

Proto-Tupi-Guarani had no palatalized velar stop

O Proto-Tupi-Guarani não tinha uma oclusiva velar palatalizada

Fernando Carvalho 

Museu Nacional. Universidade Federal do Rio de Janeiro (MN/UFRJ). Rio de Janeiro, Rio de Janeiro, Brasil

Abstract: This paper addresses one of the open issues in the reconstruction of Proto-Tupi-Guarani (PTG) segmental phonology: The status of the **k* -**kʲ* opposition. We argue that the contrast is artifactual and that the presumed evidence in favor of PTG **kʲ* can be considered as secondary developments of PTG **k* in Kayabí, Guarayu, Kagwahiva, Tenetehára, Kamayurá, and Ka'apor. We establish additional facts regarding the structure of PTG and the historical phonology of TG languages, also showing that this finding eliminates the need for an unmotivated split in Pre-PTG history, a problematic feature of current reconstructions of the Proto-Tupian consonant system.

Keywords: Comparative method. Tupi-Guarani languages. Historical phonology.

Resumo: Este trabalho tem como objetivo resolver uma das questões em aberto acerca da reconstrução da fonologia do Proto-Tupi-Guarani (PTG): a da existência ou não de um contraste entre uma oclusiva velar simples **k* e uma oclusiva velar palatalizada **kʲ*. Argumentamos que a evidência que supostamente indicaria a necessidade de reconstruir **kʲ* é mais bem explicada por meio de desenvolvimentos secundários de **k* em algumas línguas, como o Kayabí, o Kamayurá, o Tenetehára, o Guarayu, o Kagwahiva e o Ka'apor. A análise das correspondências relevantes também estabelece uma série de outros fatos acerca da estrutura do PTG e da fonologia histórica dessas línguas, além de apresentar uma avaliação crítica de algumas das etimologias tradicionalmente tidas como relevantes para a questão do estatuto do contraste **k* -**kʲ*. Por fim, mostramos que a reconstrução do PTG com **k* apenas elimina a necessidade para uma cisão não motivada em nível do Pré-PTG, uma característica problemática de propostas existentes acerca das consoantes do Proto-Tupi. Um apêndice apresenta o conjunto de etimologias utilizadas como dados para a análise apresentada.

Palavras-chave: Método comparativo. Línguas Tupi-Guarani. Fonologia histórica.

Carvalho, F. (2023). Proto-Tupi-Guarani did not have a palatalized velar stop. *Boletim do Museu Paraense Emílio Goeldi. Ciências Humanas*, 18(1), e20220013. doi: 10.1590/2178-2547-BGOELDI-2022-0013

Autor para correspondência: Fernando O. de Carvalho. Museu Nacional, Quinta da Boa Vista, São Cristóvão. Rio de Janeiro, RJ, Brasil. CEP 20940-040 (fernaoorphao@mn.ufrj.br).

Recebido em 03/03/2022

Aprovado em 06/09/2022

Responsabilidade editorial: Adam Singerman



INTRODUCTION

The goal of this paper is to resolve one of the still open issues in the phonological reconstruction of Proto-Tupi-Guarani (PTG), the shared ancestor of the largest branch of the Tupian language family. I will show that the palatalized velar stop **kʲ* (whose *status* has been recently called into question; Meira & Drude, 2015, p. 282, fn. 7), can be eliminated from the reconstructed PTG inventory, and the relevant correspondences can be more insightfully analyzed as the result of language specific developments in Kayabí, Kagwahiva, Tenetehára, Guarayu, Kamayurá, and Ka'apor. The paper is organized as follows: After a presentation of the current standing of this question ('the current view'), I will discuss the segmental correspondences in a representative sample of ten, well-attested TG languages, based on which a PTG plain velar stop **k* can be straightforwardly reconstructed. Next, I will show that overlapping correspondences with diverging reflexes in a subset of these languages can be accounted for by invoking language-specific developments of the same PTG **k*, with no need for an independent and contrasting PTG velar stop ('PTG **k* and its reflexes'). All the relevant correspondences have been extracted from cognate sets that appear in the Appendix to the paper. In the section entitled 'some implications' I briefly discuss how this finding eliminates the need to postulate an unmotivated split of Proto-Tupian ***kʲ* into PTG **k* and **kʲ*. Finally, the section 'conclusions' is devoted to the synthetic presentation of the findings in the paper.

THE CURRENT VIEW

In her overview of the then current understanding of the Tupi-Guarani language family, Jensen (1999, p. 139) notes that PTG **kʲ* is reconstructed for three morphemes: **ikié* 'to enter', **kʲér* 'to sleep' and **kʲé* 'here, near speaker'. According to her, the change in the reconstructed forms - previously uniformly reconstructed with **k* - was deemed necessary to account for the Guarayu form **kie* 'sleep' in Hoeller's (1932) data. The contrast between **kʲ* and **k* is reconstructed for PTG by Mello (2000) and by Rodrigues (2007). Mello (2000) reconstructs **kʲ* in **-kʲer* 'sleep' only, while Rodrigues (2007) has **kʲ* in **-kʲer* 'sleep' and **-ejkʲe* 'enter'. Cognate sets in languages other than Guarayu would presumably support this proto contrast, such as Kayabí *set* 'to sleep' and *se* 'to enter', and the change **e > i* in this context in Parintintin: *kir* 'sleep' and *ki* 'here'. Meira and Drude (2015, p. 281), in a paper focused on the comparison between PTG and its two closest relatives, Awetí and Mawé, note that **kʲ* has an uncertain *status* at the PTG level, being reconstructed only preceding **e* in works such as Mello (2000) and Rodrigues (2007). The authors offer a convenient summary of the *status* of the phonological problem:

Mello has only four cases of PTG **ke*: **kerap* 'to close', **keramu* 'to snore', **purake* 'electric eel' and **ukeʔi* (doubtful) 'sister/brother-in-law' (the latter apparently related to Man's Older Brother). Mello claims that **k* and **kʲ* have different reflexes in Siriono, Apiaka, Kayabí, Urubu-Kaapor and (sometimes) Tembe, but, in his data, (a) these languages are all missing in the sets for **kerap* and **ukeʔi*; (b) only Siriono occurs in the **keramu* cognate set, where it has the same reflex (*kenāmu* with *k*) as in **kʲet* (> *ke*, also with *k*); and (c) in **purake*, Tembe and Urubu-Kaapor both occur with *k* (*murake*, *purake*), while in **kʲet* only the Urubu-Kaapor reflex is different (*jer* with *j*), while the Tembe reflex is simply *ker*, with the same *k* as in **purake*. There is thus almost no evidence in Mello (2000) to support a distinction between PTG **ke* and **kie* (Meira & Drude, 2015, p. 282, fn. 7).

The situation is, in fact, more difficult for the proponents of this contrast than the Meira and Drude (2015) quote above suggests. First, note that the number of supporting etymologies falls from four to three, once it is recognized that the PTG etymon meaning 'to snore', Mello's (2000, p. 172) **keramu* 'roncar' [to snore], is not independent from the etymon **kʲer* 'dormir' [to sleep] (Mello, 2000, p. 176), but is likely a reflex of the derivative **ket-amu* 'to snore (while) sleeping', as shown by Old Guarani *aquerambu* 'roncar' [to snore], *ambu* 'ronquido' [snoring sound], *tayaçu apiĩmbu*



'de puerco' [pig's snoring sound] (Restivo, 1893 [1722], p. 482), and Old Tupi *Xequerambû* 'roncar, o que dorme' [to snore, he/she who sleeps], *Xeambû* 'roncar o porco' [to snore, the pig] (Drumond, 1952, p. 108)¹. As suggested below, the 'sleep' and 'snore' sets were phonologically segregated in Mello's (2000) reconstruction only because his set for **keramu* 'to snore' fails to include cognates from some languages such as Kayabí, which, in his view, are critical for reconstructing **k'*, while these same languages do contribute witnesses to the 'sleep' set. Second, **kenab* 'fechar' [to close] (Mello, 2000, p. 172) is very doubtful and not clearly reconstructible for PTG (see next section). Third, the supposed Ka'apor reflex *fer* 'to sleep' is a non-existent ghost form (more on this below). Fourth, as suggested by Meira and Drude (2015) and demonstrated in the remaining of this paper, the other sets are not problematic at all, pointing to language-specific developments and not to the independent reflexation of a separate PTG segment.

Before proceeding, however, I would like to highlight a generalization about the sound structure of PTG that has not been so far explicitly commented upon, but which is relevant for the evaluation of the issue at hand. This generalization will also provide a background for the synthesis of the current understanding of the putative contrast between **k* and **k'*.

An examination of all extant proposals on the reconstruction of PTG etyma (Lemle, 1971; Schleicher, 1998; Mello, 2000) reveals that the sequence **ki* is not reconstructed, as shown in Table 1, where examples for each of the reconstructed sequences **kv* (where *v* = any vowel) are given for each PTG source².

Table 1. Vocalic contexts for PTG **k* in published comparative reconstructions.

	<i>*ka</i>	<i>*ke</i>	<i>*ki</i>	<i>*ko</i>	<i>*ku</i>	<i>*ki</i>
Lemle (1971)	<i>*kab</i> 'fat'	<i>*ker</i> 'sleep'	-	<i>*moköy</i> 'two'	<i>*puku</i> 'long'	<i>*kitĩ</i> 'cut'
Schleicher (1998)	<i>*-akaŋ</i> 'head'	<i>*kér</i> 'sleep'	-	<i>*ekó</i> 'to live'	<i>*-akúβ</i> 'hot'	<i>*-akim</i> 'wet'
Mello (2000)	<i>*karu</i> 'to eat'	<i>*ike</i> 'to enter'	-	<i>*ko</i> 'garden'	<i>*ku?a</i> 'waist'	<i>*kiβ</i> 'louse'

As noted by Meira and Drude (2015), this putative contrast between PTG **k* and **k'* is attested only in the context of a following **e*, which strongly suggests that this palatalization is a secondary effect of the contextual front vowel, the only PTG front vowel that was found in this context. The PTG etyma in Rodrigues and Dietrich (1997, pp. 273-274) which exemplify the contrast are: **ike?ir* 'brother of man, younger', **ike* 'side of the body' vs. **kier* 'sleep', **ekie* 'go in'. Mello (2000, pp. 163, 172, 176, 184, 191) gives only **kier* 'sleep' for PTG **k'*, as opposed to **k* in the same context (that is, preceding **e*) in the forms for: **keramu* 'to snore', **oken* 'door', **ike* 'to enter', **kenab* 'to close', **purake* 'electric eel' and **uke?i* 'brother/sister-in-law'. Other studies give only **k*, as in Lemle (1971) and Schleicher (1998), where **ker* 'sleep' is the only case of a **ke* sequence. Jensen (1999, p. 139) presents **ikie* 'enter (to)', **kier* 'sleep (to)' and **kie* 'here, near the speaker', as putative examples of PTG **k'* but does not discuss explicitly the existence of contrasts.

¹ The "Vocabulário na Língua Brasília", or VLB, is arguably the main lexical source on the Old Tupi language. While the manuscript is dated to 1621, different lines of evidence suggest an earlier date for its original composition, perhaps as early as the mid 16th century (see Lemos Barbosa, 1948). I have used here the 1952 edition by Carlos Drumond.

² Although PTG reconstructed forms appear in a number of different works (such as Dietrich, 1990; Rodrigues & Dietrich, 1997; Rodrigues, 2007), this table includes forms from studies where the evidence for reconstructed etyma (cognate sets) is presented. Jensen (1984), although an important study, relies essentially on the reconstructions of Lemle (1971).

PTG *k AND ITS REFLEXES

We will employ a sample of TG languages for addressing this specific aspect of PTG sound structure. The set of languages compared, given below in Table 2, includes languages for which relatively significant documentation is available, and which comprehensively represent the internal diversity of the family, as indicated by their classification within the two major, extant proposals on the internal classification of TG languages (that of Rodrigues, 1984/1985, later updated in Rodrigues & Cabral, 2002, and that of Michael et al., 2015)³.

Table 2. Position of languages used for reconstruction in each of the existing internal classifications.

Code	Language	Rodrigues (1984/1985)	Michael et al. (2015)
OGU	Old Guarani	I	Southern
GUY	Guarayu	II	Non-Guaranian southern
KAA	Ka'apor	VIII	Nuclear TG ⁴
KAG	Kagwahiva	VI	Non-diasporic peripheral
KAM	Kamayurá	VII	Sister branch to nuclear TG
KAY	Kayabí	V (later: VI)	Non-diasporic peripheral
TEN	Tenetehára	IV	Non-southern diasporic
TOC	Tocantins Asurini	IV	Central
TUP	Old Tupi	III	Non-southern diasporic
WAJ	Wajãpi	VIII	Non-diasporic peripheral

The relevant correspondences, identified for the cognate sets featuring in the Appendix to this paper, are given in (1) below. Each correspondence is followed by the semantic glosses that identify the cognate sets featuring the correspondence in question⁵.

(1) Segmental correspondences

(I) TUP *k* : TEN *k* : TOC *k* : KAM *k* : WAJ *k* : KAY *k* : GUY *k* : OGU *k* : KAA *k* : KAG *k*

KNIFE; CUT; PIERCE; LONG; KILL; WOODS; LOOK FOR; PULL; FAT; BÈSTAY; HOT; GOOD; WET; BRANCH; LOUSE; CAYMAN; MONKEY; SALT; SWALLOW; KNOW; DIG; BURN (INTR.); TWO; WOMAN; ELDER BROTHER; HUSBAND'S SISTER

(II) TUP *k* : TEN *k* : TOC *k* : KAM *k* : WAJ *k* : KAY *ɣ* : GUY *kj* : OGU *k* : KAA *k* : KAG *k*

SLEEP; SIDE OF THE BODY

(III) TUP *k* : TEN *ɣ* : TOC *k* : KAM *ts* : WAJ *k* : KAY *s* : GUY *kj* : OGU *k* : KAA *ɟ* : KAG *k*

ENTER

³ A third alternative classification is that of Gerardi and Reichert (2021). In terms of the proposed subgroups it does not differ much from the other two, in particular for the lower level clades. The main difference concerns the position of Old Tupi, which appears as 'non-southern', or Amazonian TG language in the Gerardi and Reichert (2021) proposal.

⁴ The clade that contains Ka'apor (along with Guajá and Avá-Canoeiro) in the Michael et al. (2015) classification is unnamed.

⁵ Note that to limit the discussion to the issue at hand, I have only included correspondence sets for PTG *k in syllable onset position, either in morpheme/word-initial position, or in intervocalic position. PTG admits word-final codas, and *-k is frequently found in this position, though the putative palatalized segment *kʲ has never been reconstructed in this position. I am also not considering the reflexes of PTG *kʷ, which is well-supported.

- (IV) TUP *k* : TEN *k* : TOC *k* : KAM *k* : WAJ *k* : KAY *k* : GUY *k* : OGU *k* : KAA *ʃk* : KAG *k*
 WAIST; BREAST; BONE; BACK; DIRTY

The correspondence in (I) is the main (identity) correspondence that establishes PTG **k*. The correspondences in (II) and (III) are the two correspondences that have been accounted for by postulating a separate PTG segment **kʲ*. These correspondences are not only attested in fewer sets than is the case with (I), but as noted above, also happen to be contextually very limited, and occur in contexts that are complementary to those of the identity correspondence (I). The identity correspondence for **k* is attested in a variety of vocalic contexts: initially preceding *u* (WOMAN) and *i* (KNIFE); medially between *e_ũ* (TONGUE), *a_u* (HOT), *u_u* (LONG), *a_a* (CAYMAN), *a_ã* (HEAD), *e_a* (LOOK FOR), *o_õ* (SWALLOW), *a_ĩ* (WET), *u_a* (KILL), *u_i* (SALT) and (*i*)_o (DIG). The three etymologies that support correspondences (II) and (III) show the presumed reflexes of PTG **k* in a single context: that of a following **e*, as shown in Table 3, where the most important reflexes are highlighted by cell shading.

Table 3. Cognate sets instantiating correspondences II and III.

PTG (past work)	<i>*ike</i> 'Side of the body'	<i>*-kʲer</i> 'sleep'	<i>*-ekʲe</i> 'enter'
PTG (this paper)	<i>*ike</i>	<i>*-ker</i>	<i>*-ike</i>
TUP	<i>ike</i>	<i>-ker</i>	<i>-ike</i>
TEN	<i>ike</i>	<i>-ker</i>	<i>-itʃe</i>
TOC	-	<i>-ken</i>	<i>-ke</i>
KAM	<i>ike</i>	<i>-ket</i>	<i>-itse</i>
WAJ	<i>ike</i>	<i>-ke</i>	<i>-ike</i>
KAY	<i>ise</i>	<i>-set</i>	<i>-se</i>
GUY	<i>ikje</i>	<i>-kje</i>	<i>-ikje</i>
OGU	<i>ike</i>	<i>-ke</i>	<i>-ike, -ikje</i>
KAA	<i>rake (?)</i>	<i>-ker</i>	<i>ije</i>
KAG	<i>iki</i>	<i>-kir</i>	<i>-ki ~ -eki</i>

Correspondences (II) and (III) are jointly distinct from (I) due to a series of 'palatal' reflexes in Tenetehára, Kayabí, Kamayurá, Guarayu and Ka'apor. Given the complementary distribution of these correspondences, both (II) and (III) are best reconstructed as reflecting **k*, just like (I), with special, context-specific developments taking place in the diverging languages. Note that the two upper rows in the table show that **ike* 'side of the body', which has never been reconstructed with **kʲ*, shows, nevertheless, the same reflexes as **-kʲer* 'sleep', which is reconstructed with PTG **kʲ* in every study that recognizes the distinction. If **k* is reconstructed in all these cases, the following developments are implied for each of the five languages:

- (2) Context-specific developments of PTG **k*
 In Kayabí, **k > s /_*e*
 In Guarayu, **k > kj /_*e*



In Kamayurá, *k > ts /*i_*e

In Tenetehára, *k > tʃ /*i_*e

In Ka'apor, *k > ʃ /*i_*e

For Kayabí, a search through Weiss' (2005) dictionary reveals that *ke* is an unattested sequence, which supports the regular operation of *k > s /_*e. Note that in 'side of the body', which is not reconstructed with *k' in the extant literature, Kayabí has *k > s, exactly as it does in the cases of 'sleep' and 'enter', both of which are usually reconstructed as having *k' (see section 'the current view'). This shows that Kayabí offers no evidence for the recognition of two distinct PTG velar stops.

The facts of Guarayu are the same as those of Kayabí, although the languages have phonetically distinct reflexes for *ke. Jensen (1999, p. 139) claims that the postulation of PTG *k' was motivated, in part, by the existence of *k'e* in Alfred Hoeller's data on Guarayu. The problem is that there is no *ke* in Guarayu and that all cases of PTG *ke show up as *k'e* in the language (cf. *iquie* 'die Seite des menschlichen Körpers'; *aquie* 'Ich schlafe, ruhe'; *aiquie* 'Ich trete ein'; Hoeller, 1932, pp. 90, 102, 210). A search in Danielsen et al. (2019) shows that in all cases where their data has *ke*, the same form in the Hoeller (1932) materials has *k' <quie>*. It seems that *ke* → [k'] is a purely allophonic process in Hoeller's Guarayu, one that affects the pronunciation of loanwords too, as in *kesu* 'cheese' (< Spanish *queso*), where Hoeller (1929, p. 88, quoted in Danielsen et al., 2019) registers a variant <*quiezu*> 'Käse'. There is no obstacle then for the postulation of *ke > k'e in the language, with the implication that Guarayu *k'* offers no evidence whatsoever for the postulation of a separate PTG proto-segment⁶.

For Kamayurá, *k > ts /*i_*e only within morphemes, which makes it difficult for assessing the regularity of the development since the environment is very specific. There does not seem to be any other currently reconstructible PTG morpheme, other than *-ike 'to enter', where a sequence *-ike- is found. That the development did not take place inter-morphemically is shown by the fact that Kamayurá *-ket* 'to sleep', when prefixed with the Set II third person marker *i-*, retains the velar stop as such (see Seki, 2000, p. 343, for an example). This restriction to tautomorphic contexts does not seem to be unique in the family, as noted below for Tenetehára, and it is active even in languages where the effect is simply variation in the existence or not of secondary palatalization *k* → [k']. This seems to be the case of Old Guarani, where optional palatalization takes place in the reflex of *-ike 'to enter' (cf. e.g., *yque* ~ *quíe* 'entrar' [to enter], *aiquie* 'yo entro' [I enter], *Teiquie* ~ *teique* 'entrar' [entry]; Montoya, 1639, p. 376), but not in the reflex of *-ket 'I sleep', when it is preceded by the Set III⁷ first person singular prefix *wi-* (cf. *aque* 'yo duermo' [I sleep], but: *giquebo*; Montoya, 1639, p. 330).

For Tenetehára, Jensen (1999, p. 139) argues that the medial affricate in *-itʃe* 'to enter' must be a reflex of *k', and not a contextual, palatalized reflex of *k conditioned by the preceding *i. As evidence for this claim she cites the

⁶ The same considerations apply to Guarayu *quie* 'wo, irgendwo, wohin, irgendwohin' (Hoeller, 1932, p. 210), which is sometimes offered as evidence for PTG *k'e 'here, near the speaker' (Jensen, 1999, p. 139). Note, though, that Jensen (1998, p. 550) gives *k'é 'here, near the speaker'. The reconstruction of the PTG system of demonstratives raises more complex issues than those tackled here and will not be further discussed in this contribution.

⁷ PTG is reconstructed with four sets of person-indexing prefixes. Set III markers are coreferential markers that are more commonly found in certain complement clauses featuring either positional verbs (a closed class of verbs specifying the spatial position of the subject while it participates in the event of the main clause) or in so-called 'gerund' constructions, where they signal a co-reference between the dependent (gerund) subject and the main clause subject. See Jensen (1998, 1999) for details.

diachronic correspondence *ikó* < **-ikó* 'to be in motion', which would be evidence that **i* had no general palatalizing effect upon a following **k* in Tenetehára. Note, however, that the two cases are not entirely comparable, and that the palatalization **k* > *tʃ* in Tenetehára could have applied only when preceded by **i* and followed by **e*, thus making the existence of *-ikó* in the modern language unsurprising. Moreover, it is not clear what the source for this presumed form *-iko* in Tenetehára is. While often given as a separate entry, for instance, as *iko* 'morar, viver, ser, estar' [dwell, live, be, stay] (Boudin, 1978, p. 73), the *-i* in this case results from diphthong formation whenever a preceding prefix vowel is added (as *a-iko* 'eu moro' [I dwell], *u-iko* 'êle está' [he is]) (Boudin, 1978, p. 73), and it reflects, in fact, an underlying *e*, which is present when no preceding vowel occurs, as in the third person form *hêkó-* (Boudin, 1978, p. 60). Although the verb in question does have a third person *ikó*, rather than *-ekó*, when used as a positional auxiliary, this fact carries no weight in rehabilitating Jensen's proposal. As noted by Bendor-Samuel (1972, p. 130) for the Guajajara dialect of Tenetehára, the verb *ikó* has a third person *i-* in this function, and it is not implausible that the apparently root-initial *i-* in this case is just the third person prefix in question (that is: **i-eko* > *iko*). Finally, see that, as in Kamayurá and Old Guarani, palatalization of **-ke* by a preceding **i* occurs only morpheme-internally.

The more attentive reader may have noticed yet another development possibly tied to the reflexation of PTG **k*. The Kagwahiva forms in Table 3 display a diachronic correspondence **e* > *i* for the vowel following **k*⁸. Jensen (1999, p. 139) appeals to this Kagwahiva development **e* > *i* as evidence for the presence of an earlier secondary palatalization in the preceding **k*, that is, as evidence for **ki*. However, Kagwahiva shows **ki* both in sets that have been analyzed in the literature as evidence for PTG **k'ie*, such as 'sleep', and in sets that have been reconstructed as **ke*, such as 'side of the body', and thus offers no evidence whatsoever of separate and contrasting reflexes (see the etymologies in the Appendix). It is likely that PTG **ke* [k'ie] > *ki* in Kagwahiva, with the precursor phonetic palatalization of **k* preceding **e* being not only phonetically natural but attested elsewhere in family, as noted above for Guarayu. Further evidence for this intermediate stage with phonetic palatalization [k'ie] as a condition for the change is the independent evidence for **e* > *i* in the context of a preceding palatal approximant **j*, as in *-nhi'ig* 'speak' (Betts, 2012, p. 188), from PTG **-jeʔëŋ* 'to speak' (Schleicher, 1998, p. 352), *-kyhyij* 'afraid' (Betts, 2012, p. 156) < **tʃikije* 'fear' (Schleicher, 1998, p. 341) 'fear' and in the reflexive prefix *ji-* (Betts, 2012, p. 121) < **je-* 'reflexive' (Jensen, 1998, pp. 515-516)⁹.

Correspondence (IV) differs from the identity correspondence (I) only in the Ka'apor reflex *f* alternating with *k*. As noted in Meira and Drude (2015) quote in the section 'the current view', Ka'apor *f* has been suggested as this language's reflex for the presumed PTG **k'*, in contrast to **k* > *k*. Any discussion of the potential evidence offered by Ka'apor reflexes for the reconstruction of PTG **k'* must consider a well-known innovation specific to Ka'apor which consists of the palatalization of **k* to *f* when preceded by **i* (Silva, 1997, pp. 49-50; Jensen, 1999, pp. 139-140). This produces alternations in the case of **k*-initial PTG roots/stems, which show *f* in their third person forms alternating with *k*- elsewhere in their paradigms. Table 4 presents diachronic correspondences between PTG nouns and their reflexes in Ka'apor, illustrating the effects of the Set II **i-* prefix on the initial **k-*.

⁸ The conclusion that Kagwahiva *ki* sequences are necessarily derived can also be arrived at given the fact (see 'the current view') that PTG had no **ki* sequence (and **ki* is likewise not reconstructed for Proto-Maweti-Guarani; see Meira & Drude, 2015).

⁹ This suggests that *je* sequences in Kagwahiva have an independent, a later origin, in Kagwahiva, and this is supported by an analysis of known cases, such as *-jehe'o* 'cry' (Betts, 2012, p. 120) < **-jatseʔo* 'to cry' (Mello, 2000, p. 166).

Table 4. Diachronic correspondences for PTG **i-k-* > Ka'apor *i-f-*.

PTG	Ka'apor	Source
<i>*i-kuʔa</i>	<i>ʃuʔa</i> 'his/her waist' <i>ihẽ kuʔa</i> 'my waist'	Kakumasu & Kakumasu (2007, p. 45)
<i>*i-kãŋ</i>	<i>ʃaŋwer</i> 'his/her bone' <i>timã kaŋwer</i> 'leg's bone'	Kakumasu & Kakumasu (2007, p. 46)
<i>*i-kãm</i>	<i>i-ʃambi</i> 'her milk, breast' <i>kambi</i> 'milk, breast'	Kakumasu & Kakumasu (2007, p. 175)
<i>*i-kupe</i>	<i>ʃupe</i> 'his/her back' <i>nde kupe</i> 'your (sg.) back'	Kakumasu & Kakumasu (2007, pp. 200-206)

Mello (2000, pp. 257-313) gives two cases where Ka'apor would have a *f* reflex for a PTG velar stop, one in the reflex for his PTG **k'ier* 'sleep' and the other in the set for PTG **kiʔa* 'dirty'¹⁰. First, note that the claim that Ka'apor has *f* as a reflex of PTG **k* in the form for 'sleep', as in the Mello (2000, p. 176) etymology for his PTG **k'ier* 'sleep', is factually incorrect: The form attested is *-ker*, as in *u-ker* 'ele dorme' [he sleeps] (Kakumasu & Kakumasu, 2007, p. 141). In agreement with the development PTG **i-k-* > *i-f-*, what Ka'apor does have is a derivative of *-ker* 'to sleep' which shows the expected palatalization when preceded by the Set II third person prefix *i-*: *i-ferai* 'ele sonha' [he dreams], as opposed to *ihẽ kerai* 'eu sonho' [I dream] (Kakumasu & Kakumasu, 2007, p. 193). It is possible that Mello (2000) has incorrectly coded the form for 'dream' in the 'sleep' set, but one cannot be sure about it, as the cognates in Mello's (2000) etymologies are not sourced. For the set for 'dirty', the existence of the third person *ʃiʔa* 'it is dirty' (Kakumasu & Kakumasu, 2007, p. 43) suggests an error in the same direction. Therefore, the supposed evidence for PTG **k'* in the form of a Ka'apor reflex *f* in the set for 'to sleep' (see section 'the current view') is non-existent. Finally, see that in correspondence (III) the reflex of PTG **-ike* 'to enter' has the expected *f* reflex in Ka'apor for medial **-k-*.

Two etymologies call for separate discussion since they apparently breach the pattern of complementary distribution observed for the correspondences (I) and (II-III). These are the terms for 'husband's sister' and 'elder brother', which were included in correspondence (I) in (1). The two involve etyma with **ke* sequences, just like the sets for correspondences (II) and (III) (see Table 3). However, the recognition of sporadic and language-specific developments, in addition to missing forms (due either to poor documentation or actual lexical replacement), allow one to account for this exceptionality without invoking an additional PTG proto-segment. The relevant cognate sets appear in Table 5, again with cell shading highlighting the most noteworthy data.

¹⁰ A fact which is exemplary of the many inconsistencies in Mello's data and analysis is the fact that, while Ka'apor *f* takes him to reconstruct **k'* in the case of 'sleep', this is not so in the set for 'dirty', even though both are presented as evidence for a Ka'apor **k' > f* change (see Mello, 2000, p. 128).

Table 5. Cognate sets displaying unexpected correspondences for PTG **ke*.

PTG	<i>*-ukeʔi</i> 'husband's sister'	<i>*-t-iket-ʔit</i> 'elder brother'
TUP	<i>ukeʔi</i>	<i>tikeʔir</i>
TEN	<i>ukeʔi</i>	<i>tikeʔir</i>
TOC	<i>okeʔi</i>	<i>-ikeʔit</i>
KAM	<i>ukeʔi</i>	<i>-rikeʔit</i>
WAJ	<i>ukeʔi</i>	<i>-lekiʔi</i>
KAY	<i>ukiʔi</i>	<i>-ekiʔit</i>
GUY	-	<i>-ikieir</i>
OGU	<i>ukei</i>	<i>tikeir</i>
KAA	-	-
KAG	-	<i>-rekiʔir</i>

The Kayabí reflexes are the first to strike the eye: The expected reflex of PTG **ke* in the language is *se*, not *ki*. For PTG **-ukeʔi* 'husband's sister' (see Carvalho & Birchall, 2022), one finds a Kayabí form *-ukiʔi* 'cunhada da mulher' [woman's sister-in-law] (Weiss, 2005, p. 109). The Kayabí have, however, in historical times, lived in a region geographically close to that of the Kagwahiva, in the Upper Tapajós river, with which they display cultural and historical affinities (Aguilar, 2017; Menendez, 1989, pp. 6-7). Since the development **ke > ki* evidenced by the Kayabí form is a regular Kagwahiva development, the best explanation, for the moment, is that Kayabí *-ukiʔi* is a Kagwahiva loan, even though the form seems to have been lost in Kagwahiva itself.

The same unexpected sequence *ki* is again attested in the Kayabí reflex of **-t-iket-ʔit* 'elder brother'. In this case, however, Kayabí, Wajāpi and Kagwahiva show a sporadic vowel metathesis: **-t-iket-ʔit > KAY -reki-ʔit* : WAJ *-lekiʔi* : KAG *-rekiʔir*. Although sporadic, metathesis is not unparalleled within TG, having targeted at least two other etyma: **-kipi-ʔit* 'younger sister, female Ego', which has a reflex *piki-ʔit* in some languages (Carvalho & Birchall, 2022), and **tsikije* 'to fear', with reflexes such as Kaiowá *kihije* (adapted from Schleicher, 1998, p. 341; see the etymologies in the Appendix of the present paper for comments on this particular etymon).

As noted in the section 'the current view', there are four cognate sets that are usually addressed in discussions of the issue of PTG **kⁱ*, but that have not been discussed here so far: 'electric eel', 'door', 'to close' and 'to snore'. Since these are offered as cases of (non-controversial) PTG **ke*, they will not add any evidence for reconstructing **kⁱ* and, for this reason, they will be only briefly discussed here.

PTG **keramu* 'to snore' (e.g., Mello, 2000, p. 172) is, as noted before, a derivative of **-ket* 'to sleep'. Inspection of the relevant etymology in the Appendix reveals that the reflexation of **k* in this set is identical to that of **-ket*, and hence, offers no evidence for a separate reflex. One can only speculate on the reasons that have led Mello (2000) to reconstruct an apparent contrast in the initial stops of **keramu* 'roncar' [to snore] (Mello, 2000, p. 172) and **kier* 'dormir' [to sleep] (Mello, 2000, p. 176), though the lack of a Kayabí cognate for in the former set, versus the Kayabí cognate with *s-* in the latter, have mislead him into recognizing two separate correspondences.

The three other sets, although often reconstructed for PTG, have distributional problems, and these will be addressed here for the sake of completeness. They have not been included in the etymologies featuring in the Appendix.



A form like **oken* is often reconstructed for the meaning 'door' in PTG (Rodrigues & Dietrich 1997, p. 273; Mello, 2000, p. 184; Meira & Drude, 2015, p. 292), though the cognates are restricted to Old Tupi *Oquêna* 'porta' [door] (Drumond, 1953, p. 83), Tenetehára *uken* 'porta' [door] (Harrison & Harrison, 2013, p. 157), Guarayu *oquienda* 'die Türe' [the door] (Hoeller, 1932, p. 159), Old Guarani *oquêna* 'puerta' [door] (Restivo, 1893 [1722], p. 455) and Ka'apor *huken ~ hukwen* 'porta' [door] (Kakumasu & Kakumasu, 2007, p. 96). That is, the form seems essentially restricted to the non-Amazonian TG languages and to languages that are, in some internal classifications of TG languages, suggested as having a rather close relation to Old Tupi: Tenetehára and Ka'apor (see e.g., Michael et al., 2015; Gerardi & Reichert, 2021). Quite telling is the absence of a cognate in Kayabí (*-'okwat* 'porta' [door] – Weiss, 2005, p. 165) and in the Kagwahiva lects (where an extension of *-juru* 'mouth', or, like Kayabí, of *-kwat* 'hole', is used instead; see Betts, 2012, p. 125)¹¹. Although consideration of a larger sample of languages (cf. Xingu Asurini *ukina* 'porta' [door], Pereira, 2009, p. 85) and of external, non-TG evidence (Meira & Drude, 2015, p. 292) make a PTG provenance for this set virtually safe, it offers no other insight on the reconstruction of the **k-k* contrast.

As noted before ('the current view'), **purake* 'poraquê' (Mello, 2000, p. 191), the name of a kind of fish or electric eel, is one of the forms traditionally discussed in the literature where PTG **k* would be attested preceding **e*. There are, however, both formal and distributional issues. Formally, the existence of forms with initial *m* (Tenetehára *murake* 'poraquê'; Harrison & Harrison, 2013, p. 113) matching forms with a supposedly etymological *p-* (Tocantins Asurini *poraké* 'poraquê'; Cabral & Rodrigues, 2003, p. 194), often with both attested in the same language (Kagwahiva *mburaki, puraki* 'electric eel'; Betts, 2012, p. 170) calls for adequate explanation. See that *m : p* correspondences, often with doublets in the same language, are expected in cases of Class Ib dependent nouns, where *m-* seems to code an unspecified possessor of the noun in question (Jensen, 1998, pp. 500-501, 1999, pp. 152-153). However, *purake/murake*, in languages that do have this item, is an independent noun, hence the correspondence cannot be accounted for in these morphological grounds. Second, the set lacks cognates in languages such as Kamayurá, Old Guarani and Guarayu and, although limited documentation prevents a simple inference of historical hypotheses, this is enough to command caution. There are other formal properties that call for explanation, such as Wajãpi having *i* unexpectedly matching *u* in the other languages – see *pilake* '*Electropharus electricus*' (Grenand, 1989, p. 92) –, and the coexistence of two forms, *pura* and *puraque* in Old Tupi (see Cardim, 1925 [1583], p. 88; Marcgrave & Piso, 1648, p. 151).

Finally, the set for PTG **kenab* 'fechar' [to close] is very limited in distribution already in Mello (2000, p. 172). Examination of comparative data reveals that there are a number of semantically close yet formally irreconcilable sets across TG languages, with some languages participating in multiple sets. Thus, an etymon *#pemim* is suggested¹² by Old Tupi *aipemim* 'cercar assi' [to enclose] (Drumond, 1952, p. 70), Ka'apor *jupimi* 'fechar o olho' [to close eyes] (Kakumasu & Kakumasu, 2007, p. 117) and Kamayurá *-pemi* 'fechar' [to close] (Seki, 2000, p. 317), while Mello's (2000) **-kenab* is somehow¹³ related to Old Tupi *Açoquendab* 'fechar porta' [close door] (Drumond, 1952, p. 136),

¹¹ This seems like a noteworthy gap in view of the common, if implicit, practice in comparative TG linguistics of accepting, as a criterion of minimal distributional strength for etymologies, the presence of cognates from one of the westernmost Amazonian TG languages, like Kayabí and Kagwahiva, in addition to cognates from the better attested southern languages like Old Tupi and one or more of the Guaranian lects. It is not difficult to find, say, in Lemle (1971) or Schleicher (1998), cognate sets which have been accepted on such grounds, even though the total number of comparanda in the sets is limited to three or four. This seems to rely implicitly on a perception that the great geographic distance between these languages virtually guarantees that a given comparison reflects, in fact, a PTG etymon.

¹² I use '#' instead of an asterisk for tentative reconstructions.

¹³ I say 'somehow' related because the Old Tupi cognate suggests a third person object prefix **-ts-*, and all cognates suggest that the root/stem is vowel-initial, *#-ukenab* perhaps. It is also likely that this etymon is ultimately related to the form for 'door'.

Tenetehára *ukênaw* ‘fechar, tapar buraco’ [close, close a hole] (Boudin, 1978, p. 282), Old Guarani *oñoquendá* ‘cerrar ventana o puerta sin llave’ [to close window or door without a key] (Restivo, 1893 [1722], p. 207). Tenetehára *u-wàpytym* ‘fechar’ [to close] (Harrison & Harrison, 2013, p. 183) and *Wajãpi ɔ-wapi* ‘fermer’ (Grenand, 1989, p. 59) suggest a third form with the same broad meaning. The fact that a single language, such as Old Tupi or Tenetehára, can participate in more than one set with semantically similar cognates suggests that independent etyma with meanings such as ‘enclose’, ‘close’, ‘cover with lid’ got confounded, either due to semantic extensions and replacement in some of the languages, or because the relevant sources are too coarse in the semantics of the material included. Be that as it may, Mello’s (2000) **-kenaβ*, if accepted as a PTG etymon offers, at best, another instance of PTG **ke*, and no evidence whatsoever for a PTG velar contrast in this context.

SOME IMPLICATIONS

The proposal that PTG had a single velar stop **k* offers not only the best account for the relevant comparative correspondences but also eliminates inconsistencies from the previous reconstruction with a **k* - **kʲ* contrast. Jensen (1999, p. 139, fn. 22) noted, for instance, the anomalous character of the diachronic correspondence PTG **kʲer* > Tenetehára *ker* ‘sleep’, since PTG **kʲ* predicts, in her account, a reflex *tʃ* in the language. No such anomaly exists under the current proposal.

In addition, there are implications of the findings reported here for our understanding of the diversification of the Tupian language family. Rodrigues (2007, pp. 180-181) reconstructs ***kʲ* for the Proto-Tupian (PT) parent language, but this implies an unmotivated split in the PTG reflex: while ***kʲ* merges with ***k* in ***ikʲet* > **-iker* ‘irmã senior da mulher’ [older sister, female Ego], it is retained in ***kʲet* > **kʲer* ‘dormir’ [sleep], in both cases the same phonetic context of a following ***e* > **e* yields an unmotivated bifurcation of PT ***kʲ* (see also Rodrigues, 2005, p. 40; Rodrigues & Cabral, 2012, pp. 505-507). The present reconstruction of PTG eliminates this unmotivated split. If PT must be reconstructed with a ***k* - ***kʲ* contrast, PTG offers no special evidence in this respect, and the contrast was likely merged already at the Proto-Maweti-Guarani level (see Meira & Drude, 2015).

CONCLUSIONS

This paper has shown that there is no need to reconstruct a contrast between a plain velar stop **k* and a palatalized velar stop **kʲ* for the parent language of the Tupi-Guarani family. All diachronic divergences from reconstructed etyma can be accounted for as conditioned developments of PTG **k*, and one sporadic development, represented in (3) as diachronic replacements in specific segmental sequences:

- (3) Diachronic replacements proposed in this paper for individual TG languages
- *ike* > *itse* (morpheme internally, in Kamayurá)
 - *ike* > *itʃe* (morpheme internally, in Tenetehára)
 - *ike* > [ke ~ kje] (morpheme internally, in Old Guarani)
 - *ke* > [kje] in Guarayo
 - *ke* > *se* in Kayabí
 - *ke* > *ki* in Kagwahiva
 - *ik* > *if* in Ka’apor
 - *ike* > **ekʲ* metathesis in form for ‘elder brother’ in Kayabí, Wajãpi and Kagwahiva

The relatively lengthy discussion presented here in order to deal with one very specific issue on the reconstruction of PTG shows that a proper understanding of TG historical phonology requires more attention to detail and a more careful treatment of the comparative data than has been the case so far. If further progress in our understanding of the historical development of TG languages is to be attained, the practices of relying on a superficial treatment of correspondences, or what is worse, on a few supposedly conservative languages that are taken as proxies for PTG, should be left behind as features of the past of comparative TG historical linguistics.

REFERENCES

- Aguilar, A. M. G. C. (2017). Kawahíwa como uma unidade linguística. *Revista Brasileira de Linguística Antropológica*, 9(1), 139-161. <https://doi.org/10.26512/rbla.v9i1.19529>
- Araújo, A. (1895 [1686]). *Catecismo brasílico da doutrina cristã* [Facsimile edition by Julius Platzman]. B. G. Teubner.
- Bendor-Samuel, J. (1972). *Hierarchical structures in Guajajara* (Summer Institute of Linguistics Publications, 37). SIL/University of Oklahoma.
- Betts, L. (2012). *Kagwahiva dictionary*. Summer Institute of Linguistics (SIL).
- Boudin, M. (1978). *Dicionário de tupí moderno: dialeto tupí-tenetehár do Alto Gurupí*. Conselho Editorial de Artes e Ciências Humanas de São Paulo.
- Cabral, A. S. A. C., & Rodrigues, A. D. (2003). *Dicionário Asuriní do Tocantins-Português*. UFPA.
- Cardim, F. (1925 [1583]). *Tratados da terra e gente do Brasil*. Editores J. Leite & Cia.
- Carvalho, F. O., & Birchall, J. (2022). A comparative reconstruction of Proto-Tupi-Guarani kinship terminology. *LIAMES*, 22, e022001. <http://doi.org/10.20396/liames.v22i00.8666489>
- Castilho, P. (1937 [1613]). *Os "nomes das partes do corpo humano pella lingua do Brasil" de Pero de Castilho* [Edition by Plinio Ayrosa]. Empresa Gráfica da Revista dos Tribunais.
- Danielsen, S., Sell, L., & Terhart, L. (2019). Guarayu. A revised dictionary by Alfred Hoeller. *Dictionaria*, 7, 1-3590. <http://doi.org/10.5281/zenodo.4675101>
- Dietrich, W. (1990). *More evidence for an internal classification of tupi-guarani languages*. Gebr. Mann Verlag.
- Drumond, C. (Org.). (1952). *Vocabulário na Língua Brasileira* (Vol. 1, A-H). Faculdade de Filosofia, Ciências e Letras, Universidade de São Paulo. http://etnolingua.wdfiles.com/local--files/biblio%3Adrumond-1952-1953-v1b/MLBrasilia_2edDrumond_1952v1_A-H_OCR.pdf
- Drumond, C. (Org.). (1953). *Vocabulário na Língua Brasileira* (Vol. 2, I-Z). Faculdade de Filosofia, Ciências e Letras, Universidade de São Paulo. http://etnolingua.wdfiles.com/local--files/biblio%3Adrumond-1952-1953-v1b/MLBrasilia_2edDrumond_1953v2_I-Z_OCR.pdf
- Gerardi, F., & Reichert, S. (2021). The Tupi-Guarani language family. *Diachronica*, 38(2), 151-188. <https://doi.org/10.1075/dia.18032.fer>
- Grenand, F. (1989). *Dictionnaire Wayãpi-Français: Lexique Français-Wayãpi*. Peeters/SELAF.
- Harrison, C., & Harrison, C. (2013). *Dicionário Guajajara-Português*. Associação Internacional de Linguística (SIL).
- Hoeller, A. (1929). *Diccionario Guarayo-Castellano*. Ms.
- Hoeller, A. (1932). *Diccionario guarayu-castellano*. COPNAG.
- Jensen, C. (1984). *O desenvolvimento histórico da língua Wayampi* [Masters' thesis, Universidade Estadual de Campinas].
- Jensen, C. (1998). Comparative tupí-guaraní morphosyntax. In D. Derbyshire & G. K. Pullum (Eds.), *Handbook of Amazonian languages* (Vol. 4, pp. 487-618). Mouton de Gruyter.



- Jensen, C. (1999). Tupí-guaraní. In R. M. W. Dixon & A. Aikhenvald (Eds.), *The Amazonian languages* (pp. 125-163). Cambridge University Press.
- Kakumasu, J., & Kakumasu, K. (2007). *Dicionário por tópicos Kaapor-Português*. Associação Internacional de Linguística (SIL).
- Lemle, M. (1971). Internal classification of the tupi-guarani linguistic family. In D. Bendor-Samuel (Ed.), *Tupi studies I* (Summer Institute of Linguistics Publications in Linguistics and Related Fields, Vol. 29, pp. 107-129). Summer Institute of Linguistics.
- Lemos Barbosa, A. (1948). *O Vocabulário na Língua Brasileira*. Ministério da Educação e Saúde.
- Marcgrave, G., & Piso, W. (1648). *Historia Naturalis Brasiliae*. Joannes de Laet.
- 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 tupí-guaraní: aspectos fonológicos e lexicais* [Doctoral dissertation, Universidade Federal de Santa Catarina].
- Menendez, M. A. (1989). *Os Kawahiva: uma contribuição para os estudos dos Tupi Centrais* [Doctoral dissertation, Universidade de São Paulo].
- Michael, L., Chosou-Polydouri, N., Bartolomei, K., Donnelly, E., Meira, S., Wauters, V., & O'Hagan, Z. (2015). A bayesian phylogenetic classification of tupi-guarani. *LIAMES: Línguas Indígenas Americanas*, 15(2), 193-221. <https://doi.org/10.20396/liames.v15i2.8642301>
- Montoya, A. R. (1639). *Tesoro de la lengua guarani*. Juan Sanchez.
- Peggion, E. A. (1996). *Forma e função: uma etnografia do sistema de parentesco Tenharim (Kagwahiv, AM)* [Master thesis, Universidade de Campinas].
- Pereira, A. (2009). *Estudo morfossintático do Asurini do Xingu* [Doctoral dissertation, Universidade de Campinas].
- Restivo, P. (1893 [1722]). *Lexicon Hispano-Guaranicum*. Wilhelm Kohlhammer.
- Rodrigues, A. D. (1984/1985). Relações internas na família tupí-guaraní. *Revista de Antropologia*, 27/28, 33-53.
- Rodrigues, A. D., & Dietrich, W. (1997). On the linguistic relationship between mawé and tupí-guaraní. *Diachronica*, 14(2), 265-304. <http://dx.doi.org/10.1075/dia.14.2.04rod>
- Rodrigues, A. D., & Cabral, A. S. A. C. (2002). Revendo a classificação interna da família tupi-guarani. In A. S. A. C. Cabral & A. D. Rodrigues (Eds.), *Línguas indígenas brasileiras* (pp. 327-337). UFPA.
- Rodrigues, A. D. (2005). As vogais orais do Proto-Tupí. In A. D. Rodrigues & A. S. A. C. Cabral (Eds.), *Novos estudos sobre línguas indígenas* (pp. 35-46). Editora da UnB.
- Rodrigues, A. D. (2007). As consoantes do Proto-Tupí. In A. S. A. C. Cabral & A. D. Rodrigues (Eds.), *Línguas e culturas Tupí* (pp. 167-203). Curt Nimuendajú/LALI/UnB.
- Rodrigues, A. D., & Cabral, A. S. A. C. (2012). Tupian. In L. Campbell & V. Grondona (Eds.), *The indigenous languages of South America* (pp. 495-574). Mouton de Gruyter.
- Schleicher, C. O. (1998). *Comparative and internal reconstruction of the tupi-guarani language family* [Doctoral dissertation, University of Wisconsin].
- Seki, L. (2000). *Gramática do kamayurá*. Editora da Unicamp.
- Silva, B. C. C. (1997). *Urubu-Ka'apor: da gramática a história* [Master thesis, Universidade de Brasília].
- Weiss, H. E. (2005). *Dicionário Kayabí-português*. Summer Institute of Linguistics (SIL).

Appendix. Etymologies. The following Appendix contains all the cognate sets that were employed in the present work. All forms are cited as they appear in the source orthography, followed by the original source glosses and with references to where in each source a given form can be found. The abbreviations employed for language names and sources are as follows: Old Tupi (TUP): "Vocabulário na Língua Brasileira" (Drumond 1952, 1953) (VLB); Araújo (1895 [1686]) (A86), Castilho (1937 [1613]) (C13); Old Guarani (OGU): Restivo (1893 [1722]) (R22), Montoya (1639) (M39); Ka'apor (KAA): Kakumasu & Kakumasu (2007) (KK07); Guarayu (GUY): Hoeller (1929) (H29), Hoeller (1932) (H32), Danielsen et al. (2019) (DST19); Tocantins Asurini (TOC): Cabral & Rodrigues (2003) (CR03); Kagwahiva (KAG): Peggion (1996) (P96), Betts (2012) (B12); Kayabí (KAY): Weiss (2005) (W05); Wajãpi (WAJ): Grenand (1989) (G89), forms followed by 'Amapari Wajãpi' come from the author's own fieldwork notes; Tenetehára (TEN): Boudin (1978) (B78), Harrison & Harrison (2013) (HH13); Kamayurá (KAM): Seki (2000) (S00). Grammatical abbreviations are limited to 'intransitive' (INTR.), 'third person' (3), and 'singular' (sg.).

SPEAK ***-jeřěŋ**

TUP *Anheeng* 'falar' (VLB, I, 133) : TEN *u-ze'eg* 'falar' (HH13:201) : TOC *-se'éng* 'falar' (CR03:216) : KAM *je'eŋ* 'falar, fala' (S00:458) : WAJ - : KAY *-je'eğ* 'falar' (W05:36) : GUY *añee* 'Ich spreche, antworte' (H32:152) : OGU *añeê chupe* 'hablar' (R22:323) : KAA *je'ě* 'comunicação dos animais, como o pássaro, o sapo' (KK07:86) : KAG *-nhi'ig* 'speak' (B12:188)

FEAR ***-tsikije/*kitsije**

TUP *Cigquigjê* 'Medo, timor' (VLB, II, 34) : TEN *u-kyze* 'ter medo' (HH13:160) : TOC *kyysé* 'medo' (CR03:114) : KAM *-kyje* 'temer' (S00:218) : WAJ *ɔ-kiye* 'peur (avoir)' (G89:69) : KAY *-kyyje* 'temer, estar com medo' (W05:54) : GUY *ziqũye* 'Furcht' (H32:342) : OGU *quĩhiye* 'miedo' (R22:385) : KAA *kyje* 'ele tem medo de' (KK07:122) : KAG *-kyhyji* 'afraid' (B12:156)

Comments: A comparison of cognate forms shows that metathesis occurred in a subset of the languages. In the absence of external comparanda, however, it is difficult to decide which is the precise form of the etymon. Both forms are attested in the Old Tupi *corpus*, appearing in the VLB as *Aciquigjê* 'Medo ter ou auer, O mesmo he Aquicigje como algũs dizê' (VLB, II, 34).

PULL ***-ts-ekij**

TUP *Acequigy* 'Puxar' (VLB, II, 90) : TEN *u-ekyz* 'arrastar' (HH13:210) : TOC *-ekyj* 'arrancar, puxar' (CR03:66) : KAM *ekyj* 'puxar, extrair' (S00:456) : WAJ *-ekii* 'prendre; attraper, saisir' (G89:172) : KAY *-ekyi* 'puxar, desatar' (W05:22) : GUY *zequii, azequii* 'Ich löse es herab' (H32:335) : OGU *ahequiĩ* 'arrancar cosa hincada' (R22:99) : KAA - : KAG *-ekyi ~ -ekyj* 'take out, remove' (B12:72)

LOOK FOR ***-ts-ekat**

TUP *Acecar* 'buscar' (VLB, I, 60) : TEN *u-ekar* 'procurar' (HH13:210) : TOC - : KAM *ekat* 'procurar' (S00:456) : WAJ *-eka* 'chercher' (G89:171) : KAY *-ekat* 'procurar' (W05:21) : GUY *zeca, azeca* 'Ich suche ihn' (H32:327) : OGU *aheca* 'buscar' (R22:143) : KAA *kekar* 'ele caça' (KK07:92) : KAG *-ekar* 'search for' (B12:71)

Comments: Inclusion of the Ka'apor cognate is tentative, since there is no explanation for initial *k-* in the presumed cognate. The most promising hypothesis is that it derives from an univerbation of *kaʔa* 'woods' and *-ekar* 'to look for', that is 'to look for (something) in the woods' = 'hunting'.



CUT ***-kiti**

TUP *Aiquigti* 'cortar como serra, tesoura, faca et cete' (VLB, I, 83) : TEN *u-kixi* 'cortar' (HH13:158) : TOC *-kytitát* 'cortador' (CR03:112) : KAM *kytsi* 'cortar' (S00:460) : WAJ *-kisi* 'couper' (G89:230) : KAY *-kysi* 'cortar' (W05:54) : GUY *aiquỹchĩ* 'Ich schneide es' (H32:214) : OGU *aiquĩti* 'cortar asserrando ó con cuchillo' (R22:192) : KAA - : KAG *-kyti* 'cut, circumcise' (B12:157)

Comments: Note nasality in Old Guarani and Guarayu. This word-final nasality is also attested in modern Guaranian varieties, and its origin remains an open problem.

PIERCE ***-kutuk**

TUP *Aicutuc* 'furar' (VLB, I, 145) : TEN *u-kutuk* 'furar' (HH13:158) : TOC *-kotók ~ -kotóng* 'furar' (CR03:109) : KAM *kutuk* 'furar' (S00:460) : WAJ (*momu* 'percer' (G89:69)) : KAY *-kutuk* 'furar; ferrar, picar' (W05:53) : GUY *aicutu* 'Ich steche ihn' (H32:66) : OGU *aycutu* 'herir' (R22:328), *cutúg* 'herir, barrenar, punçar, sangrar' (M39:111) : KAA *kutuk* 'ele lava; ele fura' (KK07:48) : KAG *-kutug* 'pierce, stab' (B12:148)

Comments: Wajãpi *momu* 'percer' is a reflex of **mo-puk* 'bore a hole'.

BE\STAY ***-eko**

TUP *Aicô* 'estar como quer' (VLB, I, 128) : TEN *a-iko* 'eu moro, eu estou' (B78:73), *hêko* 'estar, ficar, permanecer' (B78:60) : TOC *-eká ~ -ká* 'ser, estar em movimento' (CR03:64) : KAM *-eko ~ -ko* 'ser, estar' (S00:456) : WAJ *ɔ-i-ko* 'être' (G89:58) : KAY *-eko* 'estar; estar vivo', *-ko* 'estar' (W05:21, 51) : GUY *zeko* 'leben, sein' (H29:102), *a-ico* 'Ich bin, lebe, weile, wohne' (H32:88) : OGU *aico* 'estoy' (R22:294), *Tecó* 'ser, estado de vida, condición, estar, costumbre, ley, habito' (M39:363) : KAA *reko* 'ele tem', *nixói* 'tem (não), não há' (KK07:194) : KAG *-(e)ko* 'be, exist, remain' (B12:138)

Comments: The Ka'apor form is likely a reflex of PTG **(e)ro-eko* 'to be with, to have'.

SNORE, TO ***ket-amu**

TUP *Xequerambû* 'roncar, o que dorme' (VLB, II, 108) : TEN *u-keramu* 'roncar' (HH13:158) : TOC - : KAM - : WAJ *kelamu* 'ronflement' (G89:226) : KAY *seramũ* 'roncar' (W05:99) : GUY *che quierambu* 'Ich schnarche beim Schlafen' (H32:14) : OGU *aquerambu* 'roncar', *ambu* 'ronquido', *tayaçu apiĩmbu* 'de puerco' (R22:482) : KAA - : KAG *-kirambu* 'snore' (B12:137)

SIDE (OF THE BODY) ***ike**

TUP *Igque* 'lado, ou ilharga' (VLB, II, 17) : TEN *Ikê* 'lado, costado' (B78:72) : TOC - : KAM *yke* 'lado' (S00:467) : WAJ *-ikε-lupi* 'a côté de' (G89:192) : KAY *-yse* 'lado de algo' (W05:119) : GUY *iquie* 'die Seite des menschlichen



Körpers' (H32:102) : OGU *ŷque* 'lado, costado' (R22:354) : KAA *rake* 'perto, ao lado' (KK07:134) : KAG *-ykia* 'side, the side of the body' (B12:279)

Comments: Ka'apor is a tentative cognate, as *ra-* is difficult to account for.

WAIST **-kuʔa*

TUP *Cuâ* 'Cintura' (C13:31) : TEN *iku'aw* ~ *iku'a* 'na cintura, pelo meio' (HH13:59) : TOC *-ko'á* 'cintura, cadeiras, quadril' (CR03:104) : KAM *-ku'a* 'cintura' (S00:400) : WAJ *kuʔa-kaʔi* 'taille' (G89:75) : KAY *-ku'a* 'cintura; nádegas; popa' (W05:51) : GUY *cu-a, che cú-a* 'meine Mitte, Taille' (H32:54) : OGU *Cuá* 'el medio entre los extremos' (M39:102) : KAA *xu'a* 'a cintura' (KK07:145) : KAG *-ku'a* 'buttocks, lower back from waist to legs' (B12:144)

BREAST **-kãm*

TUP *Cãma* 'Tetas' (VLB, II, 127) : TEN *i-kãm* 'mama, seio' (HH13:281) : TOC *-kóm* 'seio' (CR03:105) : KAM *-kam* 'seio' (S00:399) : WAJ - : KAY *-kam* 'seio' (W05:48) : GUY *cã* 'Brust, Tete, Euter' (H32:43) : OGU *Cãma* 'pecho de muger' (R22:423) : KAA *ixamby, kamby* 'leite dela; seio dela' (KK07:175) : KAG *kama* 'breast' (B12:131)

BONE **-kãŋ*

TUP *canga, canguera* 'osso, ossada' (VLB, II, 59) : TEN *i-kãgwer* 'osso separado do corpo' (HH13:288) : TOC *-kýng* 'osso' (CR03:111) : KAM *kang* 'osso' (S00:400) : WAJ *kãngɛ* 'os' (G89:68) : KAY *-kağ* 'osso' (W05:48) : GUY *cã, mbae canguer* 'Knochen' (H32:43) : OGU *cãng* 'hueso' (M39:88) : KAA *xanguer* ~ *ixanguer* 'osso dele' (KK07:182), *haji kanguer* 'osso do queixo dele' (KK07:44) : KAG *kağa* 'bone' (B12:129)

BACK **-kupe*

TUP *Cupe* 'costas, a parte de tras' (VLB, I, 84) : TEN *i-kupe* 'costas' (HH13:262) : TOC - : KAM *-kupe-kang* 'coluna vertebral' (S00:400) : WAJ *kupe* 'surface plate: épaule, nageoire dorsale, crête (de l'iguane)' (G89:248) : KAY *kupe* 'parte traseira, costas' (W05:53) : GUY *cupe* 'Rücken, Schulter' (H32:63) : OGU *cupe* 'espalda' (R22:290) : KAA *xupe* 'as costas dele', *ihẽ kupe* 'minhas costas' (KK07:45) : KAG *-kupea* 'back' (B12:146)

KNIFE **kitse*

TUP *Quigcê, Jtaquigcê* 'faca' (VLB, I, 133) : TEN *takihe* 'faca, facão' (HH13:139) : TOC *kyhé* 'faca', *kyé'í* 'faquinha' (CR03:111) : KAM *kye'i* 'faca' (S00:460) : WAJ *kisɛ* 'Couteau' (G89:230) : KAY *kye* 'faca; ponta de flecha ou de lança' (W05:53) : GUY *quize* 'Messer' (H32:214) : OGU *quice* 'cuchillo' (R22:198) : KAA *kyse* 'o terçado, a faca grande' (KK07:94) : KAG *itakyhea* 'long-bladed terçado machete' (B12:112)



TONGUE *-ape-kũ

TUP *Apecũ* 'Lingoa, pello membro' (VLB, II, 22) : TEN *i-apeku* 'língua' (HH13:280) : TOC *né apekó-a* 'teu céu da boca' (CR03:45) : KAM *-apekõ 'úvula', -kõ* 'língua' (S00:398, 400) : WAJ *apekũ, -kũ* 'langue' (G89:64) : KAY *apekũ* 'guelra' (W05:10), *-kũ* 'língua' (W05:51) : GUY *-apěcũ* 'die menschliche Zunge' (H32:21) : OGU *checũ* 'mi lengua' (R22:358), *âpěcũ* 'lengua y paladar' (M39:49) : KAA - : KAG *-'apekũa* 'fish gills; uvula, voice box or tongue' (B12:40), *-kũa* 'tongue' (B12:144)

Comments: For Wajãpi, *-apekũ* is the main term for 'tongue', while *-kũ* is referred to as a compound stem with the same meaning but restricted to compounds – as in *tapiʔikũ* 'langue de tapir' (a plant name) (G89:254). Grenand (1989, p. 296) postulates an etymological connection between *mokũ* 'avaler' (G89:296) and this compound root *-kũ* 'tongue'.

SLEEP *-ket

TUP *Aquer* 'dormir' (VLB, I, 106) : TEN *u-ker* 'dormir' (HH13:157) : TOC *-két, -kén, -kér* 'dormir' (CR03:103) : KAM *ket* 'dormir' (S00:460) : WAJ *-ke* 'dormir' (G89:226) : KAY *-set* 'dormir, pousar' (W05:100) : GUY *quie, aquie* 'Ich schlafe, ruhe' (H32:210) : OGU *aque* 'dormir' (R22:255) : KAA *ukwer ~ uker* 'dorme (ele)' (KK07:164) : KAG *-kir ~ -ngir* 'sleep' (B12:137)

ENTER *-ike

TUP *Aiquê* 'entrar' (VLB, I, 119) : TEN *u-ixe* 'entrar' (HH13:216) : TOC *-ké* 'entrar', *aké wehá* 'eu entrei (na casa)' (CR03: 101) : KAM *'itse* 'entrar' (S00:454) : WAJ *ɔ-y-ke* 'entrer' (G89:57) : KAY *-se* 'entrar' (W05:99) : GUY *iquie, aiquie* 'Ich trete ein' (H32:90) : OGU *ayque l. ayquie* 'entrar' (R22:280) : KAA *ixe* 'entra (ele)' (KK07:165) : KAG *-ki ~ -eki ~ -ngi* 'enter' (B12:136)

Comments: See that the Kagwahiva allomorphs have the following distribution: *-ngi* when preceded by the Causative prefix *mo-*; *-eki* when preceded by *h-* or *r-* (that is, in Absolutive constructions where the sole argument of the verb is indexed with a Set II marker), and *-ki* elsewhere.

LONG *-puku ~ *-muku

TUP *Mucũ, pucũ* 'Longa cousa' (VLB, II, 24) : TEN *i-puku* 'comprido' (HH13:72) : TOC *-pokó* 'comprido' (CR03:192) : KAM *huku* 'ser comprido' (S00:457) : WAJ *pɔkɔ, puku* 'long' (G89:64) : KAY *fuku* 'alto' (W05:30), *muku* 'longe, distante' (W05:75) : GUY *pucu* 'lang' (H32:208) : OGU *pucú* 'largo' (M39:323) : KAA *puku* 'é comprido' (KK07:132) : KAG *-puku ~ -mbuku* 'long in space or time' (B12:229)

FAT, BE *-kira

TUP *Xequirà* 'gorda ser a pessoa, ou qualquer outro animal quadrupes' (VLB, I, 149) : TEN *i-kyra* 'gordo, com saúde' (HH13:60) : TOC - : KAM *kyra* 'ser gordo' (S00:460) : WAJ *kila* 'Graisse, gras' (G89:229) : KAA - : KAY *ky'ra* 'gordo' (W05:54) : GUY *quira, mbae quira* 'Fett, Speck' (H32:213) : OGU *quĩracue* 'Grassa' (R22:319) : KAG *kyr, kyra* 'fat, chunky, stout' (B12:156)



HOT ***-akup**

TUP *Xeracub* 'quente estar', *Acuba* 'quente adiectivo' (VLB, II, 94) : TEN *h-aku* 'quente' (HH13:27) : TOC *-akop* ~ *-akom* 'quente' (CR03:33) : KAM *-akup* 'quente' (S00:67) : WAJ *aku* 'chaud' (G89:52) : KAY *-akup* ~ *-akuw-a* 'quente' (W05:5) : GUY *acu, tacu, zacu, racu* 'heiss' (H32:7) : OGU *tacu* 'caliente' (R22: 149) : KAA *haku* 'quente (está)' (KK07:188) : KAG *-akuv* 'hot or burning from sunburn or fire or fever' (B12: 29)

GOOD ***-katu**

TUP *Catu* 'bem, bene' (VLB, I, 54) : TEN *katu haw* 'bondade, o que presta, segurança' (HH13:87) : TOC *katóeté* 'bom, bem' (CR03:100) : KAM *-katu* 'ser bom' (S00:63) : WAJ *i-katu* 'Bon (être)' (G89:50) : KAY *katu* 'bom, certo' (W05:50) : GUY *catupirì* 'gut, schön' (H32:50) : OGU *ycatupirĩ* 'Bien, está bien hecho' (R22:130) : KAA *katu* 'bom (é)' (KK07:174) : KAG *-katu* 'pretty, good' (B12:135)

WET ***-akĩm**

TUP *Aquigma* 'molhada cousa', *Xeaquigm* 'molhado estar' (VLB, II, 40) : TEN *i-àkym* 'molhado (estar)' (HH13:48) : TOC *-akým* 'molhado' (CR03:35) : KAM *'akym* 'estar molhado' (S00:454) : WAJ *-akã* 'wet' (Amapari Wajãpi) : KAY *-akym* 'molhado' (W05:7) : GUY *aquỹ, ñaquỹ* 'feucht, nass, durchnässt' (H32:28) : OGU *cheaquỹ* 'mojarse' (R22:389) : KAA *iankym* 'molhado (está)' (KK07:179) : KAG *-akym* 'wet' (B12:31)

DIRTY ***-kiʔa**

TUP *Quigã* 'Çuja ou çujo' (VLB, I, 87) : TEN *ki'a* 'sujo (ser, estar)' (B78:103) : TOC - : KAM - : WAJ *kiʔa* 'saleté, être sale' (G89:228) : KAY *-ky'a* 'sujo' (W05:53) : GUY *quia* 'Schmutz' (H32:212) : OGU *quiã* 'suciedad' (R22:504) : KAA *xi'a* 'sujo (está)' (KK07:193) : KAG *-ky'a* ~ *-ngy'a* 'dirty, black with dirt' (B12:155)

WOODS ***kaʔa**

TUP *Caã* 'mata ou matos' (VLB, II, 33) : TEN *ka'a* 'mata, floresta' (HH13:89) : TOC *ka'á* 'mato' (CR03:94) : KAM *ka'a* 'mata, folha' (S00:459) : WAJ *kaʔa* 'forêt' (G89:60) : KAY *ka'a* 'mato, folha' (W05:47) : GUY *caa* 'Wald' (H32:43) : OGU *caá* 'bosque' (R22:390) : KAA *ka'a* 'o mato' (KK07:118) : KAG *ka'a* 'leaf' (B12:128)

BRANCH ***-ts-akã**

TUP *Çacã* 'Rama' (VLB, II, 96) : TEN *h-àkà* 'galho, ramo' (HH13:26) : TOC *-akó* 'galho' (CR03:33) : KAM - : WAJ *ãkã* 'branche' (G89:50) : KAY *akã* (W05:151), *'ywarakã* 'galho de árvore' (W05:5) : GUY *zãcã* 'sein Ast' (H32:316), *tãcã*



'Zweig, Ast' (H32:232) : OGU *ŷbŷra racângue* 'rama' (R22:463) : KAA *hankã* 'o riacho, igarapé, galho' (KK07:105) : KAG *-akã* 'branch of tree' (B12:26)

LOUSE **-kip*

TUP *Quigba* 'piolhos' (VLB, II, 78) : TEN *kyw* 'piolho' (HH13:99) : TOC *kýp, kýwa* 'piolho' (CR03:111) : KAM *ikip* 'piolho dele' (S00: 415) : WAJ *kii* 'pou' (G89:70) : KAY *-kyp* 'piolho' (W05:53) : GUY *qui* 'Haarlaus' (H32:212) : OGU *qui.b* 'Piojo de cabeça' (R22:434) : KAA *ky* 'piolho' (KK07:185) : KAG *-kyva* 'louse (piolho)' (B12:158)

CAYMAN **jakare*

TUP *Iacare* 'lagarto dagoa' (VLB, II, 17) : TEN *zakare* 'jacaré' (HH13:229) : TOC *sakarétíng* 'espécie de jacaré' (CR03:207) : KAM *jakare* 'jacaré' (S00:457) : WAJ *yakale* 'reptile (sp.)' (G89:500) : KAY *jakare* 'jacaré' (W05:33) : GUY *yacare* 'Kaiman' (H32:273) : OGU *yacare* 'lagarto de agua' (R22:355) : KAA *jakare* 'jacaré' (KK07:173) : KAG *jakarea* 'alligator' (B12:116)

MONKEY **kaʔi*

TUP *Caí* 'Bogio não tem gênero, outros maiores' (VLB, I, 56) : TEN *ka'i* 'macaco (genérico)' (HH13:90) : TOC *ka'í* 'macaco-prego' (CR03:95) : KAM *ka'i* 'macaco-prego' (S00:459) : WAJ *kaʔi* 'Macaque' (G89:110) : KAY *ka'i* 'macaco' (W05:48) : GUY *cai* 'Affe' (H32:45) : OGU *caý* 'mono' (R22:390) : KAA *ka'ijarar* 'macaco caiarara' (KK07:175) : KAG *ka'ia* 'general term for monkey' (B12:130)

SALT **jukit*

TUP *Iuquigra* 'Sal' (VLB, II, 111) : TEN *zukyr* 'sal' (HH13:238) : TOC *sykýt ~ sykýn ~ sykýr-a* 'sal' (CR03:232) : KAM *jokit ~ jukit* 'sal' (S00:429) : WAJ *yuki* 'sel' (G89:73) : KAY *jukyt* 'sal' (W05:45) : GUY *yuquìr* 'Salz' (H32:312) : OGU *juquĩ* 'sal' (R22:486) : KAA *jukyr* 'sal' (KK07:191) : KAG *jukyra* 'salt' (B12:124)

KILL **juka*

TUP *Ajuca* 'matar como quer' (VLB, II, 33) : TEN *u-zuka* 'matar' (HH13:202) : TOC *-soká* 'matar' (CR03:226) : KAM *juka* 'matar' (S00:459) : WAJ *-yuka* 'Tuer' (G89:76) : KAY *-juka* 'matar (uma entidade)' (W05:45) : GUY *ayuca* 'Ich töte ihn' (H32:311) : OGU *ayuca* 'matar' (R22:379) : KAA *jukwa* 'ele mata' (KK07:92) : KAG *-juka* 'kill' (B12:124)

KNOW **-kuwaap*

TUP *Aicuguab* 'conhecer' (VLB, I, 80) : TEN *u-kwaw* 'saber, conhecer' (HH13:159) : TOC *-kwaháp ~ -kwahám* 'saber, conhecer' (CR03:115) : KAM *kwahap* 'saber, conhecer' (S00:460) : WAJ *-kuwa* 'savoir' (G89:252) : KAY *-kwaap*



'saber, conhecer; entender' (W05:54) : GUY *cuaa*, *cuaaza* 'Wissen, Verständnis, Erkenntnis' (H32:56) : OGU *ayquaa* 'lo sé' (R22:484) : KAA *ukwa* 'ele sabe' (KK07:190) : KAG *-kwahav* 'know a thing, know how to do, understand' (B12:149)

Comments: On Ka'apor *ukwa*, morphologically *u-kwa*, note that secondary labialization of *k* by a preceding *u* is a synchronically active process in Ka'apor phonology.

DIG **-iβi-koj*, **-iβi-kōj*, **-iβi-koĩ* (??)

TUP *Acigbigcoi* 'Cauar' (VLB, I, 69) : TEN *u-z-àwykàz*, *u-h-àwykàz* 'cavar a terra (para plantar)' (HH13:190) : TOC *-ywykáj* 'cavar terra' (CR03:266) : KAM - : WAJ *ɔ-pikōy* 'creuser' (G89:54) : KAY *-ywykai* 'cavocar, fazer um buraco no chão, cavar' (W05:121) : GUY *azuguicoĩ* 'Ich häufe Erde an um eine Pflanze, schüte sie an' (H32:352) : OGU *ahĩbicoy* 'cavar la tierra' (R22:163) : KAA *-pyküi* 'cava (ele)' (KK07:158) : KAG *-yvykoi* 'dig' (B12:292)

Comments: For Kamayurá, Seki (2000, p. 219) gives *-jo'ok* 'cavar'. Note that the initial vowel in what is likely an incorporated root **iβi* 'earth, soil' is unstable. It changes to *u* in Guarayo (though a variant with the etymological *i* < **i* does exist: cf. Hoeller, 1929, p. 150), it has a variant with e in Old Guarani (noted by Restivo, 1893 [1722], p. 163) and it shows up as *ə* < *à*> in Tenetehára. The schwa in the Tenetehára cognate also relates to the issue of whether nasality was present or not in the etymon, a fact suggested by the Wajãpi and Guarayo cognates. Finally, see that Wajãpi *-pikōy* and Ka'apor *-pyküi*, while plausibly cognate, call for some explanation for the surprising change **β* > *p*. An association with *-ipi* 'bottom, depth' (Grenand, 1989, p. 60) could be the folk-etymological source for this change, or these reflect an independent formation with **-pi* 'bottom, depth'.

BURN (INTR.) **-kaj*

TUP *Acay* 'arder' (VLB, I, 40) : TEN *u-kaz* 'queimar-se' (HH13:157) : TOC *-káj* 'queimar' (CR03:94) : KAM *kaj* 'queimar-se' (S00:459) : WAJ *ɔ-kay* 'brûler' (G89:51) : KAY *-kai* 'queimar (sozinho), arder' (W05:48) : GUY *acai* 'Ich brenne mich' (H32:45) : OGU *Caĩ* 'quemadura', *Acaĩ* 'yo me quemo' (M39:86) : KAA *ukwái* ~ *ukái* 'queima (3sg.)' (KK07:188) : KAG *-kai* 'burn, burn oneself' (B12:130)

EAT (INTR.) **-karu*

TUP *Acarú* 'Comer' (VLB, I, 77) : TEN - : TOC *-karó* 'comer' (CR03:99) : KAM *karu* 'comer' (S00:459) : WAJ - : KAY *-ka'ru* 'mexer, revirar; mastigar' (W05:49) : GUY *acaru* 'ich esse' (H32:48) : OGU *acaru* 'comer' (R22:169) : KAA - : KAG -

SWALLOW **-mokõn*

TUP *aimocon* 'tragar' (VLB, II, 134) : TEN *-mukun* 'engolir' (HH13:165) : TOC *mokón* 'engolir' (CR03:142) : KAM *mokon* 'engolir' (S00:462) : WAJ *mokũ* 'avaler' (G89:296) : KAY *-mokon* 'engolir' (W05:65) : GUY *moco*, *amoco* 'Ich verschlucke es' (H32:131) : OGU *amocô* 'tragar' (R22: 523) : KAA *mokon* ~ *mokõ* 'engole' (KK07:125) : KAG *-mokon* 'swallow' (B12:141)

Comments: Danielsen et al. (2019) note, for Guarayu, the presence of nasalization (*amokõ*) that is not recorded in Hoeller's data.



Two **mokōj*

TUP *Mocōy* 'dous, ou duas' (VLB, I, 106) : TEN *mokoz*, *mukuz* 'dois' (HH13:110) : TOC *mokój*, *mokósa* 'dois' (CR03:142) : KAM *mokōj* 'dois' (S00:462) : WAJ *məkōy* 'deux' (G89:296) : KAY *mukūi* 'dois' (W05:75) : GUY *mokōi* *mokōi* 'pairwise' (DST19) : OGU *mocol* 'dos' (R22:255) : KAA *mokōi* 'dois' (KK07:125) : KAG *mokonha* 'two' (B12:174)

WOMAN **kujã*

TUP *Cunhã* 'Mulher, mulier' (VLB, II, 40) : TEN *kuzà* 'mulher' (HH13:93) : TOC *kosó*, *kosóa* 'mulher, fêmea' (CR03:109) : KAM *kujã* 'mulher' (S00:460) : WAJ *kuyã* 'soeur' (G89:253) : KAY *kūjã* 'mulher, fêmea' (W05:52) : GUY *cuña* 'Weib, Frau' (H32:63) : OGU *cuñã* 'muger' (R22:394) : KAA *kunjã* 'a senhora, a mulher' (KK07:50) : KAG *kunha* 'woman, married woman' (B12:146)

HUSBAND'S SISTER **-ukeʔi*

TUP *Ukēi*, *xe ukēi* (A86:273-4) : TEN *ukē'i* 'cunhada da irmã do marido' (B78:282) : TOC *-oke'ia* 'mulher do irmão (de mulher)' (CR03: 168) : KAY *-uki'i* 'cunhada da mulher' (W05:109) : KAM *-uke'i* 'esposa do irmão; irmã do marido' (S00:393) : WAJ *ukeʔi* 'belle-soeur' (G89:453) : OGU *che-uquey* 'cuñada, hermana de su marido' (R22:202)

Comments: Kayabí *ukiʔi* is possibly a Kagwahiva loan, as it shows *e > i. Note that there is no general harmonizing rule in Kayabí historical phonology that could account for this.

ELDER BROTHER **-t-iket-ʔit*

TUP *Tigueigra* (VLB, I, 14) : TEN : *Tikê'ir* 'irmão mais velho' (B78:267) : TOC *-yke'ýt* 'irmão mais velho de homem' (CR03:259) : KAM *-ryke'yt* (S00:391) : KAY *-eki'yt* 'irmão mais velho (do homem)' (W05:21) : GUY *tiquieir* 'mein älterer Bruder, sagt der Mann' (H32:254) : WAJ *ε-łekiʔi* (G89:60) : OGU *Tiqueira* (M39:392) : KAG *-reky'yra* 'Elder brother' (P96:66)

Comments: Wajãpi, Kagwahiva and Kayabí show vowel metathesis: **-t-iket* > KAY *-reki-ʔit* : WAJ *-łekiʔi*.
