2010, p. 405; reconstruction ours) || PMu \**itfi-bi* > Mu *itfi-bə*, Ku *itfi-bi* (Picango, 2019, p. 137) || Kt *tepi*, Ari *tepu* (Rocha, 2011, p. 108; Nimuendajú, 1932, p. 113, <tepú><sub>N</sub>) || PTpr \**itimbo* (!) > Ma <ötömbó> (Sekelj, 1948)

derived: PT \**ja-pə* 'root'

PMG \**capo* > Mw *sapo*, Aw *tapo*, PTG 3<sup>rd</sup> \**ts-apo* (Meira & Drude, 2015, p. 294) || PMu \**tabi*(?) (fossilized 3<sup>rd</sup> person) > Mu *təbá*, Ku *tabi* (Picango, 2019, p. 142)

(36) PT \**mbəc* 'snake'

PMG \**mboc* > Mw *moc*, Aw *möc*, PTG \**mboc* (Meira & Drude, 2015, p. 295; reconstruction modified from \**mōj*) || PMu \**pij* > Mu *pəj-bá*, Ku *pij* (Picango, 2019, p. 137) || PJu \**put-á* > Yu *hutá*, Xi *huta* (Fargetti & Rodrigues, 2008, p. 560; reconstruction ours) || (?) Kr *māj[gāra]* (Gabas Jr., 1999, pp. 13, 33) || Pu *māñ[ũp]* (Galucio, 2005, p. 182; Monserrat, 2005, p. 15) || Pa *māj[kir]*, *māj[kó:ra:]* (spp.), Gv/Zo *baj* (Bontkes, 1978, p. 11; Galucio et al., 2015, p. 271)

<sup>31</sup> Root preserved in Old Tupi *i-ʔir-a* (Drummond, 1952–1953, p. 119), Kamayurá *je-ir* (Galvão, 1953).

<sup>32</sup> Meira and Drude (2015, p. 293) mention only Sateré-Mawé *po* and PTG \**po* (and reconstruct PMG \**po*). These allomorphs are found when a possessor is specified. The unpossessed form PMG \**mbo* is reconstructed based on Sateré-Mawé *mo*, PTG \**mbo* (Ribeiro, 2010, p. 72; Schleicher, 1998, p. 347).

<sup>33</sup> Rodrigues (2005, p. 39) and Galucio et al. (2015, p. 255) consider Karo *pá-* (as in *pá-pe?* 'hand', *pá-ro* 'hands'; Gabas Jr., 1999, p. 90) and Puruborá *ba* 'arm' (Galucio, 2005, p. 167) to belong here. From a phonetic point of view, these roots are rather comparable with PMu \**pə* 'arm' > Mu/Ku *pə* (Picango, 2019, p. 136).



## Appendix 1

(Continue)

- (37) PT
- \*mbi*
- (absolute)
- \*pi*
- (relational) 'foot'

PMG *\*mbi/pi* > Mw/Aw *mi/pi*, PTG *\*mbi/pi* (Meira & Drude, 2015, p. 293)<sup>34</sup> || PMu *\*i* > Mu *í*, Ku *i* (Picanço, 2005, p. 103; Nimuendajú, 1930, p. 318; reconstruction ours) || PTpr *\*mbi(-to)* > Ma *mbi*, Wy *mbi / -bi*, Tu *si-tó*, Sk *pi-tso*, Ak *pi* (Nikulin & Andrade, 2020, p. 305) || Kt/Ari *pi* (Landin, 2005, p. 24; Rondon & Faria, 1948, p. 194; Nimuendajú, 1932, p. 109, <pi><sub>B</sub>, <u-pi><sub>N</sub>) || Kr *pi* (Gabas Jr., 1999, pp. 39, 97) || Pu *fi-be* (Galucio, 2005, p. 166) || Pa *pi(:)-pe*, Ar/Gv/Zo *bi/pí*, Sl *pi-pe*, *bi-?a* (Bontkes, 1978, p. 16; Galucio et al., 2015, p. 255)

- (38) PT
- \*mbi?a*
- (absolute)
- \*pi?a*
- (relational) 'liver'

PMG *\*mi?a/pi?a* > Mw *mi?a/pi?a*, PTG *\*mbi?a/pi?a* (Ribeiro, 2010, p. 75; Corrêa da Silva, 2010, p. 402; reconstruction ours) || PMu *\*piā* > Mu *psā*, Ku *biā* (Picanço, 2019, p. 139) || PJu *\*bi?á* > Yu *bi?á*, Xi *bia* (Fargetti & Rodrigues, 2008, p. 561; reconstruction ours) || PTpr *\*pi?a* > Ma *pía*, Wy *pia*, Tu *si?á*, Sk *pitsa*, Ak *bíta* (Nikulin & Andrade, 2020, p. 309) || Kr *pía* (Gabas Jr., 1999, p. 54) || Pu *bia* (Galucio, 2005, p. 168)

- (39) PT
- \*mbV?uɪt*
- ~
- \*mbVkuɪt*
- 'necklace'

PMG *\*mbo?it* > PTG *\*mbo?it*<sup>35</sup> (reconstruction ours) || Kt *mbo?et* ~ *mōet* (Dietrich, 2009, p. 29; Landin, 2005, p. 16)

- (40) PT
- \*mēpit*
- 'son (female ego)'

PMG *\*mēpit* > Mw *mēpit*, Aw *mēpit*, PTG *\*mēmbit* (Meira & Drude, 2015, p. 295) || (?) PJu *\*mābi-a* (!) 'daughter' > Yu *mābia*, Xi *mābia* (Fargetti & Rodrigues, 2008, p. 561; reconstruction ours) || PTpr *\*mēpit* 'child, sororal nephew/niece (female ego)' > Ma/Wy *mēpit*, Tu *mēsít*, Sk *mēpit*, Ak *mēpit* (Nikulin & Andrade, 2020, p. 305) || (?) Kr *māmæt* (!) (Gabas Jr., 1989, p. 10) || Pa *mápit* 'son (female or male ego)', Ar <mambid>, Gv *māpit* (!) (Felzke & Moore, 2019; Sekelj, 1948)

- (41) PT
- \*mēr*
- 'husband'

PMG *\*mēr* > Aw *mēr*, PTG *\*mēr* (Corrêa da Silva, 2010, p. 408; reconstruction ours) || PJu *\*men-á* > Yu *mená*, Xi *mēna* (Fargetti & Rodrigues, 2008, p. 562; reconstruction ours) || PTpr *\*mēr* > Wy/Tu/Sk/Ak *mēr* (Nikulin & Andrade, 2020, p. 305) || Kt *mān*, Ari *māen/māēō-o* (Storto, 1999, p. 26; Rondon & Faria, 1948, p. 195; Nimuendajú, 1932, p. 110, <iman><sub>B</sub>, <u-madō><sub>N</sub>) || Kr *mēn* (Gabas Jr., 1999, p. 126) || Pu *mēr* (Galucio, 2005, p. 170; Monserrat, 2005, p. 14) || Ar <umen>, Gv *mēt* (Sekelj, 1948; Felzke & Moore, 2019, p. 32)

- (42) PT
- \*mā-*
- 'causative'

PMG *\*mā-* > Mw/Aw *mo-*, PTG *\*mā-* (Corrêa da Silva, 2010, p. 412; reconstruction ours) || PMu *\*ma-* > Mu *mā-*, Ku *ma-* (Picanço, 2019, p. 143) || PJu *\*mā-* > Yu/Xi *ma-* (Fargetti, 2001, p. 186; Nimuendajú, 1923, p. 847; reconstruction ours) || PTpr *\*mā-*<sup>36</sup> > Ma *mā-*, Wy *ō-/mā-*, Tu *ō-/m-*, Sk/Ak *ō-/mā-* (Braga, 2005, pp. 163–165; Nogueira, 2011, p. 145ff.; Alves, 2004, p. 57; Singerman, 2018, pp. 121–125; Galucio, 2001, pp. 96–8; Aragon, 2014, pp. 197–198; reconstruction ours) || Kt *m-*, Ari *mū-* (Rocha, 2014; Nimuendajú, 1932, as in <i-mu-paḡn-un><sub>N</sub> 'I made him/it dirty', <u-mu-paḡn-un><sub>N</sub> 'I made myself dirty', <es-un-i-mu-kupā><sub>N</sub> 'I put (lit. made descend) him/it on the ground') || Kr *ma-* (Gabas Jr., 1999, pp. 63–65) || Gv *mā-* (Moore, 1984, p. 190; Gavião, 2019, pp. 83–84)

- (43) PT
- \*māk*
- 'to stick, to glue'

PMG *\*māk* > Mw *mōk*, PTG *\*mōk* (Corrêa da Silva, 2010, p. 407; reconstruction ours) || PMu *\*māḡ* 'to lean' > Mu *māḡ*, Ku *māḡ* (Picanço, 2019, p. 138)

<sup>34</sup> Meira and Drude (2015, p. 293) mention only Sateré-Mawé *pi* and PTG *\*pi* (and reconstruct PMG *\*pi*). These allomorphs are found when a possessor of this now is specified. The unpossessed form PMG *\*mbi* is reconstructed based on Mw/Aw *mi*, PTG *\*mbi* (Ribeiro, 2010, p. 75; Sabino, 2016, p. 111).

<sup>35</sup> Root preserved in Old Tupí (Lemos Barbosa, 1951, p. 87), Parintintin (Betts, 1981), Ka'apor (Kakumasu & Kakumasu, 2007 [1988], p. 53), and other languages.

<sup>36</sup> In all Tupian languages except Makurap, this prefix has the allomorph *ō-* which occurs before consonant-initial stems. The loss of PTpr *\*m* is not known to be a regular sound change; it can nevertheless be identified as an innovation shared by the Core Tupian languages (Wayoró, Tuparí, Sakurabiat, Akuntsú).

Appendix 1

(Continue)

(44) PT \**mĩcõ* 'curassow'

PMG \**mĩcũ* > Mw *mĩñũ*, Aw *mĩtũk*, PTG \**mĩtũ* (Meira & Drude, 2015, p. 292) || PMu \**wĩtõ* (!) > Mu *witõ*, Ku *wito* (Picanço, 2019, p. 140) || PTpr \**mĩcõ* > Ma *mĩtõ* (Nikulin & Andrade, 2020, p. 305) || Kt *mbisĩ*, Ari *mĩsũ* (Landin, 2005, p. 16; Rondon & Faria, 1948, p. 198; Nimuendajú, 1932, p. 112, <mixum><sub>B</sub>, <misyũ, mityũ><sub>N</sub>)

(45) PT \**něcuk* 'horsefly'

(?) Mw *nasik* (!) (Ribeiro, 2010, p. 76) || PJu \**nătĩk-á* > Yu *natíká* (!), Xi *natika* (Fargetti, 2001, p. 295; Nimuendajú, 1928, p. 847; reconstruction ours) || PTpr \**něcuk* > Ma *nětík*, Tu *něuk* (Nikulin & Andrade, 2020, p. 306) || Pa *nig* 'gnat', Ar <*digá*> 'gnat, fly', Gv *dik* 'gnat', Zo *dig-akĩj* 'gnat', Sl *dik* (*kíp*) 'gnat' (Bontkes, 1978, p. 13; Sekelj, 1948; Galucio et al., 2015, p. 267)

(46) PT \**ngə* 'cultivated field'<sup>37</sup>

PMG \**ngə* > Mw *no/ko*, Aw *ko*, PTG \**ko* (Meira & Drude, 2015, p. 293)<sup>38</sup> || PMu \**kĩ* 'women's field' > Mu *ká*, Ku *kĩ* (Picanço, 2019, p. 142) || PJu \**kú-á* > Yu *kúá*, Xi *kua* (Fargetti & Rodrigues, 2008, p. 563; reconstruction ours) || PTpr \**ngə* > Ma/Wy *ngə* (Nikulin & Andrade, 2020, p. 308) || Kt *nga*, Ar *ngæ* (Landin, 2005, p. 9; Rondon & Faria, 1948, p. 200; Nimuendajú, 1932, p. 111, <gá><sub>B</sub>, <ngā><sub>N</sub>) || Pu <*tá*><sub>KG</sub> (Koch-Grünberg, 1932, p. 36) || Pa *na*, Ar <*ngá*>, Gv/Zo *ga* (Bontkes, 1978, p. 4; Sekelj, 1948; Moore, 1984, p. 18; Lacerda, 2014, p. 320)

(47) PT \**ngəat* 'sun'

PMG \**nguat* > Aw *kwat*, PTG \**kʷat*, \**kʷar[atsi]* (Sabino, 2016, p. 49; Rodrigues, 2007, p. 192; Mello, 2000, p. 176; reconstruction ours) || (?) PMu \**koat[ol]* 'summer' > Mu *koató*, Ku *koato* (Picanço, 2019, p. 143)<sup>39</sup> || PJu \**kuaz[adĩ]* > Yu *kuadĩ* (!), Xi *kuazadĩ*, Mn <*hayadí*> (Fargetti & Rodrigues, 2008, p. 563; Steinen, 1886, p. 361; reconstruction ours) || PTpr \**ngəat* 'sun, sky' > Ma *ngəat*, Wy *ngiat*, Tu *kiát* 'up' (Nogueira, 2011, p. 198; Nikulin & Andrade, 2020, p. 308) || Pa *ñád*, Ar/Gv/Zo/Sl *gát* (Bontkes, 1978, p. 4; Galucio et al., 2015, p. 252)

cf. PT \**at* 'day'

PMG \**a:t* > Mw *a:t* 'sun', PTG \**ar* (Ribeiro, 2010, p. 51; Mello, 2000, p. 155; reconstruction ours) || PTpr \**a:t* > Tu *a:t* 'sky' (Alves, 2004, p. 157; reconstruction ours) || Kt *o:t*, Ari *ot* 'midday', *e-oð-o* 'rain season' (Karitiana, 2016, p. 32; Nimuendajú, 1932, p. 110, <ë-áq̃><sub>B</sub>, <ad><sub>N</sub>) || Kr át 'day' (Gabas Jr., 1999, pp. 58, 60) || (?) Gv *ará-tígi* 'in ancient times' (Moore, 1984, p. 197).

(48) PT \**nguw* 'liquid'

PTpr \**ngi* 'liquid, saliva' > Wy *ngũ*, Tu *ku*, Sk/Ak *kĩ* (Nikulin & Andrade, 2020, p. 308) || Kt *ngə* 'blood' (Landin, 2005, p. 9)

(49) PT \**nguw* 'louse'

PMG \**ngip* > Mw *ngip*, Aw *ṇa-kíp*, PTG \**kíp* (Meira & Drude, 2015, p. 293) || PMu \**kíp* > Mu *kíp*, Ku *kíp* (Picanço, 2019, p. 141) || PJu \**kíp-á* > Yu *kíp-á*, Xi *kíp-a* (Fargetti & Rodrigues, 2008, p. 562; reconstruction ours) || PTpr \**ngip* > Ma *ngip*, Wy *ā-ngup*, Tu *kup*, Sk/Ak *kíp* (Nikulin & Andrade, 2020, p. 308) || Kt *ngəp*, Ari *ngəp-o* (Storto, 1999, p. 14; Niemundajú, 1932, p. 113, <gēēbā><sub>N</sub>) || Kr *nāp* (Galucio et al., 2015, p. 252) || Pu *típ* (Galucio, 2005, p. 165) || Pa *ñib*, Sl *gip*, (?) Ar *git* ~ <*ngid*>, Gv/Zo *git* (Galucio et al., 2015, p. 252; Sekelj, 1948)

(50) PT \**nguwat* 'salt'

PMG \**wu-kīt* (cf. \**wu* 'thorn') > Mw *ukīt*, Aw *tukīt*, PTG \**jukīt* (Meira & Drude, 2015, p. 294) || PTpr \**ngi?īt* > Ma *ngiīt*, Wy *ngũ:t* ~ *ku:t*, Tu *ku?út*, Sk *kĩ:t*, (?) Ak *kĩ:c* (Nikulin & Andrade, 2020, p. 308; Moore & Galucio, 1994, p. 134) || Pu *tir[a]* 'salt made of *aricuri* palm' (Monserrat, 2005, p. 18) || Ar <*git*> (Sekelj, 1948)

<sup>37</sup> Rodrigues (2005, p. 40) lists Kr *na-cej* as a member of this cognate set; however, *na-* is not a root but rather a prefix in Karo.

<sup>38</sup> Meira and Drude (2015, p. 293) attest Sateré-Mawé *ko* and reconstruct PMG \**ko*. While the relational stem *ko* is indeed used in Sateré-Mawé, there is also its absolute counterpart *ngo* (Franceschini, 1999, p. 29; Ribeiro, 2010, p. 56), which prompts the reconstruction of PMG \**ngə*.

<sup>39</sup> PMu \**katĩ* 'sun' (Mu *káfi* 'sun/moon', Ku *kadzĩ*) (Picanço, 2019, p. 142) is sometimes seen as a cognate of PTG \**kʷat* ~ \**kʷaratsi* and Yu *kuadĩ*, Xi *kuazadĩ* (Rodrigues, 2007, p. 192), but it might be better accountable for as an Arawakan borrowing (cf. Proto-Arawakan \**kečĩ* 'sun', as reconstructed in Payne (1991, p. 420).

## Appendix 1

(Continue)

- (51) PT *\*paβæk* 'to be visible'  
PMu *\*pápik* > Mu *papák* (Picanço, 2005, p. 17; reconstruction ours) || PTpr *\*[e]paβok* > Tu *epapók* 'to return' (Alves, 2004, p. 167; reconstruction ours)
- (52) PT *\*pepʔə* 'wing, flight feather'  
PMG *\*pepo* > Mw/Aw *pepo*, PTG *\*pepo* (Corrêa da Silva, 2010, p. 403; reconstruction ours) || PJu *\*pebu-a* > Yu *pewa*, Xi *séba* (Fargetti & Rodrigues, 2008, p. 560; reconstruction ours) || PTpr *\*pepʔo* > Wy *peo*, Tu *pépʔo*, Sk *pebo*, Ak *pébo* (Moore & Galucio, 1994, pp. 133, 135; Alves, 2004, p. 236; Aragon, 2008, pp. 67, 79; Galucio et al., 2015, p. 253) || Kt *papi* 'arrow feather, fin', Ari *pæpə* 'wing' (Landin, 2005, p. 23; Rondon & Faria, 1948, p. 198; Nimuendajú, 1932, p. 112, <ipapú><sub>B</sub>, <i-papú><sub>N</sub>)
- (53) PT *\*pətic* 'heavy'  
PMG *\*potic* > Mw *potic*, Aw *potic*, PTG *\*potsic* (Meira & Drude, 2015, p. 293) || PMu *\*pítʃi* > Mu *pojʃi*,<sup>40</sup> Ku *pidʒi* (Picanço, 2005, p. 267; Snethlage, 1932, p. 80, as <ipidy><sub>L</sub>; reconstruction ours) || PJu *\*padét-ú* > Yu *i-padétú*, Xi *padeTu* (Fargetti & Rodrigues, 2008, p. 562; reconstruction ours) || PTpr *\*poti* > Ma/Wy *poti*, Tu *posí*, Sk *potsi* (Moore & Galucio, 1994, p. 133; Alves, 2004, p. 241; reconstruction ours) || Kt *piti* (Landin, 2005, p. 27) || Kr *piʔti* (Gabas Jr., 1999, p. 15) || Pa *pati-ga*, Ar <patí ~ pati-f> 'heavy, thick', Gv *patil* (Bontkes, 1978, p. 15; Sekelj, 1948; Moore, 1984, p. 146)
- (54) PT *\*pi* 'inner part'  
PMG *\*pi* > Mw *pi-pe* 'inside', [ati] *pi* 'sky', *ɲaʔa-pi* 'forest', Aw *pi-wo* 'inside', PTG *\*pi* (abs. *\*mbi*)<sup>41</sup>, *\*pi-pe* 'inside' (Ribeiro, 2010, pp. 51, 56, 78, 88; Sabino, 2016, p. 122; Mello, 2000, p. 192; reconstruction ours) || PMu *\*pi* 'mouth' > Mu/Ku *pi* (Picanço, 2019, p. 136) || PTpr *\*-pi* > Ma *-pi* (as in *borá-pi* 'mouth'), Wy *-pi* 'inside', Tu *-(p)si*, Sk/Ak *-pi* (Nogueira, 2019, p. 171; Alves, 2004, p. 246; Nikulin & Andrade, 2020, pp. 292, 299) || Kt *-pi*, as in *ej-pi* 'earth' (Landin, 2005, p. 8) || Pa *pi* 'face', Ar <ambí> 'face', <gatpi> 'sky', Gv [á] *bi* 'face', *gar-pi* 'sky', Zo *gat-pi* 'sky' (Bontkes, 1978, p. 16; Moore, 1984, p. 149; Gavião, 2019, p. 73; Lacerda, 2014, p. 320)
- (55) PT *\*picja* 'heel'  
PMG *\*píca* > Mw *pia* (!), Aw *pi-[a]pita*, PTG *\*pita* (abs. *\*mbita*) (Meira & Drude, 2015, p. 290, fn. 24; reconstruction ours) || PMu *\*(?)iða* > Mu [wə] *jdá*, Ku *ila* (Picanço, 2019, p. 137) || Kt *pi:so*, Ari *piɔso* (!) (Landin, 2005, p. 24; Rondon & Faria, 1948, p. 193; Nimuendajú, 1932, p. 109, <pioçó><sub>B</sub>, <-pi-asá><sub>N</sub>) || Pu *bija[ka]* (Galucio, 2005, p. 178)
- (56) PT *\*picjo* 'breath'  
PMG *\*piçu* > Mw *pihu*, PTG *\*pitu* (Corrêa da Silva, 2010, p. 410; reconstruction ours) || PMu *\*piðo* > Mu *pído* (Picanço, 2005, p. 262; reconstruction ours), *\*ka(?)-bíðo* 'wind' > Mu *kə-bíðo*, Ku *ka-bilo[llo]* (Picanço, 2019, p. 143) || PTpr *\*pijo* > Wy *pi:to* (!), Tu *sio*, Mk *pito* 'pulmão', Ak *pitó* 'pulmão' (Nogueira, 2019, p. 152; Caspar, n.d.; Hanke et al., 1958, p. 209; Aragon, 2008, p. 27; reconstruction ours) || Kt [nó] *pisi* (Landin, 2005, p. 19)
- (57) PT *\*pitik* 'to take, to grab'  
PMG *\*pitik* > Mw/Aw *pitik*, PTG *\*pitsik* (Corrêa da Silva, 2010, p. 409; reconstruction ours) || PMu *\*itʃik* > Mu *ijʃik* 'to hold' (Gomes, 2006, p. 113; reconstruction ours) || PJu *\*pidík-ú* > Yu *pidikú*, Xi *padiku* (!) (Fargetti & Rodrigues, 2008, p. 562; Nimuendajú, 1929, p. 866; reconstruction ours) || Kt *pitik* 'to take out, to empty' (Rocha, 2011, p. 215)
- (58) PT *\*puuk* 'to burn'  
PMu *\*pik* > Mu/Ku *pik* (Picanço, 2019, p. 101) || PTpr *\*pik-* > (?) Ma *pik-* 'to extinguish', Wy *puuk-* 'to cook' (Nogueira, 2019, p. 21; reconstruction ours) || Kr *pæk* (Gabas Jr., 1999, p. 86)
- (59) PT *\*tə* 'to go (singular)'  
PMG *\*to* > Mw *to*, Aw *to*, PTG *\*tso* (Meira & Drude, 2015, p. 293) || PMu *\*tʃi* > Mu *tʃə*, Ku *tʃi* (Picanço, 2019, p. 139) || PJu *\*tʃa* > Yu *tʃa*, Xi *ta* (Fargetti & Rodrigues, 2008, p. 561; reconstruction ours) || (?) Kr *ti* (Gabas Jr., 1999, pp. 64, 136)

<sup>40</sup> The rounded vowel found in Mundurukú is irregular (*\*paʃi* would be expected). The Kuruaya form <ipidy><sub>L</sub>, attested by Antonio Lopes da Costa (Snethlage, 1932, p. 80), shows an unrounded vowel in the first syllable of the root; guided by the PMG cognate with *\*o*, we phonologize this Kuruaya form as */i-pidi/ i-pidʒi*.

<sup>41</sup> Root preserved in Old Tupí (Lemos Barbosa, 1951, p. 132), Parintintin (Betts, 1981), and other languages.

Appendix 1

(Continue)

(60) PT \**tə*- 'third person coreferential'

PMG \**to*- > Mw *to*-, (?) Aw *to*- 'reciprocal' (Franceschini, 1999, p. 24; Sabino, 2016, p. 131; reconstruction ours) || (?) Xi *d(u)*- (Rodrigues, 1995, p. 12) || PTpr \**te*- > Wy/Tu *te*-, Sk *tse*-, Ak *te*- (!) (Nogueira, 2019, p. 34; Alves, 2004, p. 257; Galucio, 2001, p. 74; Aragon, 2014, p. 219ff.; reconstruction ours) || Kt *ta*-, Ari *tæ*- (Storto, 1999, p. 125; Rondon & Faria, 1948, p. 196; Nimuendajú, 1932, pp. 114–116, <*ta*-><sub>B/N</sub>) || Kr *to*- (Gabas Jr., 1999, p. 81)

(61) PT \**tə*- 'pronominal formative (pl.)'

(a) PT \**oðe-tə* 'we (excl.)'

PMG \**urV-to* (!) > Mw *uru-to* (!) (Franceschini, 1999, p. 63; Ribeiro, 2010, p. 91) || PMu \**otfe-ji* > Mu *otfe-dʒá*, Ku *otfe-di* (Picango, 2019, p. 143)

(b) PT \**oc-tə* 'we (incl.)'

PMG \**uc-to* 'I' > Mw *uc-to*, Aw *ito* (!) (female speech) (Franceschini, 1999, p. 63; Ribeiro, 2010, p. 90; Sabino, 2016, p. 82) || PMu \**[w]ij-ji* > Mu *wəj-dʒá*, Ku *wei-di* (Picango, 2019, p. 143) || Kt *ij-tʃa*, Ari *uj-tæ* (Storto, 1999, pp. 22, 40, 206; Nimuendajú, 1932, p. 114, <*uita*><sub>N</sub>) || Kr *iʔ-tə* (Gabas Jr., 1999, p. 39, 81)

(c) PT \**ec-tə* 'you (pl.)'

PMu \**ej-ji* > Mu *ejdʒá*, Ku *eidʒi* (!) (Gomes, 2006, p. 24; Mendes Jr., 2007, p. 31; reconstruction ours) || Kt *aj-tʃa*, Ari *æj-tæ* (Storto, 1999, p. 22, 163; Nimuendajú, 1932, p. 114, <*aita*><sub>N</sub>)

(d) cf. also Mw *ac-to* 'we (incl.)' (Franceschini, 1999, p. 63; Ribeiro, 2010, p. 48); PTpr \**ki-te* 'we (incl.)' > Ma *ki-tě[nā]*, Sk *ki-tse*, Ak *ki-tʃé* (Braga, 2005, p. 88; Galucio, 2001, p. 38; Aragon, 2008, p. 53, 2014, p. 179); Ma *e[kij-tě[nā]* 'you (pl.)' (Braga, 2005, p. 88)

(62) PT \**tək* 'to pound, to grind'<sup>42</sup>

PMG \**tok* > Mw *tok*, Aw *[a]tok*, PTG \**tsok* (Ribeiro, 2010, p. 88; Sabino, 2016, p. 69; Mello, 2000, p. 199; Rodrigues & Dietrich, 1997, p. 274; reconstruction ours) || PMu \**tʃik* > Mu *tʃək* 'to break' (Picango, 2005, p. 36; reconstruction ours) || Pju \**(pá) dák-u* > Yu *pá-dák-u*, Xi *(pa-)dak-u* (Lima, 2008, p. 22; Nimuendajú, 1929, p. 878; reconstruction ours) || Kt *tak* (Rocha, 2011, p. 23) || (?) Pa *-tagá* 'to smash' (as in *to-dagá* 'to pound'), Gv *tágá* 'to beat' (Meer, 1982, p. 37; Gavião, 2019, p. 81)

(63) PT \**tək* 'larva'

PMG \**tok* > PTG \**tsok* (Mello, 2000, p. 200; PMG reconstruction ours) || Pju \**[a]dáká* > Yu *adáká* (Fargetti, 2001, p. 56; reconstruction ours) || PTpr \**tek* > Tu *tek* (Alves, 2004, p. 257; reconstruction ours) || Kt *[ken]tak* (Landin, 2005, p. 13) || Pa *[ka]dé:g* 'coconut larva', (?) *[mā:m]nég* 'caterpillar sp.', (?) *nég* 'boil', Ar <*mandeg*> 'caterpillar', Gv *mā-dék* 'lizard' (Meer, 1982, p. 37; Bontkes, 1978, pp. 8, 11, 13; Sekelj, 1948; Moore, 1984, p. 186)

(64) PT \**tī* 'mother'

PMG \**ndi/\*ti* > Mw/Aw *ni/ti*,<sup>43</sup> PTG \**tsi* (Meira & Drude, 2015, p. 294) || PMu \**tʃi* > Mu *tʃi*, Ku *dʒi* (Picango, 2019, p. 140) || Pju \**di-á* > Yu *dʒá*, Xi *diã* (Fargetti & Rodrigues, 2008, p. 562; reconstruction ours) || PTpr \**ti* > Ma/Wy *ti*, Tu *si*, Sk *tsi*, Ak *ti* (Nogueira et al., 2019, p. 39; reconstruction ours) || Kt/Ari *ti* (Storto, 1999, p. 17; Rondon & Faria, 1948, p. 195; Nimuendajú, 1932, p. 110, <*otí*><sub>B</sub>, <1 u-ti, 3 i-ti><sub>N</sub>) || Pa/Ar/Gv/Zo *ti* 'mother; maternal aunt', Sl *ti* 'mother' (Bontkes, 1978, p. 17; Felzke & Moore, 2019)

derived: PT \**tī-ʔur* ~ \**tī-ʔur* 'maternal aunt'

PMG \**tīʔir* > PTG \**tsiʔir* (Corrêa da Silva, 2010, p. 407; PMG reconstruction ours) || Kt *teʔet* (Landin, 2005, p. 30)

(65) PT \**tik* 'resin'

PMG \**-tik* > Mw *[i]tik-he*-, Aw *[tipa]tik*, PTG \**[ʔi]tsik* (Ribeiro, 2010, p. 99; Figueiredo, 2010, p. 226; Corrêa da Silva, 2010, 404; reconstruction ours) || Pju \**díká* 'wax' > Yu *[ka]díká* 'resin', Xi *dika* 'wax', *[ari]dika* 'resin' (Fargetti & Rodrigues, 2008, p. 562; Carvalho, 2019; reconstruction ours) || PTpr \**tik* > Tu *sik*, Sk *tsik* 'sap' (Alves, 2004, p. 247; Galucio, 2001, p. 91; reconstruction ours) || Ar <*tigi*> (Sekelj, 1948)

<sup>42</sup> Rodrigues (2007, p. 182, 2008, p. 6) includes Tuparí *tek* into this cognate set. However, other sources attest the Tuparí verb for 'to pound, to grind' as *ter*- (Alves, 2004, p. 258), which goes back to PTpr \**nder*- (Nikulin & Andrade, 2020, p. 306).

<sup>43</sup> The form *ni* (not given by Meira and Drude (2015)) is used both in Sateré-Mawé and in Awetí as the absolute equivalent of the relational stem *tī*. It is attested in Franceschini (1999, p. 29), Ribeiro (2010, p. 76) and Sabino (2016, p. 113).

## Appendix 1

(Continue)

## (66) PT \*tuur 'to cook'

PMG \*t<sub>i</sub>r 'to roast' > Aw [te]t<sub>i</sub>r, PTG \*te]ts<sub>i</sub>r, \*mi-ts<sub>i</sub>r 'roasted' (Sabino, 2016, p. 183; Mello, 2000, p. 162; Corrêa da Silva, 2010, p. 408; reconstruction ours) || PTpr \*t<sub>i</sub>r > Ma t<sub>i</sub>r 'to roast', Tu s<sub>ur</sub>- (Braga, 2005, p. 201; Nikulin & Andrade, 2020, p. 291, fn. 14) || Kt [o]tet (Rocha, 2011, p. 214) || Kr t<sub>i</sub>t (Gabas Jr., 1999, p. 105) || Pa t<sub>i</sub>(: )r 'hot, to cook', Ar/Gv/Zo t<sub>i</sub>r<sub>i</sub> 'to burn, to roast', Sl t<sub>i</sub>:t 'to burn' (Bontkes, 1978, pp. 8, 17; Moore, 2005, p. 520; Sekelj, 1948)

## (67) PT \*tuur 'digging stick'

PMG \*t<sub>i</sub>r > Mw t<sub>i</sub>r-pi, Aw t<sub>i</sub>r, PTG \*ts<sub>i</sub>r (Corrêa da Silva, 2010, p. 406; reconstruction ours) || PTpr \*t<sub>i</sub>r > Ma t<sub>i</sub>r 'spade', Wy t<sub>ur</sub> 'machete', Tu s<sub>ur</sub> (Braga, 2005, p. 201; Nogueira, 2019, p. 56; Rodrigues, 2002, p. 290; reconstruction ours)

## (68) PT \*uβicjo 'wind'

PMG \*iβicū > Mw iwitu, Aw iwii (!), PTG \*iβitu (Ribeiro, 2010, p. 98; Sabino, 2016, p. 102; Meira & Drude, 2015, p. 295; reconstruction ours) || PTpr \*iβijo > Tu uβiō (Nikulin & Andrade, 2020, p. 292)

## (69) PT \*ũp 'dirty, black'

PMG \*-i<sub>p</sub> > Mw wār-[t]i<sub>p</sub> 'night', (?) Aw [tatik]i<sub>p</sub> 'dark' (Ribeiro, 2010, p. 93; Sabino, 2016, p. 24; reconstruction ours) || PTpr \*i<sub>p</sub> > Tu ũp (Alves, 2004, p. 271; reconstruction ours) || Kt ē:m, Ari ēm- (Storto, 1999, p. 15; Nimuendajú, 1932, p. 114, <i-ēmā><sub>N</sub>); Ari pæ-ēm(-) 'Black person' (Rondon & Faria, 1948, p. 196; Nimuendajú, 1932, p. 110, <tapa-ēme><sub>B</sub>, <i-paēmā><sub>N</sub>)

## (70) PT \*upek 'duck'

PMG \*ipek > Aw ipek, PTG \*ipek (Corrêa da Silva, 2010, p. 156; reconstruction ours) || PTpr \*ipek > Wy upek, Tu (?)u<sub>p</sub>ék, Sk ipek, Ak ipék (Moore & Galucio, 1994, p. 133; Alves, 2004, p. 277; Hanke et al., 1958, p. 200; Aragon, 2014, p. 78) || (?) Pa i:péja, Ar <ipei>, Zo ipeja (Bontkes, 1978, p. 7; Sekelj, 1948; Lacerda, 2014, p. 321)

## (71) PT \*utuū 'deer'

PMG \*iti: > Mw iti:, (?) Aw ti[wapa] (Meira & Drude, 2015, p. 292) || PMu \*i<sub>t</sub>i > Ku i<sub>d</sub>í (Costa, 2002, p. 99; reconstruction ours) || PTpr \*iti: > Ma iti:, Wy iti:, Sk itsi:, Ak iti (Galucio et al., 2015, p. 266; Aragon, 2014, p. 148; reconstruction ours) || (?) Kt/Ari nde (!) (Landin, 2005, p. 18; Rondon & Faria, 1948, p. 198; Nimuendajú, 1932, p. 112, <dé, de><sub>B</sub>, <ndē><sub>N</sub>) || Kr iti (Gabas Jr., 1999, p. 10) || Pu idi (Galucio, 2005, p. 175; Monserrat, 2005, pp. 15–16) || Pa i<sub>t</sub>i[áb] ~ i<sub>t</sub>i:[áb], Ar i<sub>t</sub>i ~ itií, Gv iitií, Zo itii (Bontkes, 1978, p. 8; Galucio et al., 2015, p. 266)

## (72) PT \*wākūjā 'agouti'

PTpr \*wākīñā > Ma mākīñā, Wy η\*wākīñā, Sk mākīñā/pakīñā (dialectal variants), Ak pakāñā (Nikulin & Andrade, 2020, p. 295) || Kr wakāja (Gabas Jr., 1999, p. 13) || Pu wakīñā (Galucio, 2005, p. 182; Monserrat, 2005, p. 16) || Pa wakī(:), Ar βakīí, Gv βaakīí, Zo βakīí, Sl βakii (Bontkes, 1978, p. 18; Galucio et al., 2015, p. 262)

## (73) PT \*wati 'moon'

PMG \*wa:ti > Mw wa:ti, Aw tati, PTG \*jatsi (Meira & Drude, 2015, p. 294) || PMu \*wa<sub>t</sub>i > Mu wafí,<sup>44</sup> Ku wad<sub>z</sub>i (Santos, 2013, p. 71; Nimuendajú, 1932, p. 107; Costa, 2002, p. 99; reconstruction ours) || Kt oti, Ari oti (Storto, 1999, p. 14; Rondon & Faria, 1948, p. 194; Nimuendajú, 1932, p. 109, <oti><sub>B</sub>, <oti><sub>N</sub>) || Pa gāti-kád (male speech) (!), Ar gāti, gát-ti (!), Gv/Zo gāti (!) (Bontkes, 1978, p. 4; Galucio et al., 2015, p. 258)

## (74) PT \*wetik 'sweet potato'

PMG \*wećik (!) > Aw tezik, PTG \*jetik (Corrêa da Silva, 2010, p. 405; reconstruction ours) || PMu \*we<sub>t</sub>ik > Mu we<sub>f</sub>ik-, Ku wed<sub>z</sub>ik (Picanço, 2019, p. 136) || Pu <witiká><sub>KG</sub>, <widiká><sub>ES</sub> (Koch-Grünberg, 1932, p. 44; Snethlage, 1934) || Pa wa<sub>t</sub>i(:)η-a (!), Ar βitīñ-ṛa, Gv βitig-ā, Zo βetīñ-āa fīp (!), Sl tīñ-ṛa (Bontkes, 1978, p. 18; Galucio et al., 2015, p. 272)

<sup>44</sup> This form is documented only in the Madeira dialect of Mundurukú. In the Tapajós dialect, this term has been replaced with káfí 'sun/moon' (Picanço, 2005, p. 27). Interestingly, the form <wasuptá><sub>N</sub> 'star', attested for the Madeira dialect in Nimuendajú (1932, p. 107), has been likewise replaced with kasoptā (Picanço, 2005, p. 128).

Appendix 1

(Conclusion)

(75) PT \**wi* 'ax'

PMG \**wi* > Mw *wi-ʔiʔ* 'handle', Aw *ki*, PTG \**ʔi* (Ribeiro, 2010, p. 95; Meira & Drude, 2015, p. 292) || PTpr \**wi* > Ma *βi*, Tu *wi(:)*, Sk/Ak *kʷi* (Nikulin & Andrade, 2020, p. 299) || Kt *hi* 'knife (dated)', Ari *hi* (Landin, 2005, p. 11; Rondon & Faria, 1948, p. 195; Nimuendajú, 1932, p. 111, <hi><sub>B</sub>, <hi-><sub>N</sub>) || Pu *wi-a* (Galucio, 2005, p. 181; Monserrat, 2005, p. 19)

(76) PT \**wica* 'stone'

PMG \**wica* > Aw *kita*, PTG \**ita* (Corrêa da Silva, 2010, p. 404; reconstruction ours) || PMu \**wita* 'grater' > Mu *witá* (Picanço, 2005, p. 43, fn. 6; reconstruction ours); PMu \**wita-ʔa* 'stone' > Mu *witá-ʔa*, Ku *wita-ʔa* (Picanço, 2019, p. 141) || Kt *i(t)so* 'whetstone', Ari *isp(-o)* (Landin, 2005, p. 12; Rondon & Faria, 1948, p. 194; Nimuendajú, 1932, p. 110, <iʔó><sub>B</sub>, <isoá, isaá><sub>N</sub>) || Kr *ijá* (Gabas Jr., 1999, p. 182) || (?) Pa *ifáa*, Ar *ʔaʔa*, Gv *ifjáa*, Zo *ifjáa*, Sl *ʔaʔa ~ isáʔa* (!) (Bontkes, 1978, p. 4; Galucio et al., 2015, p. 259)

(77) PT \**wip* 'cooked'

PMG \**wip* 'cooked' > Mw *wip*, Aw *[o]kip*, PTG \**tsip* (Meira & Drude, 2015, p. 292) || PMu \**wip* > Mu *wip* (Gomes, 2006, p. 259, fn. 13; reconstruction ours) || Pju \**[u]wipu* > (?) Yú *uh-u* (!) 'be roasted', Xi *uwih-u* (Fargetti & Rodrigues, 2008, p. 561; Nimuendajú, 1929, p. 869; reconstruction ours) || Kt *hip* 'to roast', (?) Ari *mbuð-o-hip-op-o* 'frying pan' (Rocha 2011, p. 33; Nimuendajú, 1932, p. 111, <burohíppa><sub>N</sub>) || Kr *wip* (Gabas Jr., 1999, p. 17) || (?) Gv *βîl-p* (Moore, 1984, p. 261)

(78) PT \**wuu* 'to blow'

PTpr \**wi* > Ma *βi*, Tu *u* 'play a blowing instrument' (Nikulin & Andrade, 2020, p. 299) || Kt *he:* (Rocha, 2011, p. 33)

(79) PT \**wucjã* 'shell'

PMG \**wicã* > Aw *kitã*, PTG \**itã* (Corrêa da Silva, 2010, p. 404; reconstruction ours) || PTpr \**wijnã* > Tu *unã* (Rodrigues, 2002, p. 290; reconstruction ours) || Ari *mũsõ* (Rondon & Faria, 1948, p. 199, <mixon><sub>B</sub>)

(80) PT \**ʔək* 'tuber'

PMG \**ʔək* > Aw *ʔok*, PTG \**ʔok* (Meira & Drude, 2015, p. 293) || PMu \**ʔik* 'belly' > Mu *ʔək*, Ku *ʔik* (Picanço, 2019, p. 136) || PTpr \**ek* > Tu *ek* (Alves, 2004, p. 164; reconstruction ours)

derived: PT \**mānĩ-ʔək* 'manioc'

PMG \**mānĩʔək* > Aw *maniʔok*, PTG \**māndiʔok* (Meira & Drude, 2015, p. 293) || PMu \**másik* > Mu *məsák*, Ku *másik* (Picanço, 2019, p. 140) || Pju \**maniaká* 'manioc' > Yú *maniáká ~ majáká*, Xi *maniaka ~ majaka* (Steinen, 1886, p. 363; Nimuendajú, 1928, p. 833, 1932, p. 585; Fargetti & Rodrigues, 2008, p. 562; reconstruction ours)

(81) PT \**ʔu* 'water'

PMG \**ʔi* > Mw *i[ʔi]*, Aw *ʔi*, PTG \**ʔi* (Meira & Drude, 2015, p. 295) || PMu \*(*ʔ*)*i-ði* > Mu *idj-bí*, Ku *idji* (!) (Picanço, 2019, p. 135) || Pju \**i-á* > Yú *ijá*, Xi *ija* (Fargetti & Rodrigues, 2008, p. 560; reconstruction ours) || PTpr \**ʔi* > Ma *i:*, Wy *u-gu*, Tu *ʔu* 'drink, chicha', Sk *i-kí*, Ak *i-kí* (Braga, 2005, p. 201; Nogueira, 2019, p. 8; Alves, 2004, p. 276; Galucio, 2001, pp. 42, 70, 77; Aragon, 2014, p. 78; reconstruction ours) || Kt/Ari *e* 'rain' (Landin, 2005, p. 7; Rondon & Faria, 1948, p. 194; Nimuendajú, 1932, p. 110, <hé><sub>B</sub>, <é><sub>N</sub>); Kt/Ari *e-se* 'water, river' (Storto, 1999, p. 138; Rondon & Faria, 1948, p. 194; Nimuendajú, 1932, p. 110, <exê, ecê><sub>B</sub>, <esê><sub>N</sub>) || Kr *i-ci* (Gabas Jr., 1999, p. 10) || (?) Pu *ifja* 'creek', *i[wic]* 'big river' (Monserrat, 2005, p. 17) || Pa *i:*, Ar *ii*, Gv *ii-pâàp* 'water' (*ii* 'river, chicha'), Zo *ii* (Bontkes, 1978, p. 21; Galucio et al., 2015, p. 258; Moore, 1984, pp. 167, 245; Lacerda, 2014, p. 320)

derived: PT \**ʔu-ʔa* 'gourd'

PMG \**ʔa* > Aw *ʔa*, PTG \**ʔa* (Corrêa da Silva, 2010, p. 405) || PTpr \**ʔiʔa* > Ma *ia[pe]*, Wy *ʔuβa* 'traditional recipient for storing', Tu *úʔa* 'recipient for storing chicha' (Braga, 2005, p. 201; Nogueira, 2019, p. 11, 39; Alves, 2004, p. 272; reconstruction ours) || (?) Gv *iʔá-àp* 'calabash' (Rodrigues, 2007, p. 187)<sup>45</sup>

<sup>45</sup> We were unable to locate this form in any primary source on Gavião (2019). However, it corresponds well to the data of other Tupian languages and thus we provisionally leave it here, given that Rodrigues (2007) may have had access to unpublished Gavião (2019) data. An anonymous reviewer remarks, however, that the Gavião (2019) form in question is composed of *i* 'chicha' and *-áàp* 'convex or concave object' and, thus, might not be cognate with the remaining forms.

## Appendix 2. PT consonants.

PT	PMG	PMu	PJu	PTpr	Kt/Ari	Kr	Pu	Mo
*p	*p	*p/*b <sup>A</sup> /*ɔ <sup>B</sup>	*p/*b <sup>A</sup>	*p	p	p	b/b ~ f <sup>C</sup>	p
*[mb] *m <sup>E</sup>	*[mb] *m <sup>E</sup>	*p *m <sup>E</sup>	*p *m <sup>E</sup>	*[mb] *m <sup>E</sup>	[mb] m <sup>E</sup>	m	m	b, Pa m m <sup>E</sup>
*β	*β	?	?	*β	?	?	?	?
*w	*w	*w/*ɔ <sup>A</sup>	*w	*w	h/ɔ <sup>D</sup> /m <sup>E</sup>	w	w	β, Pa w
*t	*t, *t <sup>C</sup>	*tʃ/*j <sup>A</sup> /tʃ <sup>C</sup>	*tʃ/*d <sup>A</sup>	*t	t, Kt tʃ <sup>F</sup>	t	d/n <sup>E</sup>	t
*[nd] *n <sup>E</sup>	*[nd]/*ð <sup>A</sup> *n <sup>E</sup>	*tʃ ?	*tʃ *n <sup>E</sup>	*[nd] *n <sup>E</sup>	*[nd] *n <sup>E</sup>	*[nd] *n <sup>E</sup>	?	d, Pa n ?
*ð	*ð	*tʃ	*β	*t	t	?	?	(?) t
*r	*r	*r	*r	*r	r	r	r	l, Pa r
*tʃ	*tʃ	*f	(?) *f	?	(?) s	?	?	?
*c	*c	*t	*t	*c	s	j	?	∅
*j	*c, *c <sup>G</sup>	*ð	*s ~ f	*j/*[ŋ] <sup>E</sup>	s/n <sup>E</sup>	j ~ c	j ~ t ~ f	dz ~ ts/j <sup>E</sup> /tʃ <sup>C</sup>
*k	*k	*k	*k	*k	k	k		k
*kʲ	*k	*k, *f <sup>H</sup>	*k	*ʔ ~ *∅	k	k	k	k
*k̥	*ʔ	*ʔ ~ *∅	∅	*k	?	?	?	∅
*[ŋg] *ŋ <sup>E</sup>	*[ŋg] *ŋ <sup>E</sup>	*k	*k *n <sup>E</sup>	*[ŋg]	[ŋg] n	n	t ?	g, Pa ŋ n
*ʔ	*ʔ	*ʔ ~ *∅	∅	*ʔ	?	*ʔ ~ *∅	*ʔ ~ *∅	∅

<sup>A</sup> = between vowels; <sup>B</sup> = before or flanked by front vowels; <sup>C</sup> = before \*j; <sup>D</sup> = in non-final syllables or between front vowels; <sup>E</sup> = before nasal nuclei; <sup>F</sup> = after j; <sup>G</sup> = after \*i or \*c (PT \*cj > PMG \*c̣); <sup>H</sup> = before PT front or central vowels.

PT	PMG	PMu	PJu	PTpr	Kt	Kr	Pu	Mo
*p	*p	*p/*m <sup>A</sup>	*p/*m <sup>A</sup>	*p	p/m <sup>A</sup>	p/m <sup>A</sup>	p	p, Pa b
*t	*t	*t/*n <sup>A</sup>	*β/*n <sup>A</sup>	*t	t/n <sup>A</sup>	t/n <sup>A</sup>	t	t, Pa d
*c	*c	*j/*ɔ <sup>B</sup>	*t/*j <sup>A</sup>	*c/*ɔ <sup>B</sup>	j/n <sup>A</sup> /ɔ <sup>B</sup>	j/ɔ <sup>B</sup>	c	j
*k	*k	*k/*ŋ <sup>A</sup>	*k/*ɔ <sup>A</sup>	*k	k/ŋ <sup>A</sup>	k/ŋ <sup>A</sup>	k	k, Pa g

<sup>A</sup> = after nasal nuclei; <sup>B</sup> = after \*j.

