

# Macabéa

Revista Eletrônica do Netli, Volume 8, Número 2, Jul.-Dez., 2019

## ESTUDOS DIACRÔNICOS DE LÍNGUAS INDÍGENAS BRASILEIRAS: UM PANORAMA



### DIACHRONIC STUDIES OF THE INDIGENOUS LANGUAGES OF BRAZIL: A PANORAMA

ANDREY NIKULIN  
UNIVERSIDADE DE BRASÍLIA, Brasil

FERNANDO O. DE CARVALHO  
UNIVERSIDADE FEDERAL DO AMAPÁ, Brasil

[RESUMO](#) | [INDEXAÇÃO](#) | [TEXTO](#) | [REFERÊNCIAS](#) | [CITAR ESTE ARTIGO](#) | [O AUTOR](#)

RECEBIDO EM 09/04/2019 • APROVADO EM 13/05/2019

---

#### Abstract

---

This study has a two-fold purpose: to provide the reader with a panorama of the state of the art in diachronic studies of the Indigenous languages of Brazil and to promote a rigorous application of known, accepted methods employed by historical linguistics to these languages. We discuss at some length issues such as the *proof of language relationship*, the *internal classification*, the *phonological* and *syntactic* reconstruction, as well as the *philological studies* that aim at detecting diachronic changes. Special attention will be given to the comparative method.

---

## Resumo

---

O artigo possui o duplo objetivo de introduzir o leitor a um panorama do estado da arte em estudos diacrônicos das línguas indígenas brasileiras, além de promover uma aplicação rigorosa dos métodos já consolidados da linguística histórica a essas línguas. Para tanto, abordamos separadamente os problemas de *demonstração de parentesco genético* entre as línguas, de *subagrupamento*, de *reconstrução fonológica e sintática* de proto-línguas e de *estudos filológicos* que visam a detectar mudanças linguísticas. Atenção especial será trazida ao método histórico-comparativo.

---

## Entradas para indexação

---

**KEYWORDS:** Indigenous languages. Historical linguistics. Comparative method.

**PALAVRAS CHAVE:** Línguas indígenas; Linguística histórica; Método histórico-comparativo.

---

## Texto integral

---

### INTRODUÇÃO

Em termos de diversidade linguística, o Brasil é um dos países que mais se destaca no mundo. Estima-se que o território nacional abriga mais de cento e cinquenta<sup>2</sup> línguas originárias (isto é, desconsiderando-se aquelas cuja presença no continente deve-se à colonização europeia e a processos migratórios posteriores), subdivididas em pelo menos dezenove unidades genéticas distintas. Entretanto, apesar do alto grau de diversidade linguística do país, os estudos diacrônicos de suas línguas ainda são incipientes em comparação com as línguas de algumas outras regiões, tais como a África ou a Eurásia.

Este artigo possui um duplo objetivo. Em primeiro lugar, o de oferecer um panorama (necessariamente seletivo) dos estudos já desenvolvidos na área da linguística histórico-comparativa das línguas indígenas faladas no Brasil. Em segundo lugar, o de promover uma aplicação mais rigorosa dos métodos clássicos e eficientes, amplamente utilizados nas pesquisas em linguística histórica, às línguas indígenas brasileiras. Para tal fim, discutiremos criticamente uma amostra dos estudos incluídos no nosso panorama, com a atenção focada a aspectos metodológicos de importância central na aplicação do método comparativo.

O restante deste artigo está estruturado da seguinte maneira. Na **primeira seção** (**'Classificação genética'**), abordaremos a *classificação genética* das línguas indígenas brasileiras, dedicando subseções separadas aos problemas da demonstração da *existência* do parentesco entre línguas (**primeira subseção: 'Demonstração de parentesco'**) e da determinação do *grau* de parentesco, isto é, da identificação de relações genéticas mais e menos profundas dentro de famílias e troncos já delimitados (também chamada de *classificação interna*; **segunda**

**subseção**). Trataremos da reconstrução de elementos de proto-línguas na **segunda seção** (**'Reconstrução de proto-línguas: método comparativo'**), discutindo tanto a reconstrução fonológica (**primeira subseção**) como a reconstrução sintática (**segunda subseção**). A **terceira seção** (**'Rastreamento de mudanças linguísticas: método filológico'**) trará um panorama dos estudos diacrônicos das línguas indígenas brasileiras baseados em dados filológicos. Concluiremos o artigo com a **quarta seção** (**'Conclusão'**), em que resumimos os pontos principais levantados em seções anteriores.

## CLASSIFICAÇÃO GENÉTICA

A tarefa de classificação genética das línguas do mundo consiste na identificação de conjuntos de línguas que possuem uma origem genética comum. Como resultado dessa atividade, a linguística histórica oferece informação diretamente relevante para a investigação da pré-história da humanidade, uma vez que a identificação de uma ancestralidade comum entre línguas permite a inferência de que as populações que as falam partilham também de uma história em comum. Nesse sentido, os estudos em classificação das línguas complementam os resultados de diversas outras disciplinas, da arqueologia e da genética à antropologia e à mitologia comparada.

O problema colocado diante dos linguistas pode ser sintetizado da seguinte maneira: do conjunto de todas as similaridades lexicais (ou seja, aquelas observadas no vocabulário), nos elementos gramaticais, nos padrões fonológicos, morfológicos e sintáticos, das línguas do mundo, quais são indicadoras de que essas línguas possuem um parentesco, uma ancestralidade comum? Quais, por outro lado, resultam apenas de contatos históricos entre seus falantes, quais são meramente fortuitas e quais podem ser atribuídas a regularidades tipológicas na estrutura dos sistemas gramaticais? O linguista histórico tem diante de si a tarefa de discriminar entre os diferentes tipos de similaridades encontradas em função da sua origem provável, identificando aquelas que constituem evidência de uma origem comum.<sup>3</sup> A partir dessas evidências, é possível afirmar que determinadas línguas pertencem a um mesmo agrupamento genético (na tradição acadêmica brasileira, os agrupamentos mais profundos e mais diversos são conhecidos como *troncos* e os menos profundos e menos diversos como *famílias*, embora a diferença entre esses dois termos não seja rígida).

A linguística histórica não apenas procura estabelecer se há relação de parentesco genético entre determinadas línguas, mas também busca resolver o problema da *classificação interna* (ou *subagrupamento*) das unidades genéticas previamente identificadas a fim de reconstruir, no maior detalhe possível, o processo de diversificação das línguas ancestrais (proto-línguas) de tais unidades.

Estudos linguísticos demonstraram que algumas das línguas indígenas brasileiras possuem uma origem comum, formando pelo menos doze famílias e troncos: **Macro-Jê, Tupi, Aruak, Caribe, Tukano, Pano-Tacana, Nadahup, Txapakura, Nambikwara, Arawá, Yanomami** e **Guaicuru** (desta última família,

apenas a língua Kadiwéu é falada no Brasil).<sup>4</sup> Outras línguas, tais como o Tikuna e o Pirahã, pertencem a famílias muito pequenas (**Tikuna-Yuri** e **Mura-Pirahã**), sendo as únicas línguas sobreviventes das mesmas. As relações genéticas externas de algumas outras línguas e pequenos grupos de dialetos, tais como Mýky/Irantxé, Kanoê, Kwazá, Aikanã, Trumai, Arutani e o Maku de Roraima, não foram identificadas até o presente, razão pela qual estas línguas são conhecidas como *línguas isoladas*. O *status* de algumas línguas e pequenas famílias ainda é debatido. Por exemplo, de acordo com alguns pesquisadores, as línguas Bororo, Kariri, Yaathê, Guató e Puri integrariam o tronco linguístico Macro-Jê, enquanto outros autores não as incluem neste tronco (para uma discussão, ver RIBEIRO, VOORT, 2010, seção ‘Rastreamento de mudanças linguísticas: método filológico’). O complexo dialetal Katukina-Kanamari poderia formar parte da família Katukina-Harakmbet (ADELAAR, 2000); contudo, ainda não há consenso quanto à validade dessa proposta.

Nas próximas seções apresentaremos o estado da arte nos estudos relacionados à classificação genética das línguas indígenas brasileiras, abordando a demonstração de parentesco na primeira subseção e o problema de subagrupamento na segunda subseção.

## DEMONSTRAÇÃO DE PARENTESCO

As primeiras hipóteses relacionadas à classificação genética das línguas da América do Sul remontam ao trabalho de Gilij (1780–1784), que identificou duas famílias importantes do continente: a família Maipure (hoje conhecida como Aruak) e a família Caribe. Desde Gilij até a primeira metade do século XX, o problema da classificação das línguas sul-americanas foi abordado por diversos outros autores, tais como Lucien Adam, Paul Rivet, Čestmír Loukotka e J. Alden Mason (ver ADAM, 1890; RIVET, 1924; LOUKOTKA, 1963, 1968; MASON, 1950, entre muitos outros trabalhos). Contudo, esses estudos eram limitados pela extrema escassez da documentação das línguas do continente. Além de serem poucos em quantidade, os registros aos quais os autores supracitados tinham acesso foram, em sua maioria esmagadora, elaborados por viajantes e naturalistas europeus que nunca tiveram a oportunidade de trabalhar detidamente com as línguas em questão, impossibilitando, assim, uma análise fonológica, morfológica e sintática segura dos dados por eles registrados. Esses fatores, naturalmente, constituíram um impedimento incontornável para as pesquisas em linguística histórica.

A situação se reverteu com os avanços na documentação das línguas indígenas sul-americanas na segunda metade do século XX, e deveu muito ao trabalho pioneiro de Aryon Dall’Igna Rodrigues. Dois dos principais agrupamentos linguísticos presentes exclusiva ou majoritariamente no Brasil, as famílias Tupi e Macro-Jê, tiveram as suas primeiras caracterizações mais explícitas no trabalho deste autor, que os definia como “troncos linguísticos”.<sup>5</sup> Utilizando métodos mais confiáveis do que aqueles empregados por outros pesquisadores, como a

“comparação multilateral” de Joseph Greenberg (ver RANKIN, 1992; MATISOFF, 1990; CAMPBELL, 1988 para críticas deste procedimento), e baseando-se em mais e melhores dados do que aqueles disponíveis, por exemplo, para a classificação de Loukotka (1968), Rodrigues pôde, ao longo de décadas de trabalho, propor um delineamento básico para os dois troncos em questão.

Qualquer proposta de parentesco genético entre duas ou mais línguas, para ser comprovada, tem de atender aos seguintes critérios:

- (1) a identificação de um conjunto de correspondências sonoras (necessariamente recorrentes) entre as línguas a serem comparadas;
- (2) a identificação de um conjunto de prováveis cognatos entre as línguas em questão, isto é, de itens lexicais tais que
  - (a) apresentem as correspondências sonoras supracitadas,
  - (b) possuam semântica idêntica ou muito próxima,
  - (c) pertençam ao vocabulário básico (isto é, não cultural);<sup>6</sup>
- (3) a ausência de hipóteses concorrentes mais plausíveis.

Muitos autores que tratam da questão da demonstração de parentesco genético acrescentam a essa lista mais um critério:

- (4) a identificação de um conjunto de semelhanças morfológicas (idealmente paradigmáticas).

Em nossa opinião, embora a existência de paralelos morfológicos seja de grande valor para a corroboração das propostas de parentesco, este critério não pode ser considerado universal. Primeiramente, casos de erosão morfológica rápida são bem atestados na história das línguas do mundo (por exemplo, na história recente do inglês, do francês moderno e do africâner), mostrando que sistemas morfológicos complexos podem ser perdidos ou reestruturados em um período relativamente breve. Além disso, o critério morfológico não é aplicável no caso de línguas isolantes (por exemplo, no caso de algumas línguas da família Nadahup) pelo simples motivo de essas línguas quase não possuírem morfologia afixal. Isto é, embora a presença de correspondências morfológicas, em especial quando apresenta organização paradigmática, constitua uma evidência forte para que se reconheça a existência de um vínculo genético entre duas ou mais línguas (ou famílias), a *ausência* deste tipo de similaridade não indica que as línguas ou famílias sob comparação não possuam tal vínculo, nem servem de indicação de que investigações adicionais que examinem dados, por exemplo, de natureza lexical, estejam necessariamente fadadas ao fracasso. Para uma discussão maior sobre a superioridade do critério lexical em relação ao critério morfológico, referimos o leitor a Dybo e Starostin (2008, p. 124–138).

A Tabela 1 apresenta uma lista, com breves comentários, de trabalhos usualmente apontados como aqueles que primeiro apresentaram evidências de parentesco entre duas ou mais línguas indígenas brasileiras (ou entre famílias inteiras). Além do pioneirismo das contribuições, procuramos selecionar aqueles que mais se aproximam do ideal exposto acima, no que diz respeito ao tipo de evidência e argumentação apresentados em favor da hipótese de parentesco em questão.<sup>7</sup>

Tabela 1. Seleção de trabalhos que demonstram as relações genéticas entre as línguas indígenas brasileiras

obra	línguas ou famílias comparadas	agrupamentos cuja existência é demonstrada	comentário
GILIJ, 1780–1784	Maipure, Guipunave, Avare, Cavere (Cabre), Mojeño	Aruak (parcialmente)	trabalho baseado em línguas Aruak do Orinoco e também no Mojeño, da bacia do Mamoré (atual Bolívia), contém uma comparação cuidadosa de elementos lexicais e gramaticais
	Tamanaco, Avaricoto, Pajure, Galibi	Caribe (parcialmente)	trabalho baseado apenas em um subconjunto das línguas Caribe faladas na bacia do Orinoco (atual Venezuela), contém uma comparação cuidadosa de elementos lexicais e gramaticais, incluindo o reconhecimento de correspondências regulares
DE LA GRASSERIE, 1890	Pano, Conibo, Pacavara, Maxuruna, Caripuna, Mayoruna Domestica, Mayoruna Fera	Pano	comparação de dados lexicais de campos semânticos diversos (partes do corpo, flora, fauna, parentesco) e de classes lexicais diversas, incluindo “verbos” e “adjetivos”; o autor indica ainda a existência de correspondências sonoras regulares (ver OLIVEIRA, 2014 para a reconstrução mais atual e compreensiva do Proto-Pano)
BRINTON, 1891	Arawá, Paumari	Arawá (parcialmente)	Brinton (1891) foi o primeiro a falar em um “tronco ( <i>stock</i> ) linguístico Arawá”, baseado em dados do Paumari ( <i>Pammary</i> ) e Arawá ( <i>Araua</i> ); ver Dixon (2004) para uma reconstrução do Proto-Arawá
ADAM, 1899	Abipón, Mocovi, Toba, Mbayá	Guaicuru	embora a existência da família Guaicuru fosse conhecida desde o século XVIII, a compilação produzida por Lucien Adam é a primeira investigação sistemática de similaridades entre as línguas em questão, mais completa do que trabalhos anteriores como, por exemplo, o de Lafone Quevedo (1896)
CRÉQUI-MONTFORT, RIVET, 1913	Tapakura, Kitemoka, Wanyam, Moré, Napeka, Rokorona	Txapakura	primeiro trabalho extenso das línguas da família Txapakura (cuja existência já havia sido proposta por Alcide d’Obrigny e Alexander F. Chamberlain); não se aplica o método comparativo, mas as línguas são próximas entre si o suficiente para não deixar dúvida quanto a sua origem comum



DAVIS, 1966	Apinajé, Kĩsêdjê, Canela, Xavante, Kaingang	Jê	primeira demonstração convincente do parentesco entre o Kaingang e as demais línguas Jê (ver NIKULIN, 2019 para uma reconstrução atualizada)
DAVIS, 1968	Jê, Maxakali, Karajá	Macro-Jê (parcialmente)	ver Ribeiro (2012) para maiores evidências em favor da inclusão do Karajá, Nikulin e Silva (2019) para maiores evidências em favor da inclusão do Maxakali
GIRARD, 1971a	línguas Caribe	Caribe	primeira demonstração formal do parentesco; ver Meira e Franchetto (2005) e Meira et al. (2010) para uma reconstrução atualizada
MIGLIAZZA, 1972	Yanam/Ninam, Sanima, Yanomam, Yanomami	Yanomami	embora o parentesco entre as línguas/variedades que formam a família Yanomami seja óbvia, o trabalho de Migliazza (1972) oferece a primeira exploração sistemática da diferenciação entre os membros dessa pequena e pouco diversa família; sua divisão em quatro línguas/variedades básicas parece ter se mantido, ao menos como referência, na literatura posterior sobre esse agrupamento
RODRIGUES, 1980	Munduruku, Tupi-Guarani	Tupi (parcialmente)	primeira demonstração formal do parentesco, confirmando a proposta de Rodrigues (1958)
PAYNE, 1991	24 línguas Aruak	Aruak	primeira demonstração formal do parentesco entre as línguas Aruak na acepção moderna dessa família (ver MATTESON, 1972 para uma tentativa anterior que envolvia línguas que hoje não são reconhecidas como membros da família)
ADELAAR, 2000	Katukina, Harakmbet	Katukina-Harakmbet	apresenta 40 comparações de elementos lexicais, quase que exclusivamente nominais, alguns dos quais apresentam apenas cognatos parciais; há o apontamento inicial de correspondências segmentais
RODRIGUES, 2002	Tupari, Tupi-Guarani	Tupi (parcialmente)	primeira demonstração formal do parentesco, confirmando a proposta de Rodrigues (1958)
SEKI, 2002	Jê, Krenak	Macro-Jê (parcialmente)	ver Nikulin e Silva (2019) para maiores evidências em favor da inclusão do Krenak
RODRIGUES, 2005, 2007	Tupi-Guarani, Aweti, Mawé, Munduruku, Juruna, Tupari, Arikém, Mondé, Ramarama, Puruborá	Tupi	primeira demonstração formal detalhada do parentesco, confirmando a proposta antecipada pelo mesmo autor (RODRIGUES, 1958)
ADELAAR, 2008	Chiquitano, Jê e outras línguas Macro-Jê	Macro-Jê	primeira demonstração convincente do parentesco entre o Chiquitano e as demais línguas Macro-Jê

DIETRICH, 2009	Karitiana, Tupi-Guarani	Tupi (parcialmente)	primeira demonstração formal do parentesco, confirmando a proposta de Rodrigues (1958)
RIBEIRO, VOORT, 2010	Jabuti, Jê e outras línguas Macro-Jê	Macro-Jê	primeira demonstração convincente do parentesco entre a família Jabuti e as demais línguas Macro-Jê, confirmando a hipótese de Nimuendajú (2000 [1935])
EPPS, BOLAÑOS, 2017	Hup, Yuhup, Dâw, Nadëb	Nadahup	embora o fato de as quatro línguas Nadahup serem relacionadas tenha sido estabelecido ainda no começo do século XX por Th. Koch-Grünberg, o trabalho de Epps e Bolaños (2017) é o primeiro a demonstrar que não há evidências suficientes para agrupá-las com as línguas Kákua, Nìkak, Puinave e Hodì, como se fazia em várias classificações anteriores

Quando se trata de unidades genéticas de alta profundidade temporal, torna-se imprescindível a utilização de reconstruções intermediárias (isto é, de reconstruções de subagrupamentos de baixa profundidade, sobre os quais vide a subseção ‘Classificação interna’) para a demonstração do parentesco: a quantidade das evidências positivas invariavelmente diminui com o aumento da distância temporal por diversas razões (perda lexical, acúmulo de mudanças sonoras, semânticas e gramaticais). Para melhor ilustrar isto, providenciamos na Tabela 2 abaixo alguns dados lexicais das línguas Xavante e Maxakali que, como se sabe hoje, pertencem ao tronco Macro-Jê.

Tabela 2. Alguns dados lexicais das línguas Xavante e Maxakali.

glosa	filho	dente	barriga	fome	semente	cabeça	osso	cinza
<b>Xavante</b>	<i>'ra</i> /ʔra:/	<i>'wa</i> /ʔwa/	<i>di('i)</i> /niʔ/	<i>mra(m)</i> /mrãṃ/	<i>nhama/dzö</i> /jãm/	<i>'rã(i)</i> /ʔrãj/	<i>hi</i> /hi/	<i>'ru(i)</i> /ʔruj/
<b>Maxakali</b>	<i>kutok</i> /ktuk/	<i>xox</i> /cuc/	<i>tex</i> /tek/	<i>putup</i> /ptip/	<i>xap</i> /cap/	<i>putox</i> /ptuc/	<i>kup</i> /kip/	<i>putohok</i> /ptuk/

Obs: os dados da Tabela 2 estão listados em sua representação ortográfica e fonológica. Em Xavante, algumas palavras possuem alomorfes diferenciados conforme o contexto prosódico.

É possível observar algumas semelhanças entre os dados dessas duas línguas. Um leitor atento poderia hipotetizar, a partir dos dados apresentados na Tabela 2 acima, que diversas oclusivas do Maxakali (*k-*, *c-*, *p-*) poderiam corresponder a Xavante ʔ- antes de consoantes (*ktuk* ~ *ʔra:*, *cuc* ~ *ʔwa*, *ptuc* ~ *ʔrãj*, *ptuk* ~ *ʔruj*; exceção: *ptip* ~ *mrãṃ*) e buscar explicações para as demais correspondências (Xavante *a* ~ Maxakali *u*; Xavante *r* ~ Maxakali *t*). Entretanto, se um investigador postulasse uma relação genética entre o Xavante e o Maxakali **apenas** com base nos dados na Tabela 2, provavelmente sua hipótese não encontraria ampla aceitação na comunidade acadêmica.



Felizmente, ambas línguas integram famílias relativamente ramificadas (o Xavante pertence ao subramo Central do ramo Cerrado da família Jê; o Maxakali é parte de um mesmo ramo com o Krenak), permitindo uma comparação entre as respectivas proto-línguas. Na Tabela 3, detalhamos a evolução reconstruída das formas apresentadas na Tabela 2 acima (as formas cognatas são destacadas em negrito).

Tabela 3. A evolução de alguns dados lexicais das línguas Xavante e Maxakali.

glosa	filho	dente	barriga	fome	semente	cabeça	osso	cinza
<b>Xavante</b>	<b>'ra</b> /ʔra:/	<b>'wa</b> /ʔwa/	<b>di(i)</b> /niʔ/	<b>mra(m)</b> /mrãĩm̃/	<b>nhama/dzö</b> /ʔam/	<b>'rã(i)</b> /ʔrãj/	<b>hi</b> /hi/	<b>'ru(i)</b> /ʔruj/
<b>Proto-Jê Central</b>	<b>*kra:</b>	<b>*kwa</b>	<b>*di(ki)</b>	<b>*mrã(m)</b>	<b>*ɲãmã/*jɶ</b>	<b>*krã(j)</b>	<b>*hi</b>	(?)
<b>Proto-Cerrado</b>	<b>*kra</b>	<b>*jwa</b>	<b>*tik</b>	<b>*prãĩm̃</b>	<b>*cim</b>	<b>*krãj</b>	<b>*ci</b>	<b>*mbrɔ</b>
<b>Proto-Jê</b>	<b>*kra</b>	<b>*jɔ</b>	<b>*tik</b>	<b>*prãĩm̃</b>	<b>*cim</b>	<b>*krĩɲ</b>	<b>*ci</b>	<b>*mbrɶ</b>
<b>Proto-Maxakali-Krenak</b>	<b>*kruk</b>	<b>*juɲ</b>	<b>*tek</b>	<b>*prim</b>	<b>*jam</b>	<b>*krĩɲ</b>	<b>*jet ~ *jek</b>	<b>*proɲ</b>
<b>Maxakali</b>	<b>kutok</b> /ktuk/	<b>xox</b> /cuc/	<b>tex</b> /tek/	<b>putup</b> /ptip/	<b>xap</b> /cap/	<b>putox</b> /ptuc/	<b>kup</b> /kip/	<b>putohok</b> /ptuk/

Obs. As reconstruções acima são de Nikulin (2017, para o Proto-Jê Central e o Proto-Cerrado; 2019, para o Proto-Jê) e Nikulin e Silva (2019, para o Proto-Maxakali-Krenak). As reproduzimos aqui com algumas modificações referentes ao sistema de transcrição.

Uma breve inspeção das formas reconstruídas apresentadas na Tabela 3 é suficiente para o investigador perceber o alto grau de semelhança superficial entre as línguas das famílias Jê e Maxakali-Krenak. É fácil ver que, com o aumento da distância temporal, a quantidade das evidências inequívocas de parentesco diminui. Nos primeiros cinco pares de exemplos, isto ocorre em razão do acúmulo das mudanças sonoras (tais como as mudanças \*ɔ > \*wa em Proto-Cerrado, \*jw > \*kw em Proto-Jê Central, \*k > ʔ em Xavante; \*r > t em Maxakali). Nos itens 'cabeça' e 'osso', o Maxakali sofreu uma perda lexical (os itens \*krĩɲ 'cabeça' e \*jet ~ \*jek 'osso' são preservados apenas em Krenak, onde seus reflexos possuem a forma *krên*, *jek*, respectivamente). Já a forma \*mbrɔ 'cinza' do Proto-Cerrado, embora seja preservada em Xavante, sofreu uma mudança semântica: seu reflexo Xavante *-pro* /-prɔ/ ocorre em palavras compostas e denota substâncias sólidas reduzidas a partículas tênues. Portanto, a suposta correspondência entre Maxakali *p-* e Xavante ʔ-, discutida acima, é **fictícia**: os dois pares de palavras em que ela se baseava, após uma análise mais detalhada, resultam não serem verdadeiros cognatos.

Finalmente, chamamos a atenção do leitor ao fato de existirem casos em que uma língua passa por uma grande quantidade de mudanças linguísticas, dificultando a identificação da língua como um membro da família já identificada.

Ao discutir o caso da língua Xetá, Rodrigues (1978, p. 8) atenta para os fatores que contribuíram para que outros pesquisadores (GUÉRIOS, 1959; LOUKOTKA, 1960) duvidassem da classificação dessa língua como um membro do subramo Guarani:

- 1) mudanças fonológicas que tornaram algumas palavras não imediatamente reconhecíveis como afins a suas correspondentes em guarani <...>: **a'wāci** *neblina*, em guarani (g.) **ibi'tĩ**; **ñocha** *espécie de palmeira*, g. **yu'yi**; **wa'rēke** *bebida (de coquinho)*, g. **ĩ'ba ri'kwe** *suco de fruta*; **há'wiča** *grande*, g. **hubi'cya**; **rāiča** *frio*, g. **ro'ʔi'ca**;
- 2) substituição de nomes tabuizados por outras designações, geralmente locuções descritivas; p. ex. **'haikā 'pīrītai** *caça/carne malhada*, em lugar de **'ñagwa** *onça*, que corresponde ao guarani **ya'wa(re'te)**, **'haikā čape'ai** *caça/carne que tem casca*, em lugar de **'tato** *tatu*, que corresponde ao g. **ta'tu**; **'haikā 'huuai** *caça/carne que faz huu*, em lugar de **ka'piwai** *capivara*, que corresponde ao g. **kapi'ʔi'ba**;
- 3) substituição de nomes simples por locuções metafóricas que aludem a fatos míticos; p. ex.: **'ñane 'čape ta'kiē** *nossa luz, irmão mais velho* para o sol, em g. **kwara'hi**; **'ñane 'čape 'tēwē** *nossa luz, irmão mais moço* para a lua, em g. **ya'ci**; **ta'tōka** *aldeia das estrelas* para o céu, em g. **ĩ'bag**; **ta'tōka ma'ŋetai** *introdutor (?) do céu* para o gavião, em g. **tawa'to**;
- 4) substituição de nomes simples ou complexos por derivados ou compostos descritivos; p. ex.: **haikā'ča** (**haikā+eči+a**) *assador de caça/carne*, em lugar de **'ata** *fogo*, que corresponde ao g. **a'ta**; **'ñane 'tata** *nossas estrelas*, cujo segundo elemento corresponde ao segundo componente do g. **ya'ci-tat'a** *estrela*; **ñepra'ka** *o colhedor* (g. **yepora'ka'ha**) para o machado, em g. **yi**; **či mi'rata** *que eu faço andar comigo* (g. **cye remiero'a'ta**) para *minha esposa*, em g. **cye remire'ko** (literalmente *a que eu faço estar/viver comigo*); **porō'ha** *o pizador* [sic] (g. **pī'rū'ha**) para o pé, em g. **pi**; **mā'ha** *o vedor* (g. **ma'ʔē'ha**) para o olho, em g. **e'ca**; **poči'a** *o defecador* (g. **po'ti'ha**) para o ânus, em g. **e'bi'kwa** (literalmente *orifício das nádegas*); **'moi či'waki** *cobra de braços* para o lagarto, em g. **te'yu**;
- 5) substituição de nomes simples ou complexos por outros em consequência de extensões semânticas associadas ou não a mudanças culturais; p. ex.: **ta'pegwa** *esteira para dormir*, em g. **pī'ri(pē'mi)**, mas correspondente ao g. **tatape'kwa** *abano para o fogo*: a mesma peça tem ambas as serventias entre os índios da Serra dos Dourados; **'tagwa** *flauta*, em g. **mi'mi**, mas correspondente ao g. **ta'kwa** *taquara*: as flautas da Serra dos Dourados são feitas de taquara; **ka'nomi** *homem*, em g. **kuima'ʔe**, mas correspondente ao g. **kunumĩ** *menino*; **'ñčoi** *sapo*, em g. **kuru'ru**, mas correspondente ao g. **yu'ʔi rā**. (RODRIGUES, 1978, p. 8)

Outras línguas com perfis similares ao do Xetá, isto é, que, em virtude de drásticas mudanças induzidas tanto por fatores internos (mudanças sonoras, perdas de morfemas gramaticais e substituição lexical) ou externos (contato linguístico), tiveram seu parentesco com um dos grandes agrupamentos linguísticos do Brasil colocadas em cheque incluem o Sirionó para a família Tupi-Guarani, e tanto o Chamicuro quanto o Yanesha' para a família Aruak. Embora seja pouco provável que alguma das línguas que atualmente são consideradas como isoladas (não classificadas) pertença, de fato, a um dos troncos linguísticos já delimitados, tendo sido sujeitas a uma combinação destes processos que dificultam a busca por elementos cognatos, tal possibilidade ainda não pode ser definitivamente descartada.

## CLASSIFICAÇÃO INTERNA

Como já dissemos acima, o problema da *classificação interna* (ou *subagrupamento*) consiste em se determinar, dentro de um conjunto de línguas que já se sabe serem relacionadas por uma origem comum, quais entre elas possuem uma relação de proximidade mais íntima. O ponto crucial é o de que não é necessário (e nem suficiente) que se demonstre que duas ou mais línguas são *mais parecidas* em alguma medida. O que o linguista histórico busca são evidências de *um período comum de desenvolvimento*, isto é, que mostrem que as línguas em questão são continuações de uma língua ancestral (conhecida como *proto-língua intermediária*) que se diversificou em um passado mais recente do que a proto-língua de todo o agrupamento genético. Para demonstrar que determinadas línguas tiveram, no passado, um período comum de desenvolvimento, é necessário identificar um conjunto de *inovações compartilhadas*, ou seja, mudanças linguísticas que ocorreram em um estágio anterior destas (e somente destas) línguas. As inovações se opõem às *retenções* (isto é, elementos linguísticos que não sofreram mudança no percurso da proto-língua até as línguas em questão), que não possuem valor classificatório.<sup>8</sup> É importante observar que para discernir as inovações das retenções, é necessário saber qual é o estado ancestral do fenômeno em questão (isto é, possuir uma reconstrução parcial da proto-língua). Por sua vez, a reconstrução do estado ancestral depende, em grande parte, do subagrupamento da família considerada, resultando em uma certa circularidade do processo. Esse problema não é fácil de se contornar, cabendo ao pesquisador avaliar qual dos possíveis cenários de diversificação é mais compatível com os dados (ver CHACON, LIST, 2015 para uma pioneira tentativa de automatização do processo, em que um dos algoritmos testados alcançou resultados positivos na dupla tarefa de reconstruir a evolução das consoantes das línguas Tukano e de determinar o subagrupamento dessa família linguística).

Em alguns casos a estreita relação entre determinadas línguas é tão transparente que se torna possível agrupá-las de forma preliminar, sem que as inovações que caracterizam o agrupamento sejam identificadas. É desta forma que foram delimitadas as famosas famílias dos troncos Tupi e Macro-Jê.

Um dos produtos da aplicação do método comparativo da linguística histórica é o estabelecimento de uma *cronologia relativa* para as mudanças que afetaram uma ou mais línguas. Isso é possível, entre outras razões, porque algumas mudanças que ocorrem em determinado ponto na história de uma língua criam as condições para que mudanças posteriores ocorram. Assim, no Macuxi, membro da família Caribe falado no estado de Roraima, certas mudanças sonoras alteraram encontros consonantais, mudando, por exemplo, *-tp-* em *-ʔp*. Para que isso ocorresse, no entanto, foi preciso que, primeiro, esses encontros consonantais fossem produzidos como resultado de uma outra mudança de queda das vogais. Um sufixo como *\*-tɨpu* (passado perfectivo) mudou primeiro para *\*-tɨpu*, e apenas depois para *-ʔpu*. A mudança de queda da vogal (no caso, a vogal *i* do sufixo *\*-tɨpu*) precisou ocorrer antes, para que a mudança do encontro consonantal ocorresse (o leitor pode consultar GILDEA, 1995 para uma interessante discussão desse e de outros processos similares em várias línguas da família Caribe). Temos, assim, uma cronologia relativa em que a mudança vocálica precede historicamente a mudança consonantal.

A cronologia relativa dos eventos de diversificação é comumente vista desde a perspectiva do chamado *modelo arbóreo*. Esse modelo parte da premissa de que cada língua (exceto os *pidgins* e as línguas crioulas) descende de exatamente uma língua ancestral, representando o processo de diversificação linguística como uma série de *divergências*, sem admitir a possibilidade de uma *convergência* de duas (ou mais) línguas.<sup>9</sup> Dessa forma, é possível diagramar a cronologia relativa dos eventos de diversificação em um *cladograma*, em que cada nó corresponde a uma língua atestada ou a uma língua ancestral de um dado agrupamento no momento de sua dissolução.

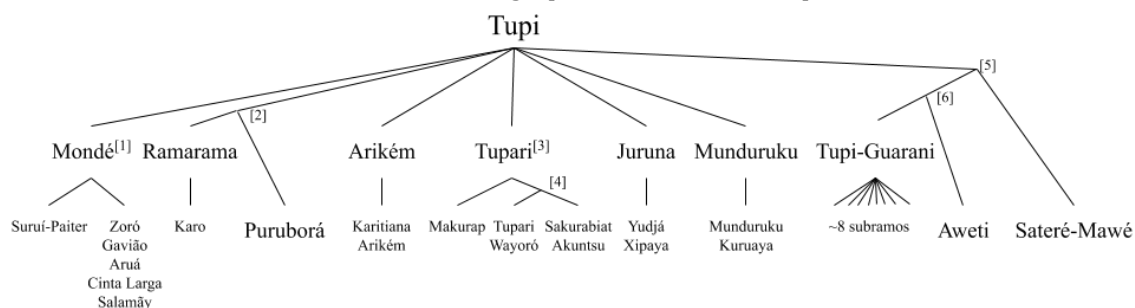
Há estudos que objetivam identificar não apenas a cronologia relativa dos eventos de diversificação, mas também sua *cronologia absoluta*. Em algumas regiões do mundo, tal objetivo pode ser alcançado, em alguns casos, com a utilização de dados extralinguísticos (históricos ou arqueológicos) através da datação dos eventos históricos que subjazem os eventos de diversificação linguística, como, por exemplo, a separação do subramo oriental/balcânico das demais línguas neolatinas pode ser associada com a retirada das legiões romanas em 271 d.C. Entretanto, no caso das línguas indígenas brasileiras não há registros escritos que permitam ancorar os eventos de diversificação linguística (exceto os mais recentes) a eventos históricos, o que implica na necessidade de recorrer a outros métodos de datação. Uma das ferramentas propostas para esse fim é conhecida sob o nome de *glotocronologia* (SWADESH, 1952, 1955). Ela parte da premissa de que existe uma taxa constante de decaimento (perda) lexical, cujo valor pode ser determinado (calibrado) a partir dos dados das línguas com uma história mais bem documentada. Os estudos produzidos nessa linha incluem os de D'Ans (1973, 1975) para a família Pano e Meira (1998, capítulo 8) para o subramo Taranoano da família Caribe, entre outros. No entanto, muitos linguistas não aceitam os resultados dos estudos glotocronológicos como válidos, apontando a alguns problemas nas premissas teóricas envolvidas, bem como à incompatibilidade entre os resultados de aplicação desse método e a realidade histórica em alguns casos específicos.<sup>10</sup>

Outros estudos aplicam métodos quantitativos (filogenéticos) sem se comprometerem com a datação absoluta dos eventos de divergência. Tais estudos, via de regra, se baseiam em dados lexicais, embora haja estudos que contemplam níveis linguísticos diferentes, tais como a fonologia. Os estudos filogenéticos se classificam em *baseados em distâncias* (LIMA-ANGENOT, 1995 para a família Txapakura; DANIËLSEN *et al.* 2011, para a família Aruak; GALUCIO *et al.*, 2015 para o tronco Tupi) e *baseados em caracteres* (RIBEIRO, 2006 para a família Pano; WALKER, RIBEIRO, 2011 para a família Aruak; MICHAEL *et al.*, 2015 para a família Tupi-Guarani; CHACON, LIST, 2015 para a família Tukano; BIRCHALL *et al.*, 2016, para a família Txapakura). No entanto, no estágio atual do desenvolvimento da filogenética não há consenso sobre qual dos métodos existentes gera classificações mais confiáveis, e a maioria dos estudos supracitados visa apenas verificar se os métodos filogenéticos neles utilizados dão resultados compatíveis com as classificações obtidas usando métodos mais tradicionais (não filogenéticos). Por fim, vale notar que a correta aplicação de métodos computacionais depende de uma aplicação prévia do método histórico-comparativo, uma vez que os dados de entrada sobre os quais operam tais métodos, como conjuntos de cognatos ou desenvolvimentos diacrônicos na fonologia e a na morfologia (os chamados “caracteres filéticos”), são parte do *produto* que resulta da aplicação sistemática dos métodos tradicionais. Assim, os métodos computacionais constituem uma ferramenta complementar, que não visa substituir a linguística histórica tradicional e já estabelecida. Essa observação crucial torna mais clara ainda a necessidade de se proceder com cautela ao avaliar os resultados da aplicação destes métodos quantitativos às famílias linguísticas do Brasil e da América do Sul, uma vez que resta ainda muito trabalho a ser feito na aplicação do método histórico-comparativo a esses grupos linguísticos.

Abaixo, a título de exemplo, apresentamos algumas informações acerca do estado atual do conhecimento acerca do subagrupamento de alguns dos grandes troncos brasileiros: Tupi, Aruak e Macro-Jê.

Os estudos mais recentes das línguas do tronco **Tupi** convergem na classificação apresentada no Quadro 1 (não discutiremos aqui o subagrupamento da família Tupi-Guarani).

**Quadro 1.** Subagrupamento do tronco Tupi



As famosas dez famílias constituintes (Mondé, Ramarama, Puruborá, Arikém, Tupari, Juruna, Munduruku, Tupi-Guarani, Aweti, Sateré-Mawé) foram delimitadas

por Aryon Dall’Igna Rodrigues em razão de sua relativa homogeneidade interna, sem a identificação das inovações que as caracterizassem. Posteriores estudos identificaram nós e bifurcações adicionais, identificados com números entre colchetes no Quadro 1 acima, que detalhamos a seguir.

- [1] Divisão binária da família Mondé em dois subramos: Suruí-Paiter e Mondé Nuclear (MOORE, 2005). Dentro do ramo Mondé Nuclear, a língua Salamã é a mais divergente (fato não representado no Quadro 1 por motivos gráficos). Esta divisão é contestada por Anonby (2012).
- [2] Uma proximidade estreita entre a família Ramarama e a língua Puruborá é defendida por Galucio e Gabas Júnior (2002).
- [3] Divisão binária da família Tupari em dois ramos: Makurap e Tupari Nuclear (ANDRADE, a sair; GALUCIO, NOGUEIRA, 2011). Um exemplo de uma inovação compartilhada por todas as línguas Tupari menos o Makurap é a extensão do alomorfe *\*i-* do prefixo de terceira pessoa para os temas da chamada classe II, substituindo o alomorfe *\*c-* (> Makurap *t-*): cf. Proto-Tupari *\*c-aʔip* ‘filho dele’ > Makurap *t-aip* (terceira pessoa de *\*jaʔip* ‘filho’), mas Proto-Tupari Nuclear *\*i-jaʔip* > Sakurabiat *i-taip*, Tupari *i-aʔip*. A situação em Makurap é claramente conservadora, havendo paralelos exatos em outras línguas do tronco Tupi (Munduruku *t-*, Sateré-Mawé *h-*, entre outros cognatos).
- [4] Divisão binária do ramo Tupari Nuclear em dois subramos: Tupari-Wayoró (ANDRADE, a sair) e Sakurabiat-Akuntsu (ANDRADE, a sair; GALUCIO, NOGUEIRA, 2011). O subramo Tupari-Wayoró é caracterizado por claras inovações lexicais, tais como *\*kĩrĩṅã* ‘unha’ e *\*akop-k-ap* ‘fogo’ (substituindo Proto-Tupari *\*mbo-ape* ‘unha’ e *\*ojat* ‘fogo’). O subramo Sakurabiat-Akuntsu é caracterizado por múltiplas inovações lexicais compartilhadas, tais como *\*mãtso* ‘noite’ (substituindo Proto-Tupari *\*ṅgicak*) e *\*kĩcpit* ‘peixe’, além de claras inovações fonológicas, incluindo a epêntese consonantal no ambiente *i\_V* (Sakurabiat *s*; Akuntsu *t*) e um desenvolvimento peculiar de Proto-Tupari *\*t, \*j* (Sakurabiat *s, t*; Akuntsu *ʃ, t*).
- [5] Uma proximidade estreita entre a família Tupi-Guarani e as línguas Sateré-Mawé e Aweti (DRUDE, 2006; CORRÊA-DA-SILVA, 2010; MEIRA, DRUDE, 2015), que formam um agrupamento conhecido na literatura como “Maweti-Guarani” ou “Mawé-Guarani”. Há inovações fonológicas e lexicais claras que o caracterizam, tais como a emergência de uma regra sincrônica de palatalização de um segmento coronal reconstruível como *\*c* quando precedido por *\*i* (*\*cet* ‘nome’ → *\*ui-ćet* ‘meu nome’) ou a substituição de Proto-Tupi *\*tik* ‘muriçoca’ por *\*watiũ*, com o mesmo significado.
- [6] Uma proximidade particularmente estreita entre a família Tupi-Guarani e a língua Aweti (DRUDE, 2006; CORRÊA-DA-SILVA, 2010; MEIRA, DRUDE, 2015), evidenciada por inovações fonológicas, tais como os reflexos

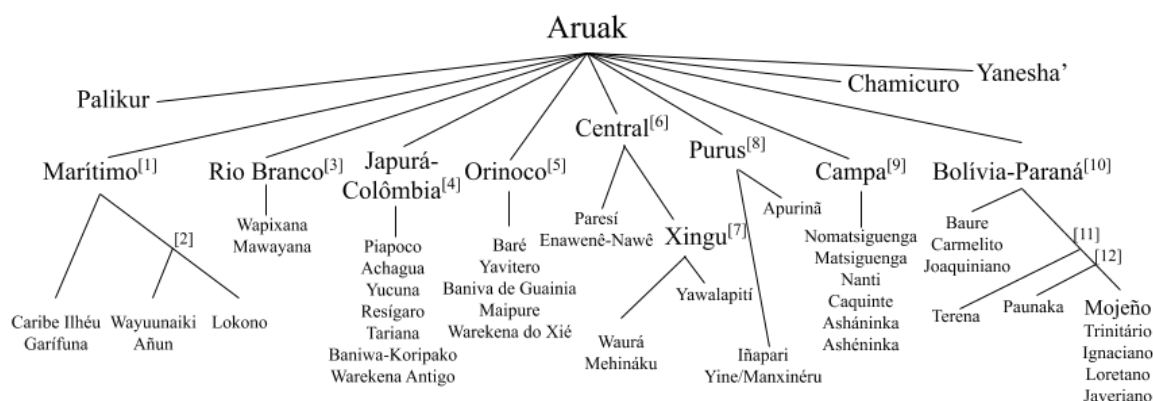


de \*c, \*ć do Proto-Mawé-Guarani entre vogais (em ambos ramos, \*-c- > \*-∅- em ambientes orais, \*-j- em ambientes nasais; \*-ć- > \*-t- em todos os ambientes) ou a mudança semântica dos reflexos de \*wup ('vermelho' → 'amarelo'); em Proto-Aweti-Guarani, o significado 'vermelho' era expresso pelo item \*piðãη<sup>11</sup>).

Até o presente não foi confirmada a validade da proposta de Rodrigues (2007) referente a uma divisão binária do tronco Tupi, informalmente conhecida como a hipótese Oeste/Leste. De acordo com essa proposta, as línguas Tupari, Arikém, Mondé, Ramarama e Puruborá, por um lado, e as línguas Juruna, Munduruku, Mawé, Aweti e Tupi-Guarani, por outro lado, formariam unidades genéticas. A confirmação ou a rejeição (completa ou parcial) da proposta de Rodrigues crucialmente depende da reconstrução do estado ancestral do tronco; entretanto, atualmente não existe uma reconstrução do Proto-Tupi universalmente aceita. A investigação dos autores em curso identificou evidências plausíveis que permitem agrupar as famílias Tupari e Arikém, bem como as famílias Munduruku e Juruna, embora haja também evidências menos sugestivas que corroborariam a hipótese de Rodrigues referente à divisão binária do tronco Tupi em ramos Oeste e Leste.

Para as línguas **Aruak** a maior parte das classificações ancoram-se principalmente sobre a proximidade geográfica entre as línguas (esse é o caso de AIKHENVALD, 1999) ou sobre retenções lexicais, isto é, proporções de cognatos compartilhados entre pares de línguas, em geral computados sobre uma lista prévia de 'significados básicos' ou de étimos reconstruídos. Essa segunda abordagem é a de Payne (1991), que após reconstruir 203 étimos para o Proto-Aruak propõe uma classificação da família Aruak. No Quadro 2 abaixo, detalhamos os agrupamentos mais bem demonstrados, com base em argumentos propriamente linguísticos, incluindo um compartilhamento de similaridades lexicais e gramaticais que, no entanto, permanecem apenas intuitivamente compreendidos. Na ausência de uma reconstrução confiável para o Proto-Aruak, não é possível, no momento, identificar inovações compartilhadas para os ramos/subgrupos provisoriamente identificados. Em virtude da grande diversidade interna da família, aceita-se, atualmente, que a prioridade para o trabalho histórico consiste em aplicar o método histórico-comparativo para os subgrupos individuais, formados por línguas de relação obviamente próxima, e só então prosseguir, com segurança, para os níveis mais inclusivos, alcançando o Proto-Aruak em última instância. Além disso, é pouco provável que o Proto-Aruak tenha se dividido em 11 línguas/ramos descendentes simultaneamente, porém a existência de agrupamentos mais abrangentes nunca foi formalmente demonstrada, uma tarefa que deverá ser assumida em futuras pesquisas.<sup>12</sup>

Quadro 2. Subagrupamento da família Aruak



As propostas de inovações (ou, ao menos, similaridades) que caracterizam os agrupamentos apresentados no Quadro 2 são discutidas a seguir.

- [1] As línguas faladas nas proximidades da costa Caribe e na América Central (e, historicamente, nas ilhas caribenhas), a saber, o Lokono, o Wayuunaiki, o Añun (ou Paraujano), o extinto Taíno e o diversos *letos* do Caribe Insular, são consensualmente classificados como membros de um subgrupo Caribenho (AIKHENVALD, 1999; CAMPBELL, 1997, 2012). Não existe, no entanto, qualquer demonstração de inovações compartilhadas que definam concretamente este subgrupo, além de indicações genéricas, por exemplo, sobre a perda de *\*n* medial nas línguas em questão (Proto-Aruak *\*t̥fina-ru* 'mulher' > Way(uunaiki) *hié-ri*, Lok(ono) *hia-ro*, Car(ibe Insular) *hĩá-ru*; Proto-Aruak *\*-nene* 'língua' > Way *-a-jee*, Lok *-ije*, Car *-ĩeĩe*; PAYNE, 1991, p. 422, 426).
- [2] O ramo constituído pelas línguas Wayuunaiki e Añun (Paraujano), de um lado, e o Lokono, de outro lado, é caracterizado pelo emprego do prefixo da primeira pessoa do singular *\*ta-* (substituindo o prefixo *\*nu-*, presente em quase todas as línguas Aruak). Carvalho (2016d) reconstrói um elemento *\*-ja*, possivelmente de natureza dêitica, que funciona como base para a derivação de pronomes pessoais independentes nas línguas deste subgrupo, e que esse poderia ser, de fato, uma inovação compartilhada desse ramo. Carvalho (2017c) mostra que o numeral *\*kabini* 'três' e uma formação particular para o numeral 'quatro' são também inovações compartilhadas que permitem diferenciar um subgrupo formado do Lokono, do Wayuunaiki e do Añun.
- [3] A mera inspeção de listas vocabulares parece ser suficiente para indicar que o Wapixana e o Mawayana estariam mais próximos um do outro do que de qualquer dos demais membros da família. No entanto, até o momento, não há qualquer identificação explícita de inovações que caracterize o subgrupo Rio Branco.

- [4] Ramirez (2001) propõe a existência de um subgrupo “Japurá-Colômbia” reunindo nove línguas Aruak do noroeste amazônico. O autor propõe algumas inovações fonológicas que seriam compartilhadas pelas línguas em questão. O primeiro problema é que muitas dessas são encontradas (co-ocorrendo) em diversos outros subgrupos, como o enfraquecimento da velar *\*k* diante de vogais anteriores e a perda da oposição entre *\*p* e *\*b* provisoriamente reconstruída para o Proto-Aruak. De maneira mais séria, no entanto, o trabalho de Ramirez (2001) apresenta um sem número de erros metodológicos e de tratamentos dos dados, o que é suficiente para que todas as suas conclusões sejam encaradas com um grau ponderado de ceticismo (ver MICHAEL, 2009 para uma crítica geral; CARVALHO, 2018b, p. 8–9 para um problema específico no tratamento da fonologia histórica do Resígaro).
- [5] Zamponi (2003) apresenta o que, segundo o autor, seriam evidências de inovações lexicais, e algumas poucas inovações morfológicas, que permitiriam agrupar o Maipure, o Yavitero, o Baniva de Guainia (com seu co-dialeto Warekena do Xié) e o Baré em um mesmo ramo. A maior parte das supostas inovações não são exclusivas desse subgrupo (por exemplo, enfraquecimento de *\*k* precedendo *\*i*; uma forma idêntica ou próxima a *-(k)anu* para ‘braço’) e não constituem argumentos fortes o suficiente para reconhecer esse subgrupo com base em inovações.
- [6] O subgrupo Central, até o momento baseado essencialmente no fator de proximidade geográfica, reuniria o Paresi e o Enawenê-Nawê, que seriam co-dialetos de uma mesma língua (“Aruak do Juruena”; CARVALHO e BRANDÃO, 2018; BRANDÃO *et al.*, 2018) com o ramo do Xingu [7].
- [7] O ramo Xingu inclui, por um lado, o Waurá, o Mehináku e o extinto Kustenaú como membros do que parece ser um mesmo complexo dialetal e, por outro lado, o mais diferenciado Yawalapiti (CARVALHO, 2016a). A fusão de *\*ɨ* e *\*i* em *i* (com posterior *\*e > ɨ*) e a coronalização das labiais *\*p > r* e *\*m > ɲ* no contexto de *\*i* (mas não de *i < \*ɨ*) caracterizam o Yawalapiti unicamente com relação ao Waurá-Mehináku-Kustenaú (CARVALHO, 2016b).
- [8] Tradicionalmente associado a um ramo mais inclusivo, o hipotético “ramo Pré-Andino”, que incluiria também as línguas Campa [9], o subgrupo Purus já foi objeto de algum trabalho histórico-comparativo (FACUNDES, 2000, 2002; BRANDÃO, FACUNDES, 2007). Embora pareça um grupo coeso lexicalmente, não foi ainda estabelecido com base em inovações compartilhadas.
- [9] As línguas Campa (Kampa) formam um subgrupo bastante coeso, compartilhando formas lexicais (*\*iNɣato* “árvore”, *\*-taki* “casca”, *\*-toNki* “osso”, etc.) e mesmo gramaticais que não possuem cognatos óbvios em outros ramos da família. A existência de uma oposição de animacidade (além das categorias masculino e feminino), além de um contraste de clusividade no sistema pronominal (“primeira pessoa + terceira pessoa” vs. “primeira pessoa + terceira pessoa + segunda pessoa”) são características

gramaticais distintivas. Esta última característica é em geral apontada como resultado de influência areal de línguas dos Andes, como as variedades Quéchuá. Fonologicamente, a presença de um monotongo complexo (ditongo?) *\*ī* (MICHAEL, 2011), de alternâncias morfofonológicas *\*k ~ \*g*, *\*p ~ \*β* e de sequências mediais de oclusivas nasais seguidas de oclusivas orais homorgânicas (*\*NC*) indicam caminhos para investigações que possam apontar inovações fonológicas características deste subgrupo.

- [10] O subgrupo Bolívia-Paraná reúne línguas faladas entre os rios Mamoré e Guaporé, na Bolívia, e ainda o Terena, falado no Brasil próximo ao rio Paraguai (bacia do rio Paraná). Há, aparentemente, evidência para um desenvolvimento contextual *\*a > \*o* que seria compartilhado pelas línguas deste ramo (CARVALHO, no prelo), mas, na ausência de reconstruções da proto-língua do subgrupo e de uma melhor compreensão do Proto-Aruak, o fator condicionante deste desenvolvimento não foi ainda identificado. Um fator complicador é a existência, primeiro, de harmonia vocálica e *ablaut a ~ o*, este último associado a um sistema de expressão da distinção de modo *realis/irrealis*, e, segundo, de desenvolvimentos posteriores, como a fusão de *\*a* e *\*o* em *a* no dialeto Ignaciano do Mojeño (CARVALHO & ROSE, 2018), e que tornam as correspondências envolvendo *a* e *o* nesse grupo de línguas sensivelmente mais complexas.

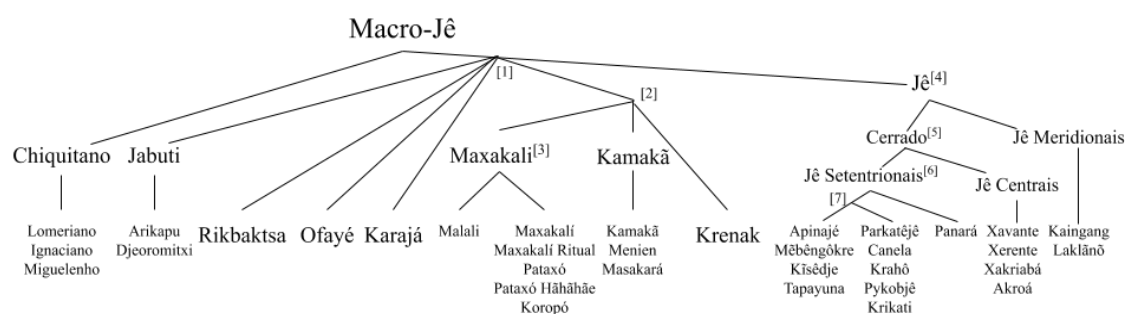
- [11] Mais certa, no entanto, é a divisão, interna ao Bolívia-Paraná entre, por um lado, o Baure (e seus co-dialetos, como o Carmelito e o Joaquiniano), e, por outro, o ramo provisoriamente chamado de “Achane” (CARVALHO, 2017e), que incluiria o Mojeño, o Paunaka e o Terena. Embora o Terena tenha tradicionalmente sido classificado como mais distante, a exclusão desta língua de um subgrupo formado por Baure, Mojeño e Paunaka justifica-se apenas com base em critérios geográficos. Além de inúmeros itens lexicais compartilhados (CARVALHO, 2018a, 2018d, 2017e, no prelo), outras similaridades incluem: a forma e a distribuição de alguns classificadores, uma forma *-VCV* para o sufixo de objeto de primeira pessoa plural (CARVALHO, 2017e, p. 82–83), aparentemente única entre as línguas da família, similaridades em domínios lexicais estruturados (CARVALHO, 2018d), e uma alomorfa no sufixo reflexivo, *-wo ~ -pu*, em que o segundo alomorfe tem sua distribuição restrita ao contexto do morfema de *irrealis -a*.

- [12] A proximidade maior entre o Mojeño (que inclui os dialetos Trinitário, Ignaciano, Loretano e Javeriano, além do extinto Mojeño Antigo registrado por missionários Jesuítas no final do século XVII) e o Paunaka é sugerida não apenas por inúmeros compartilhamentos lexicais mas também por similaridades gramaticais sem correlatos óbvios em outras línguas da família, como um “nome relacional” usado para a posse indireta de nomes alienáveis, e um prefixo de “terceira pessoa não-específica”, de forma *ti-*, e que não possui cognatos nos sistemas pronominais de qualquer outra língua Aruak (ver CARVALHO, 2018a).

Pelo que foi resumido acima, sustentamos que a observação de Campbell (2012: 71) segundo a qual a classificação interna da família Aruak seria bem menos clara do que a de outras grandes famílias da América do Sul permanece válida e atual.

O tronco **Macro-Jê** compreende as seguintes famílias bem definidas (muitas delas consistem de apenas uma língua): Chiquitano, Jabuti, Rikbaktsa, Ofayé, Karajá, Maxakali, Kamakã, Krenak, Jê. O Quadro 3 detalha o estado da arte dos estudos em subgrupamento deste tronco.

Quadro 3. Subgrupamento do tronco Macro-Jê



A seguir, comentamos a estrutura apresentada acima.

- [1] Divisão binária do Proto-Macro-Jê em Chiquitano e Macro-Jê Nuclear. Uma inovação marcante que caracteriza o ramo que chamamos aqui de “Macro-Jê Nuclear” diz respeito à chamada *consoante temática* (também conhecida como *prefixo relacional*), que ocorre em todos os temas cujos cognatos em Chiquitano são iniciados por uma vogal: Chiquitano *-otu* ‘língua’ ~ Proto-Macro-Jê Nuclear *\*ñ-ōctək*; Chiquitano *-ē?ē* ‘mão’ ~ Proto-Macro-Jê Nuclear *\*ñ-ĩm*. A consoante temática nas línguas Macro-Jê Nucleares pode desaparecer nas formas flexionadas para a segunda e para a terceira pessoa; em outros ambientes, sua ocorrência é obrigatória.
- [2] A proximidade estreita entre a família Maxakali e a língua Krenak é sugerida por Nikulin e Silva (2019). As inovações que caracterizam o ramo em questão incluem a evolução das consoantes do Proto-Macro-Jê (*\*p*, *\*t* > *\*m*, *\*n* em ambientes nasais; *\*mb*, *\*nd* > *\*p*, *\*t*), uma mudança semântica no uso dos verbos *\*tē* ‘ir/vir (singular)’, *\*mũ* ‘ir/vir (plural)’ > *\*mũ* ‘ir’, *\*nē* ‘vir’, bem como uma extensão da esfera de uso das orações nominalizadas, que foram reanalisadas como orações independentes em Proto-Maxakali-Krenak (as antigas orações independentes sobreviveram nessas línguas apenas em algumas construções, como, por exemplo, no imperativo). Para a família Kamakã, dispomos de dados muito limitados, que, no entanto, são compatíveis com a classificação do Kamakã como um membro do ramo Maxakali-Krenak.
- [3] A separação antiga do Malali das demais línguas Maxakali é sugerida por algumas retenções fonológicas, tais como a ausência de

epêntese vocálica nos encontros consonantais do tipo *\*Cr* (> Malali *C*, Proto-Maxakali Nuclear *\*Cid* > Maxakali *Cit*).

- [4] A divisão família Jê em dois ramos, Cerrado e Jê Meridional, é evidenciada por claras inovações em ambas proto-línguas (tais como o desenvolvimento de Proto-Jê *\*ɔ* > Proto-Cerrado *\*wa*, Proto-Jê Meridional *\*a*, entre muitas outras inovações fonológicas e lexicais, ver NIKULIN, 2019); a língua Ingain, já extinta (não representada no Quadro 3), é a mais divergente entre as línguas Jê Meridionais, apresentando várias retenções do Proto-Jê, tais como os reflexos de Proto-Jê *\*kēt* ‘pedra’ e *\*mbīt* ‘astro’.
- [5] A divisão binária do Proto-Cerrado em dois ramos (Jê Setentrional e Jê Central) se justifica pela existência de inovações nas duas proto-línguas, tais como a nasalização *\*am*, *\*um*, *\*im* > *\*ām*, *\*ūm*, *\*ĩm* em Proto-Jê Setentrional e a reorganização do sistema vocálico e consonantal em Proto-Jê Central (NIKULIN, 2017).
- [6] A separação antecipada do Panará do Proto-Jê Setentrional é evidenciada por algumas inovações lexicais, tais como a substituição de *\*kakūm* ‘estação seca’ (refletido como Panará *akūŋ*, cf. Proto-Jê Central *\*wahuṃ*) por *\*aŋra*.
- [7] Após a separação do Panará, o Proto-Jê Setentrional (*stricto sensu*) sofre uma divisão binária em dois ramos, Timbira e Trans-Tocantins (incluindo o Apinajé, o Mêbêngôkre, o Kîsêdjê e o Tapayuna). O Proto-Timbira e o Proto-Trans-Tocantins são caracterizados por múltiplas inovações compartilhadas, tais como *\*c* > *h* em Proto-Timbira ou a substituição de *\*a-mbɔ* ‘comer (intransitivo)’ por *\*ap-ku* em Proto-Trans-Tocantins. A estrutura de cada um desses ramos não é indicada no Quadro 3, mas há evidências que permitem identificar alguns pequenos subramos dentro deles, tais como Pykobjê-Krikati, Canela-Krahô ou Kîsêdjê-Tapayuna (ver NIKULIN, SALANOVA a sair).

Para concluir, ressaltamos que entre as limitações dos trabalhos que tratam da classificação das línguas indígenas brasileiras, uma das mais sérias é a escassez de reconstruções adequadas de nível intermediário, isto é, para os ancestrais comuns dos subgrupos ou ramos destas famílias. Essa limitação contribui para uma baixa resolução nas classificações internas, além da produção de padrões muitas vezes erráticos de correspondência e de etimologias com muitos elementos não explicados, como é o caso do trabalho de Payne (1991) sobre as línguas Aruak, já discutido em alguns pontos acima. É de se esperar que futuras pesquisas privilegiem uma abordagem *bottom-up*, dando uma devida atenção à reconstrução das proto-línguas intermediárias, à topologia das árvores filogenéticas e à distribuição das inovações dentro de cada agrupamento genético.



Uma vez que a constituição de um determinado agrupamento linguístico é estabelecida (ainda que parcialmente), é possível proceder à reconstrução de sua língua ancestral. Por motivos de espaço, nesta seção abordaremos apenas a reconstrução da fonologia e da sintaxe das proto-línguas, não discutindo, portanto, os estudos em reconstrução semântica ou morfológica.

## RECONSTRUÇÃO FONOLÓGICA

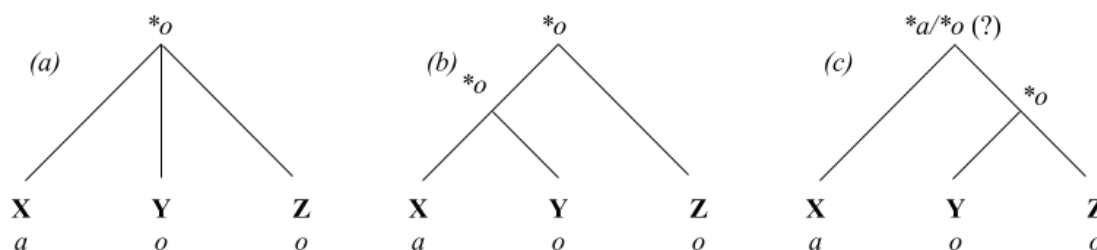
O método central para a reconstrução da fonologia de protolínguas é o *método comparativo*. Uma vez que um grupo de línguas tenha sido reconhecido como geneticamente válido e que um conjunto de prováveis elementos cognatos tenha sido identificado (ver subseção ‘Classificação interna’), sua aplicação consiste, *grosso modo*, em duas etapas: (1) a identificação das correspondências sonoras entre as línguas que estão sendo comparadas e (2) a interpretação diacrônica dessas correspondências, ou seja, a atribuição de segmentos (sequências de segmentos, elementos suprasegmentais) específicos reconstruídos para cada uma das correspondências em casos de contraste entre as correspondências, e a identificação dos desenvolvimentos condicionados (os chamados *splits*) quando as correspondências estiverem em distribuição complementar.<sup>13</sup> Dessa forma são estabelecidas as mudanças sonoras no percurso da protolíngua reconstruída às línguas modernas. A tarefa do linguista histórico é a de propor um cenário plausível para o desenvolvimento dos sons da protolíngua em todas as línguas descendentes, minimizando, na medida do possível, a necessidade de recorrer a mudanças sonoras tipologicamente incomuns.

É de boa praxe começar a reconstrução pelas protolínguas de baixa profundidade temporal, procedendo, sucessivamente, à reconstrução das línguas ancestrais dos agrupamentos cada vez mais abrangentes (dito isso, é necessário frisar que o uso dos dados externos na reconstrução intermediária não é excluído de forma categórica, sendo que esses dados podem ser determinantes em casos onde há múltiplas opções viáveis de reconstrução). Embora na atualidade haja um consenso quanto à importância desse procedimento (conhecido também como *reconstrução bottom-up*), ele não tem sido amplamente adotado nos trabalhos referentes à reconstrução das protolínguas de grandes famílias ou troncos sul-americanos, com algumas notáveis exceções.

Outro ponto importante é a relevância do subagrupamento da família cuja proto-língua está sendo reconstruída para as decisões reconstitutivas. Consideremos o seguinte exemplo imaginário. Suponhamos que uma determinada família compreende três línguas (denominadas X, Y, Z), havendo entre elas uma correspondência diacrônica  $Xa:Yo:Zo$ . Considerando que a mudança sonora  $a > o$  é tão plausível quanto  $o > a$ , qual segmento deveríamos reconstruir para a protolíngua? A resposta dependerá da configuração topológica da árvore

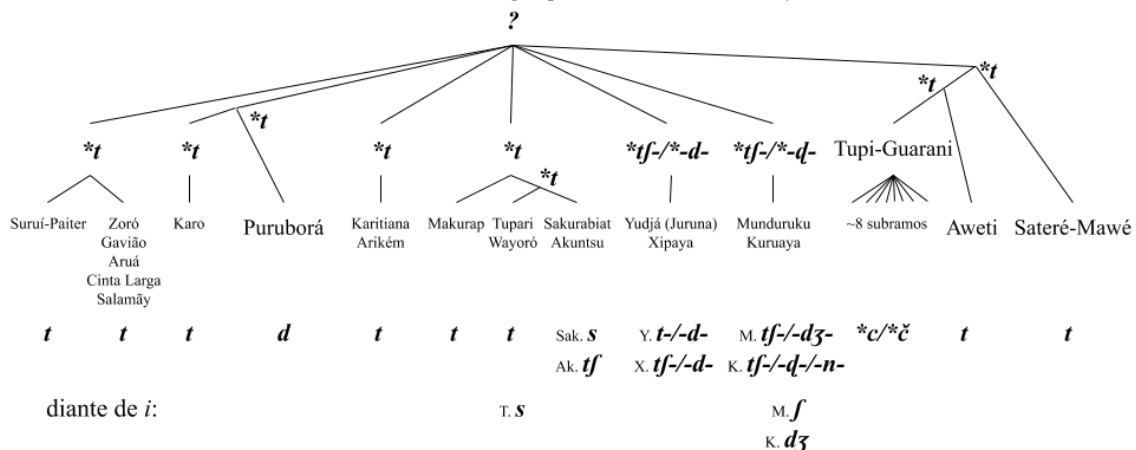
filogenética da família. Os cenários mais econômicos para cada uma das três possíveis configurações estão apresentados no Quadro 4 abaixo.

**Quadro 4.** Subagrupamento e reconstrução



Se o subagrupamento da família não envolve nenhum nó intermediário (ou seja, se as línguas X, Y e Z são mutuamente equidistantes; opção (a) no Quadro 4), a solução mais plausível seria reconstruir *\*o* para a protolíngua, visto que nesse caso teríamos que postular apenas uma inovação (*\*o* > *a* na língua X) e não duas inovações paralelas, que seriam necessárias caso reconstruíssemos *\*a* para a protolíngua (observe que este último cenário também não pode ser descartado e deve ser preferido, por exemplo, se existe uma correspondência concorrente *X o* : *Y o* : *Z o*). Se as línguas X e Y são mais estreitamente relacionadas uma à outra, formando um nó na árvore filogenética da família (opção (b) no Quadro 4), a solução mais econômica é reconstruir *\*o* tanto para a protolíngua da família inteira, como para a protolíngua intermediária de X e Y: nesse caso é necessário postular apenas uma inovação, sendo que qualquer cenário alternativo exigiria, no mínimo, duas inovações. Já no caso de uma proximidade especial entre as línguas Y e Z (opção (c) no Quadro 4) é possível propor reconstrução sólida apenas para a protolíngua intermediária de Y e Z (*\*o*, sem nenhuma mudança nas línguas-filhas). Para decidir se a protolíngua apresentava *\*a* ou *\*o* neste caso, teremos que recorrer a informações adicionais (sistêmicas ou externas), já que qualquer uma dessas opções exigiria apenas uma inovação na família (seja em X, seja em proto-Y-Z).

Em certos trabalhos influentes sobre a reconstrução das protolínguas das famílias linguísticas brasileiras, o subagrupamento dessas famílias tem sido ignorado, levando os autores a postularem reconstruções pouco prováveis de alguns segmentos. Uma das reconstruções que sofre, de forma sistemática, desse vício é a reconstrução das consoantes do Proto-Tupi de Rodrigues (2007). Examinemos a correspondência atribuída por esse autor a Proto-Tupi *\*c*(<sup>?</sup>)/*\*ç*(<sup>?</sup>):<sup>14</sup> Suruí-Paiter *n*-/*t*- : Mondé *t* : Karo *t* : Puruborá *d* (*n* em ambientes nasais) : Karitiana *t* : Makurap *t* : Tupari *t* (*s* diante de *i*) : Wayoró *t* : Sakurabiat *s* : Akuntsu *tʃ* : Yudjá *t*-/*d*- : Xipaya *tʃ*-/*d*- : Munduruku *tʃ*-/*dʒ*- (*f* diante de *i*) : Kuruaya *tʃ*-/*d*- (*tʃ*-/*n*- em ambientes nasais, *dʒ* diante de *i*) : Proto-Tupi-Guarani *\*tʃ*/*\*tʃ* : Aweti *t* : Sateré-Mawé *t*. Para maior clareza, reproduzimos essa mesma correspondência na árvore genética do tronco Tupi no Quadro 5 abaixo.

**Quadro 5. Subagrupamento e reconstrução**

Damos alguns exemplos na Tabela 3.

Tabela 3. Alguns conjuntos de cognatos nas línguas Tupi

glosa	sair, chegar	ir	morder	tomar banho	veado	pesado	lua	mãe	muri- çoca
<b>Suruí-Paiter</b>					(?)itʃi-ab	patiga	ŋati-kat	ni/-ti	nig
<b>Mondé</b>					i(:)ti:		gati	ti	
<b>Karo</b>			to		iti	pi?ti			tik
<b>Puruborá</b>	anēm-a				iði				
<b>Karitiana</b>	tām'voar', otām			oti		pĩti	oti	ti	tik
<b>Makurap</b>				tʃ-ato-a	iti:	poti		ti	
<b>Tupari</b>			toko-a 'mascar'	ato-a		posi		si	sik
<b>Wayoró</b>			togo 'macaxeira mascada'	ato-a	iti:	poti		ti	tik

Sakurabiat			sogo-a	aso-a	isi:	posi		si	
Akuntsu			tʃog-a	atʃo-a	(?) iti:			ti	tik
Yudjá		tʃa	e-tʃuku			padetu		dʒa	
Xipaya		ta	e-tuku			padetu		diã	
Munduruku	tʃêm, aɖʒêm	tʃə		aɖʒok		poʃi		ʔʃi	ʃik
Kuruaya	tʃêm	tʃi		aɖok	idʒi		wadʒi	dʒi(?)	dʒik
Proto-Tupi-Guarani	*tʃêm, *w-atʃêm	*tso	*tʃuʔu	*atʃuk		*potsij	*jatʃi	*tʃi	
Aweti	têm	to	tuʔu	atuk	t̃i-wapat	potij	tati	t̃i	
Sateré-Mawé	têm	to	tuʔu		iti:	potij	wati	-t̃i	

Obs: Omitimos aqui a marcação de tons nas línguas tonais.

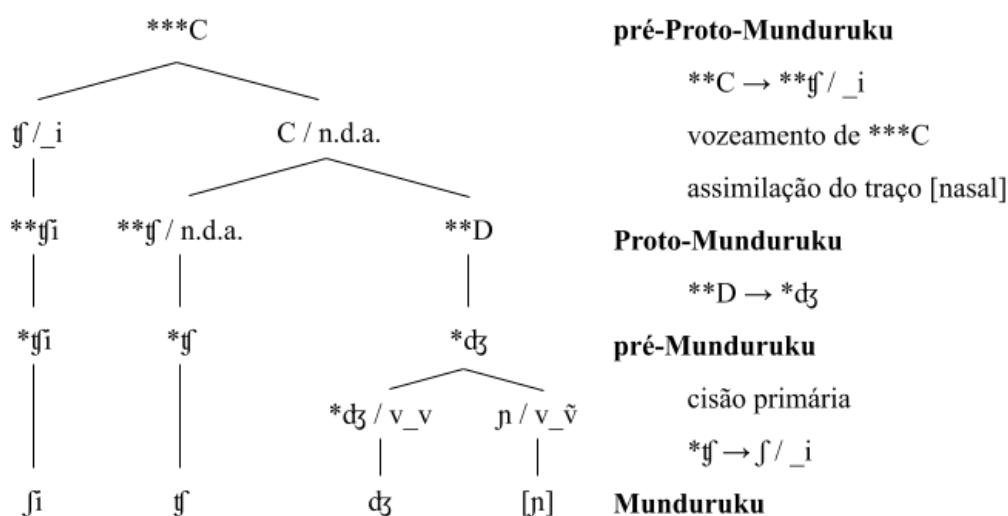
Podemos observar que quase todos os descendentes imediatos do Proto-Tupi (com a exceção do Proto-Munduruku e do Proto-Juruna) apresentam o reflexo (*\**)*t*, sugerindo fortemente que a consoante em questão deve ser reconstruída como Proto-Tupi *\*t*. Além das famílias Munduruku e Juruna (que, a princípio, poderiam formar um ramo), as línguas que possuem outros reflexos que não (*\**)*t* (excetuando-se os ambientes específicos, tais como *\_i*) são: Puruborá, Sakurabiat, Akuntsu e Proto-Tupi-Guarani. Entretanto, todas essas línguas pertencem a agrupamentos mais abrangentes (Ramarama-Puruborá, Tupari e Mawé-Guarani, respectivamente), em que todas as demais línguas apresentam, de fato, o reflexo *t*, permitindo reconstruir *\*t* para o Proto-Ramarama-Puruborá, Proto-Tupari (ANDRADE, a sair) e Proto-Mawé-Guarani (MEIRA, DRUDE, 2015).

Dessa forma, a reconstrução de *\*c(ʔ)/\*ç(ʔ)* de Rodrigues (2007) não é corroborada pelos dados das línguas Tupi se levarmos em consideração a estrutura interna desse tronco. Críticas semelhantes podem ser estendidas a outras consoantes coronais (e algumas não coronais) reconstruídas por esse autor.<sup>15</sup>

Os resultados da aplicação do método comparativo podem ser verificados e/ou complementados através de *reconstrução interna*.<sup>16</sup> Este método consiste na

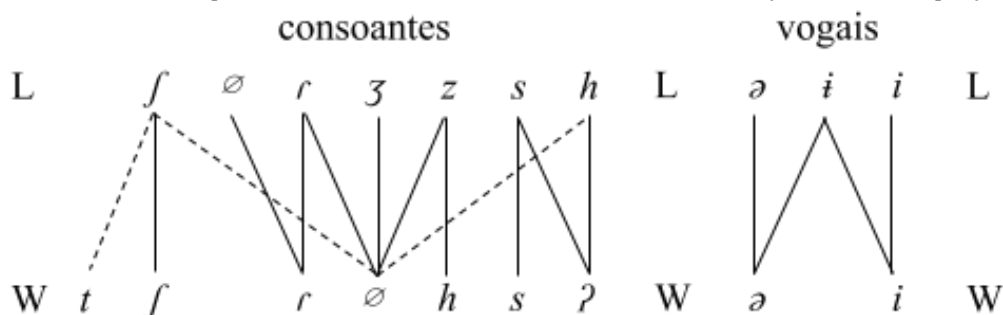
análise minuciosa dos fenômenos internos a uma língua (tais como alternâncias morfofonológicas, padrões alofônicos e lacunas distribucionais) que sugerem uma origem diacrônica dos mesmos. Uma limitação importante da reconstrução interna é que esse método (a não ser quando combinado com o método comparativo) não dá nenhuma informação sobre a cronologia das mudanças que ele revela, exceto por sua cronologia relativa. Por exemplo, aplicando o método de reconstrução interna a uma determinada alternância observada em Proto-Tupi-Guarani, é impossível dizer (sem considerar dados adicionais) se a mudança sonora que deu origem a essa alternância ocorreu na história independente de Proto-Tupi-Guarani, Proto-Aweti-Guarani, Proto-Mawé-Guarani ou até mesmo na história do Proto-Tupi. Para uma aplicação metodologicamente correta da reconstrução interna, referimos o leitor ao trabalho exemplar de Picanço (2005), que analisa as alternâncias, os padrões de alofonia e as lacunas distribucionais em Munduruku e Kuruaya (ambas da família Munduruku), chegando a conclusões sólidas referentes a mudanças sonoras ocorridas nessas duas línguas. Para tanto, a autora distingue entre diversos níveis cronológicos (Munduruku, pré-Munduruku, Proto-Munduruku, pré-Proto-Munduruku). Entre outras coisas, Picanço (2005) reconstrói a evolução do segmento que continua Proto-Tupi *\*t* (ver discussão acima) na língua Munduruku (Quadro 6).

**Quadro 6.** Evolução de Proto-Tupi *\*t* na família Munduruku (PICANÇO, 2005, p. 189)



A fim de ilustrar como funciona a aplicação do método comparativo aos dados de línguas reais, trazemos um exemplo, extraído de um trabalho de Meira (2005), que diz respeito à reconstrução do Proto-Bakairi, a proto-língua de dois dialetos da língua Bakairi (família Caribe, Mato Grosso), o Oriental (doravante L, Área Indígena Bakairi) e o Ocidental (doravante W, Área Indígena Santana).

No Quadro 7, representamos as correspondências não biunívocas entre os dois dialetos do Bakairi (isto é, aquelas em que um segmento em um dialeto pode corresponder a múltiplos segmentos do outro) identificados por Meira (2005).

**Quadro 7.** Correspondências não triviais entre os dialetos Bakairi (MEIRA, 2005, p. 5)

Meira (2005) analisa cada uma dessas correspondências a fim de estabelecer se elas devem ser atribuídas a segmentos (ou sequências de segmentos) da proto-língua que deixaram de existir nos dialetos atuais (exemplificando os chamados *mergers*) ou se elas resultam de desenvolvimentos condicionados dos segmentos da proto-língua (os chamados *splits*). Em casos ambíguos, o autor recorre a dados externos de outras línguas da família Caribe (Hixkaryana, Makuxi, Ikpeng), bem como aos registros de Steinen (1892; ver seção ‘Rastreamento de mudanças linguísticas: método filológico’, para uma discussão sobre a aplicação do método filológico às línguas indígenas). Os resultados de Meira (2005) são resumidos na Tabela 4.



Tabela 4. Reconstrução de alguns segmentos do Proto-Bakairi (PB) segundo Meira (2005)<sup>17</sup>

PB	L	W	exemplos	
*s	s	s	L <i>pasega</i> /W <i>pasega</i> 'porco', L <i>sagunu</i> /W <i>sagunu</i> 'areia', L <i>se</i> /W <i>se</i> 'pau'	
*ʔs	s	ʔ	L <i>ise</i> /W <i>iʔe</i> 'partícula de futuro', L <i>figaseda</i> /W <i>fiʔaʔeda</i> 'eu o chamei', L <i>imasedo</i> /W <i>iməʔedo</i> 'grande'	
*ʃ	ʃ	ʃ	L <i>feu</i> /W <i>feu</i> 'quati', L <i>āgaferi</i> /W <i>agaferi</i> 'teu cérebro', L <i>fogo</i> /W <i>fogo</i> 'pai!'	
*ʔʃ	ʃ	ʔ	L <i>odafi</i> /W <i>odaʔi</i> 'adentro', L <i>sefi</i> /W <i>seʔi</i> 'deixa eu ver'	
*ʔh	h	ʔ	L <i>tuhu</i> /W <i>tuʔu</i> 'pedra', L <i>taho</i> /W <i>taʔo</i> 'faca', L <i>poroho</i> /W <i>poroʔo</i> 'raposa'	
*z	z	∅	padrão geral	L <i>kozekə</i> /W <i>koekə</i> 'esp. de veado', L <i>tutuze</i> /W <i>tutue</i> 'saber', L <i>ize</i> /W <i>ie</i> 'vontade'
		h	antes de a	L <i>azagə</i> /W <i>ahagə</i> 'dois', L <i>meza</i> /W <i>meha</i> 'esp. de peixe', L <i>mazagə</i> /W <i>mahagə</i> 'esp. de muriçoca'
*ʒ	ʒ	∅	L <i>piazi</i> /W <i>pəai</i> 'pajé', L <i>orozi</i> /W <i>oro</i> 'caju', L <i>mozi</i> /W <i>moi</i> 'aranha'	
*r	r	∅	entre V nasais	L <i>pārē</i> /W <i>pāē</i> 'piranha', L <i>pōrā</i> /W <i>pōā</i> 'colar, miçanga', L <i>kirīwī</i> /W <i>kīwī</i> 'beija-flor', L <i>pērā</i> /W <i>pēā</i> 'mel'
		∅	padrão geral	L <i>paru</i> /W <i>paru</i> 'água, rio', L <i>igiri</i> /W <i>igiri</i> 'espinho', L <i>orozi</i> /W <i>oro</i> 'caju', L <i>era</i> /W <i>era</i> 'martim-pescador'
	∅	em sílabas pretônicas	L <i>fimukə</i> /W <i>firimukə</i> 'estrela', L <i>pienare</i> /W <i>pərenare</i> 'paca', L <i>kamikə</i> /W <i>karaməka</i> 'carrapato'	
*i	i	i	L <i>aki</i> /W <i>aki</i> 'cutia', L <i>ikila</i> /W <i>ikila</i> 'barro', L <i>pili</i> /W <i>pili</i> 'esp. de tucano', L <i>iwi</i> /W <i>iwa</i> 'montanha'	
*i	i	∅	padrão geral	L <i>ədi</i> /W <i>ədi</i> 'o quê?', L <i>iki</i> /W <i>iki</i> 'dormir', L <i>egi</i> /W <i>egi</i> 'xerimbabo', L <i>igiri</i> /W <i>igiri</i> 'espinho'
		ə	após C labiais	L <i>widi</i> /W <i>wədi</i> 'esposa', L <i>mitu</i> /W <i>mətu</i> 'mutum', L <i>pirəu</i> /W <i>pəreu</i> 'flecha', L <i>tāwī</i> /W <i>tāwī</i> 'fumo'
*ə	ə	∅	L <i>wəgə</i> /W <i>wəgə</i> 'sobre', L <i>əpa</i> /W <i>əpa</i> 'macaxeira', L <i>əgi</i> /W <i>əgi</i> 'quem?'	

Meira (2005) ainda aplica o método de reconstrução interna, avaliando a possibilidade de reconstruir apenas uma fricativa para o pré-Proto-Bakairi, que teria sofrido um *split* condicionado ( $**s > *s, *ʃ, *h, *z, *ʒ$ ). O autor conclui por apresentar uma amostra do léxico Proto-Bakairi, incluindo tanto as formas reconstruídas como seus reflexos no dialeto Oriental, no dialeto Ocidental e nos registros de Steinen (1892).

Terminamos esta seção com uma lista das referências mais importantes na reconstrução de proto-línguas das famílias linguísticas brasileiras (Tabela 5).

Tabela 5. Seleção de trabalhos sobre a reconstrução fonológica

família (tronco)	agrupamento	obras e comentários
Tupi	tronco inteiro	Rodrigues (2005, 2007) apresenta uma primeira proposta de reconstrução do Proto-Tupi, seguida também por Corrêa-da-Silva (2010). Essa proposta tem enfrentado críticas por não seguir os princípios <i>bottom-up</i> (MEIRA, DRUDE, 2015; ANDRADE, a sair); ver também a discussão acima sobre os problemas na interpretação diacrônica das correspondências. Não houve propostas reconstitutivas alternativas.
	Mondé	Não há propostas reconstitutivas, embora Anonby (2012) discuta alguns processos diacrônicos internos à família Mondé.
	Tupari	Moore e Galucio (1993) apresentam uma proposta inicial da reconstrução do Proto-Tupari, posteriormente desenvolvida por Galucio e Nogueira (2011) e Andrade (a sair).
	Juruna	Fargetti e Rodrigues (2008) apresentam uma proposta reconstitutiva para as consoantes do Proto-Juruna. Uma proposta alternativa, para as oclusivas, é apresentada em Carvalho (a sair). O mesmo autor está atualmente envolvido em uma reconstrução fonológica, morfológica e lexical do Proto-Juruna.
	Munduruku	Aspectos importantes da reconstrução do Proto-Munduruku são discutidos por Picanço (2005, 2010, 2013). Mendes Junior (2007) constata as correspondências entre o Munduruku e o Kuruaya, reunindo alguns pares de cognatos, mas não apresenta uma proposta reconstitutiva.
	Mawé-Guarani	A única proposta reconstitutiva que segue os princípios <i>bottom-up</i> é a de Meira e Drude (2015). Os autores observam que a reconstrução de Corrêa-da-Silva (2010) é idêntica, em termos fonológicos, à reconstrução do Proto-Tupi de Rodrigues (2005, 2007) e não pode ser representativa do Proto-Mawé-Guarani, embora o trabalho de Corrêa-da-Silva (2010) reúna uma grande quantidade de conjuntos de cognatos.

	Tupi-Guarani	As propostas reconstitutivas de Lemle (1971), Jensen (1989) e Mello (2000) convergem em muitos aspectos. Schleicher (1998) apresenta algumas críticas, em especial quanto à reconstrução de uma oposição entre duas africadas <i>*tʰ</i> e <i>*tʃ</i> , que ele argumenta ser desnecessária, e aplica o método de reconstrução interna.
Macro-Jê	Jê	A proposta reconstitutiva mais recente é a de Nikulin (2019). Ela difere substancialmente daquela de Davis (1966), sendo baseada em dados sincrônicos de maior qualidade e de maior quantidade de línguas. Isto foi possível graças a importantes avanços na documentação e na descrição das línguas Jê nas últimas duas décadas.
	Cerrado	O único trabalho dedicado à reconstrução do Proto-Cerrado é o de Nikulin (2017).
	Jê Setentrional	Nikulin e Salanova (a sair) apresentam uma proposta reconstitutiva mais elaborada do que aquela de Nikulin (2016).
	Jê Meridional	Aspectos da reconstrução fonológica do Proto-Jê Meridional são discutidos por Wiesemann (1978), D'Angelis (2008) e Jolkesky (2010). Jolkesky (2010) apresenta ainda um grande número de conjuntos de cognatos com as respectivas reconstruções; Nikulin (2016) defende importantes modificações na reconstrução das vogais do Proto-Jê Meridional.
	Maxakali-Krenak	O único trabalho dedicado à reconstrução do Proto-Maxakali-Krenak é o de Nikulin e Silva (2019).
	Jabuti	Voort (2007) apresenta uma reconstrução do Proto-Jabuti, embasado em dados significativamente mais numerosos do que Oliveira e Cabral (2015).
Aruak	família inteira	Há uma proposta de Payne (1991). Embora preliminar, é superior àquela de Matteson (1972), que inclui línguas não relacionadas (como se sabe hoje) na comparação. O trabalho de Payne (1991) deve, no entanto, ser tratado com cautela, por conter etimologias de abrangência duvidosa, julgamentos de cognação questionáveis e admitir muita irregularidade na postulação de correspondências.
	Lokono-Wayuunaiki	Há uma proposta reconstitutiva de Captain (2005).
	Campa	Heitzman (1973) apresenta uma reconstrução das consoantes do Proto-Campa, porém quase toda baseada em dados do contínuo dialetal Asháninka-Ashéninka. Michael <i>et al.</i> (2010) e Michael (2011) propõe uma reconstrução detalhada da fonologia segmental do Proto-Campa, mas, como Heitzman (1973), não apresentam uma reconstrução lexical.

	Purus	Aspectos da reconstrução fonológica e lexical do Proto-Purus são discutidos por Facundes (2000, 2002) e Brandão e Facundes (2007).
	Bolívia-Paraná	Jolkesky (2016) propõe uma reconstrução para a proto-língua deste subgrupo (que ele chama de “Mamoré-Guaporé”). Um problema de ordem geral com a proposta é a exclusão do Terena (ver a discussão do quadro 2 na subseção ‘ <b>Classificação interna</b> ’). Questões mais específicas sobre as propostas reconstitutivas de Jolkesky (2016) são abordadas por Carvalho (2018a) e Carvalho & Rose (2018). O primeiro analisa as correspondências entre o Paunaka, o Mojeño e o Terena (CARVALHO, 2018a), com um foco na fonologia histórica do Paunaka, mas não chega a propor uma reconstrução de proto-formas. Carvalho e Rose (2018) apresentam uma reconstrução detalhada da fonologia do Proto-Mojeño.
Caribe	família inteira	A proposta pioneira de Girard (1971a) é aperfeiçoada por Meira e Franchetto (2005) e Meira <i>et al.</i> (2010).
	Taranoano	Meira (1998) apresenta uma reconstrução exaustiva da fonologia e da morfologia flexional do Proto-Taranoano, língua ancestral do subgrupo formado pelas línguas Tiriyó, Akuriyó e Karihona.
	Bakairi	Meira (2005) apresenta uma reconstrução exaustiva da fonologia do Proto-Bakairi.
Tukano	família inteira	A proposta reconstitutiva mais recente e completa é a de Chacon (2015). Propostas anteriores, baseadas em dados menos numerosos e de qualidade mais baixa, incluem as de Waltz e Wheeler (1972) e de Malone (1986).
Pano-Takana	Pano	Após uma proposta pioneira de Shell (1965), que incluía dados de apenas alguns ramos da família, houve uma proposta atualizada de Oliveira (2014), que leva em consideração também os dados do Kaxarari e do ramo Mayoruna.
	Takana	Há propostas de Key (1968) e de Girard (1971b).
Txapakura	família inteira	Há propostas de Lima-Angenot (1997), Angenot e Lima-Angenot (2000) e Birchall <i>et al.</i> (2016).
Nadahup	família inteira	Martins (2005) apresenta uma reconstrução do Proto-Nadahup, porém Epps e Bolaños (2015, p. 475, nota 11) apontam a graves erros metodológicos nesse trabalho. As autoras ainda apresentam algumas formas do Proto-Nadahup reconstruídas por elas de forma tentativa, observando que um estudo comparativo mais detalhado só poderá ser elaborado com um avanço na documentação do Nadëb.
Nambikwara	família inteira	Há propostas reconstitutivas de Price (1978) e Costa e Wetzels (2008). Além disso, aspectos de reconstrução do Proto-Nambikwara foram abordados por Stella Telles (UFPE) em uma série de apresentações.

Bororo	família inteira	Camargos (2013) apresenta uma proposta reconstrutiva preliminar.
Guaicuru	família inteira	Há uma proposta de Viegas Barros (2013).
Arawá	família inteira	Dixon (2004) apresenta uma reconstrução da fonológica e lexical do Proto-Arawá. Dienst (2005) reavalia a reconstrução do inventário de fricativas, e trata de desenvolvimentos específicos do ramo Madihá (Deni-Kulina).

Além dos trabalhos supracitados, é de nosso conhecimento que atualmente estão sendo elaboradas reconstruções atualizadas das seguintes proto-línguas:

- Tupi: Proto-Tupari (Rafael Andrade, USP), Proto-Juruna (segundo autor), Proto-Tupi (ambos autores);
- Aruak: Proto-Purus e Proto-Achané (segundo autor);
- Macro-Jê: Proto-Timbira (Nandra Ribeiro, UFPA), Proto-Jabuti e Proto-Macro-Jê (primeiro autor);
- Proto-Bororo (primeiro autor).

## RECONSTRUÇÃO SINTÁTICA

Os estudos em reconstrução sintática ganharam destaque na linguística apenas recentemente. As premissas teóricas e epistemológicas de uma das abordagens mais promissoras nesta área, a *Gramática de Construções Diacrônica*, são discutidas em detalhe por Barðdal e Gildea (2015). Os autores observam que as unidades sintáticas, ao serem representadas dentro do paradigma da Gramática de Construções (CROFT, 2001), tornam-se passíveis de aplicação do método comparativo (e também do método de reconstrução interna), permitindo uma reconstrução. Quando se trata do método comparativo, o procedimento envolve a identificação de construções cognatas; no caso da reconstrução interna, o pesquisador deve identificar isomorfias entre diferentes construções que assinalam uma reanálise ocorrida num estágio anterior da língua. As mudanças sintáticas diacrônicas envolvem três mecanismos:

- reanálise, ou seja, uma mudança estrutural sem mudanças na superfície;
- extensão (incluindo o nivelamento paradigmático);
- empréstimo.

Um dos percursos mais bem estudados de evolução de construções sintáticas, o qual diz respeito à transformação de orações nominalizadas em orações independentes através de uma reanálise (GILDEA, 2008). Nas línguas

indígenas brasileiras, este tipo de desenvolvimento diacrônico foi reconstruído para uma série de línguas Caribe (GILDEA, 1998), para a língua isolada Trumai (GUIRARDELLO, 1999), para algumas línguas Jê Setentrionais (CASTRO ALVES, 2010) e para o Maxakali (NIKULIN, SILVA, a sair). Em todas essas línguas, as orações independentes para as quais é possível reconstruir este percurso de evolução tipicamente apresentam as seguintes características:

- um padrão ergativo-absolutivo de alinhamento morfossintático;
- a mesma morfossintaxe é utilizada para a codificação do argumento absolutivo — isto é, o único argumento (sujeito) dos verbos intransitivos ou o objeto dos verbos transitivos — e do possuidor de nomes (posse inalienável);
- o argumento ergativo (isto é, o sujeito dos verbos transitivos) é expresso por um sintagma posposicional ou por um sintagma nominal em um caso oblíquo;
- o verbo ocorre em uma forma especial que muitas vezes pode ser identificada com uma nominalização (no sentido etimológico), tipicamente apresentando pouca morfologia flexional.

No caso da língua Trumai, Guirardello (1999) recorre à reconstrução interna (uma vez que as relações externas dessa língua ainda não foram identificadas) e hipotetiza que todas as orações independentes do Trumai provêm de orações nominalizadas em construções clivadas, baseando-se em isomorfias morfossintáticas entre os sintagmas nominais e sintagmas verbais (em orações independentes ou subordinadas), bem como no fato de os verbos possuírem pouca morfologia flexional (como é esperado no caso das nominalizações). Em (1), reproduzimos o percurso diacrônico reconstruído por Guirardello (1999, capítulo 5) para as construções não clivadas do Trumai.

(1)	<b>estágio I</b>	oblíquo	Cop ( <i>ka/chĩ</i> )	[SN V-Nmlz]
	<b>estágio II</b>	oblíquo	Cop ( <i>ka/chĩ</i> ) Foc ( <i>in</i> )	[SN V-Nmlz]
			<i>alternativamente analisável como</i>	
		A/DAT	Cop ( <i>ka/chĩ</i> ) Foc ( <i>in</i> )	[O/S V]
	<b>estágio III</b>	A/DAT	Foc/T ( <i>ka in/chĩ in</i> )	[O/S V]

Cop = cópula, Foc = foco, Foc/T = foco e tempo, A = agente de verbo transitivo, O = objeto de verbo transitivo, S = único argumento de verbo intransitivo, V = verbo, DAT = dativo, SN = sintagma nominal, Nmlz = nominalização.



Por motivos de espaço, não podemos discutir em maior detalhe os trabalhos que abordam outros aspectos da reconstrução sintática ou morfossintática. Referimos o leitor às obras de Jensen (1990, 1998) para a reconstrução morfossintática do Proto-Tupi-Guarani, de Birchall (2015) para uma análise diacrônica da marcação de pessoa em verbos nas línguas Tupi, de Goergens (2011) e de Durand (2016) para uma análise diacrônica do alinhamento morfossintático nas línguas Aruak.

## RASTREAMENTO DE MUDANÇAS LINGUÍSTICAS: MÉTODO FILOLÓGICO

O *método filológico*, que consiste na análise minuciosa de registros históricos de estágios anteriores das línguas modernas (ou de línguas que não são faladas na atualidade) tem sido fundamental para a compreensão dos mecanismos que subjazem as mudanças linguísticas e para a verificação da validade dos demais métodos utilizados na linguística histórica.

Contrariamente ao que acontece na África e na Eurásia, a aplicabilidade do método filológico às línguas da América do Sul é limitada, visto que anteriormente à invasão europeia a transmissão de informação se dava por outros meios que não a escrita nesse continente. Por este motivo, os materiais linguísticos mais antigos referentes às línguas originárias da região remontam somente ao século XVI. Além disso, a quase totalidade dos registros dos séculos XVI–XIX foi produzida por falantes não nativos (e, em grande parte, por pessoas sem treinamento linguístico e nem proficiência na língua do registro), resultando na qualidade relativamente baixa dos dados dessa época, com raríssimas exceções.

Entretanto, mesmo com as limitações existentes alguns desses materiais possuem um grande valor para a linguística histórica,<sup>18</sup> pois eles viabilizam um rastreamento das mudanças linguísticas que ocorreram em determinadas línguas.

Um caso interessante diz respeito à língua falada pelos Kayapó do Sul, que ocupavam um vasto território no Brasil Central e tiveram um histórico conflituoso de contato com as frentes de colonização entre os séculos XVII e XX. Este povo era tido como extinto pela sociedade não-indígena até o trabalho de Heelas (1979), que levantou a hipótese de que os indígenas do povo Panará (Terra Indígena Panará) poderiam ser descendentes diretos dos Kayapó do Sul. Estudos posteriores (ver VASCONCELOS, 2013 e referências) confirmaram a hipótese; Vasconcelos (op. cit.) ainda constata que a língua Panará compartilha uma inovação fonológica importante (*\*r > j* em alguns ambientes) com as variedades do Kayapó do Sul outrora faladas em Santana do Paranaíba (MS) e no Triângulo Mineiro, divergindo neste sentido da variedade do Kayapó do Sul que era utilizada em Mossâmedes (GO) (ver também CARVALHO, 2016c).

Embora sejam colhidos em uma época relativamente recente, os dados das variedades Kayapó do Sul que provavelmente representam um estágio anterior do Panará possuem algumas características arcaicas que não estão presentes no Panará tal como é falado hoje. Embora a aplicação do método comparativo permita reconstruir a história fonológica dessa língua em algum detalhe sem que seja

necessário recorrer aos dados do Kayapó do Sul (NIKULIN, 2016, p. 170–171; 2017, p. 149–150), a análise desses dados não apenas corrobora as hipóteses reconstrutivas já existentes, mas também permite determinar a cronologia de certas mudanças linguísticas.

Por exemplo, graças ao método comparativo é sabido que em algum momento da história do Panará a vogal *a* foi nasalizada antes de consoantes pós-oralizadas (tais como *mp*, *nt*, *ns*, *ŋk*), processo que foi acompanhado de epêntese de *n*- no começo das palavras (NIKULIN, 2017, p. 149).<sup>19</sup> No entanto, seria impossível determinar a cronologia absoluta dessa mudança sonora se não fosse pelos registros do Kayapó do Sul. Comparem-se as seguintes formas reconstruídas para o Proto-Jê Setentrional (PJS) com seus respectivos reflexos em Kayapó do Sul (dialeto mineiro, BARBOSA, 1918)<sup>20</sup> e com seus reflexos Panará.

(2) Nasalização de *\*a* e epêntese de *n*-

a. PJS <i>*kaŋgro</i> ‘quente’	> Kayapó do Sul <i>aŋkjo</i>	> Panará <i>nãŋkjo</i>
b. PJS <i>*kaŋã</i> ‘cobra’	> Kayapó do Sul <i>aŋã</i>	> Panará <i>nãŋkã</i>
c. PJS <i>*kambrô</i> ‘sangue’	> Kayapó do Sul <i>ampjô</i>	> Panará <i>nãmpju</i>
d. PJS <i>*aŋgrô</i> ‘caitetu’	> Kayapó do Sul <i>aŋkjô</i>	> Panará <i>nãŋkjô</i>
e. PJS <i>*aŋgã</i> ‘miçanga’	> Kayapó do Sul <i>n/d</i>	> Panará <i>nãŋkã</i>

Os dados em (2) evidenciam que no percurso do Proto-Jê Setentrional para o Kayapó do Sul já haviam ocorrido algumas das mudanças sonoras próprias ao Panará, tais como a queda de *\*k*- antes de *-a-*, o ensurdecimento das nasais pós-oralizadas e, somente nos dialetos meridionais do Kayapó do Sul, a mudança *\*r > j* diante de vogais posteriores (CARVALHO, 2016c; NIKULIN, 2016, 2017). Por outro lado, há não menos de **três** mudanças sonoras, também características do Panará, que aparentemente ainda não haviam ocorrido no estágio dessa língua atestado por Barbosa. Trata-se das seguintes mudanças:

- a mudança sonora *ŋ > ŋk* (NIKULIN, 2017, p. 149), como em PJS *\*kaŋã* ‘cobra’, *\*ŋrô* ‘tucano’ > Panará *nãŋkã*, *ĩŋkjô-pepeti*; observe que o Kayapó do Sul conserva *ŋ* neste primeiro item;
- a nasalização *a > ã* precedendo consoantes pós-oralizadas (esta mudança é necessariamente posterior a *ŋ > ŋk*; em caso contrário, o desenvolvimento esperado seria PJS *\*kaŋã > \*aŋkã* e não *nãŋkã*);
- a epêntese de *n*- no início de palavra precedendo *ã*.

Os dados de Kayapó do Sul, portanto, oferecem evidências importantes que permitem atribuir as três mudanças sonoras acima ao período mais recente (século XX) da história da língua Panará.

Para as línguas da família Aruak, Carvalho (2017a, b) e Carvalho e Rose (2018) baseiam-se em transcrições deixadas por missionários, antropólogos e viajantes para complementar compreensão dos processos de mudança fonológica aos quais essas línguas se submeteram. No caso do Terena, por exemplo, língua de um povo Aruak do Mato Grosso do Sul, cujos ancestrais falavam uma língua documentada sob o nome de “Guaná”, Carvalho (2017a) mostra que evidências puramente internas à língua, em especial alternâncias morfofonológicas entre uma fricativa glotal *h* e fricativas coronais *s* ou *f*, permitem inferir a operação de um processo diacrônico de debucalização, *\*s, \*f > h*. Esse desenvolvimento teve como alvo os alofones desvozeados destas fricativas coronais, que possuem reflexos com constricção oral (não-debucalizados) na língua moderna apenas em contextos de vozeamento espontâneo induzido por nasalidade contextual (*háʔa* ‘pai dele/dela’, *nʒáʔa* ‘meu pai’, em que a nasalidade marca o possuidor de primeira pessoa do singular). No entanto, a presença de fricativas coronais *s* e *f* no Terena moderno (em elementos que não podem ser explicados como empréstimos, como o nome *fâne* ‘gente’ e o verbo *sîmo* ‘ele/ela chega’) poderia ser vista, em princípio, como um problema para esta análise, uma vez que a aplicação de uma mudança sonora regular de debucalização nos levaria a esperar a ausência completa de tais segmentos do estrato lexical herdado, isto é, Aruak, dessa língua. A análise de documentos produzidos, por exemplo, pelo etnólogo alemão Max Schmidt (1903) e pelo engenheiro e militar brasileiro Alfredo Taunay (1875) mostra que quando a mudança em questão operou na língua o Terena possuía, além das fricativas *s* e *f* que passaram a *h*, um par de segmentos africados, *ʦ* e *ʧ*. Esses últimos seriam, assim, a fonte das fricativas do Terena moderno, como mostramos na Tabela 6.

Tabela 6. Comparação dos dados de Schmidt (1903) com os do Terena moderno

glosa	Schmidt (1903)	Terena moderno	correspondências
avó	<ootsé>	ôse	<ts> : s
mandioca	<tšupú>	fúpu	<tš> : f
peixe	<šooé>	hõe	<š> : h
amarelo	<siaiti>	hijáʔiti	<s> : h

A comparação entre os dados do Terena moderno e os dados documentais tomados da língua como falada na passagem entre os séculos XIX e XX mostra, entre outras coisas, que o processo de debucalização que afetou as fricativas da língua foi, de fato, um processo de mudança sonora, regular, já que as fricativas ainda presentes no Terena moderno são em realidade reflexos de africadas (ver CARVALHO, 2017a, b para detalhes).

É muito importante ressaltar que qualquer aplicação do método filológico aos registros antigos deve ser precedida não apenas por uma análise grafêmica,<sup>21</sup> mas também por uma análise crítica do sistema de transcrição neles utilizado, com uma atenção especial às limitações dos autores desses registros (ver MEIRA, 2005, 11–19 para um tratamento exemplar dos dados desse tipo). Em particular, é

primordial entender que os dados que provêm de tais fontes não podem ser facilmente convertidos em representações fonológicas e devem ser decodificados com base em informações externas ao próprio registro para se chegar aos valores fonéticos dos grafemas utilizados. Se esse cuidado não é tomado, o pesquisador incorre em diversos riscos, tais como a identificação de correspondências fictícias.

## CONCLUSÃO

Neste trabalho fizemos uma tentativa de apresentar um panorama do estado da arte nos estudos em linguística histórica referentes às línguas indígenas brasileiras. Para tanto, buscamos dar um enfoque maior aos trabalhos mais recentes, que por motivos cronológicos não são contemplados em trabalhos anteriores tidos como referência (RODRIGUES, 1986; URBAN, 1998).

Na **primeira seção** (**'Classificação genética'**), além de apresentar os trabalhos que tratam da classificação das línguas indígenas brasileiras, procuramos demonstrar que essa tarefa depende crucialmente dos estudos em reconstrução. Por sua vez, na **subseção 'Reconstrução fonológica'** não apenas apresentamos os trabalhos que tratam da reconstrução de proto-línguas, mas também destacamos a importância dos estudos em subagrupamento para a aplicação do método comparativo e dos princípios de reconstrução *bottom-up*. Uma breve discussão acerca da reconstrução sintática encontra-se na **subseção 'Reconstrução sintática'**. Na **terceira seção** (**'Rastreamento de mudanças linguísticas: método filológico'**), discutimos as possíveis contribuições da aplicação do método filológico aos estudos em diacronia das línguas indígenas, apesar das limitações existentes. Por motivos de espaço, não pudemos abordar diversos tópicos interessantes, discutidos na literatura existente, tais como as hipóteses de parentesco distante, reconstrução morfológica ou a classificação interna da família Tupi-Guarani. No entanto, as respectivas referências são ubíquas nos trabalhos que citamos aqui, e o leitor interessado não terá dificuldade em encontrá-las.

## Notas

<sup>1</sup> Agradecemos a Mário André Coelho da Silva pelas sugestões que contribuíram para o aprimoramento deste texto. O primeiro autor agradece também à CAPES pelo apoio financeiro.

<sup>2</sup> Em razão da inexistência de critérios universais que poderiam distinguir entre línguas e dialetos, não é possível determinar este número com precisão. Crevels (2012, p. 179), citando um documento da FUNAI de 2005, apresenta o número de 180 línguas, que por sua vez é questionado como sendo excessivo por Moore *et al.* (2008), para quem 150 é um número "difícilmente ultrapassável" caso levemos em consideração o critério da inteligibilidade mútua para reconhecer dialetos/variedades de uma mesma língua. Galucio *et al.* (2018) estimam esse número em 150-160.

<sup>3</sup> Entre as diversas causas/fatores que explicam as similaridades encontradas entre duas ou mais línguas, a herança comum (parentesco) e o contato linguístico, esse último consistindo na transmissão de material linguístico entre duas ou mais línguas não necessariamente relacionadas por parentesco, são usualmente descritas como os *fatores históricos* por trás de similaridades entre línguas, e são, ambos, objetos de investigação na linguística histórica. Duas outras causas de similaridades são chamadas de *fatores não-históricos*, a saber: a mera convergência acidental e fortuita entre línguas, e as similaridades explicadas por regularidades tipológicas. Uma parcela significativa dos métodos da linguística histórica consiste de meios para efetuar a delicada tarefa de distinguir entre esses tipos de causas de similaridades interlinguísticas.

<sup>4</sup> Não mencionamos aqui, por limitações de espaço, as famílias presentes no Brasil historicamente mas cujas línguas deixaram de ser utilizadas na atualidade no território brasileiro em decorrência do processo colonizatório. Essa lista inclui a família Bora-Muinane, à qual pertence a língua ancestral do povo Miranha (ainda utilizada na Colômbia; os Miranha brasileiros passaram a falar português), e algumas pequenas famílias do leste brasileiro, tais como a família Kariri e a família Puri.

<sup>5</sup> Além da expressão *família linguística*, de uso generalizado e não-controverso para se referir a grupos de línguas que possuem um ancestral em comum, termos como *tronco linguístico* (tradução proposta para o inglês *stock*) e mesmo *filó* (inglês *phylum*) tornaram-se correntes na linguística brasileira, em especial para nomear os *troncos* Tupi e Macro-Jê. A diferença entre *família* e *tronco* seria de inclusividade filogenética, o *tronco* contendo mais subgrupos/línguas ou incorporando um grupo mais heterogêneo de unidades, com maior diversidade lexical e gramatical. O uso dessa nomenclatura foi introduzido no país por Aryon D. Rodrigues, baseado em convenções similares adotadas pelo linguista norte-americano Morris Swadesh. Em virtude dessa tradição já estabelecida, empregaremos aqui o termo *tronco* em adição ao termo *família*.

<sup>6</sup> É de conhecimento geral que em todas as famílias linguísticas bem estabelecidas há conjuntos de prováveis cognatos que violam as correspondências sonoras conhecidas, possuem semântica divergente e/ou não pertencem ao vocabulário básico. Tais cognatos, entretanto, não devem ser usados como evidência para uma demonstração inicial de parentesco e devem ser abordados somente após a apresentação da evidência central. Via de regra, uma vez que a reconstrução por meio do método histórico comparativo avança, esses conjuntos de prováveis cognatos são eventualmente rejeitados como apenas cognatos aparentes, ou são explicados em função de outros fatores, como mudanças sonoras adicionais, alterações analógicas, empréstimos dialetais e mudanças semânticas menos óbvias.

<sup>7</sup> Não incluímos na Tabela 1 os trabalhos que, apesar de possuírem o propósito de demonstrar o parentesco entre duas (ou mais) línguas ou famílias, não apresentam uma discussão das correspondências sonoras (RODRIGUES, 1958 para o tronco Tupi; RODRIGUES, 1966 para a língua Cinta-Larga e a família Mondé em geral) ou que contêm graves erros metodológicos, tais como segmentações morfológicas não justificadas, reconstruções fantasiosas (GUDSCHINSKY, 1971 para a relação entre a língua Ofayé e a família Jê; BOSWOOD, 1973 para a relação entre o Rikbaktsa e a família Jê; ver RIBEIRO, 2005; RIBEIRO, VOORT, 2010, p. 548, nota 32 para uma crítica) ou padrões erráticos de correspondências (MARTINS, 2011 para a relação entre a língua Guató e o tronco Macro-Jê).

<sup>8</sup> Pode-se dizer que o problema da classificação interna de línguas já reconhecidas como aparentadas é muitas vezes mais complexo do que o problema da classificação externa, isto é, o de discernir as relações entre línguas ainda não reconhecidas como aparentadas.

Se, como já discutido, no caso de hipóteses de parentesco precisamos distinguir entre quatro explicações possíveis para as similaridades notadas entre línguas (herança comum, contato, acidente ou convergência acidental, e similaridades tipológicas), no caso da classificação interna temos que distinguir ainda entre propriedades herdadas do *ancestral comum de toda a família* (propriedades que não tem relevância para estabelecer uma classificação interna) daquelas que são herdadas do *ancestral comum intermediário*, isto é, do ancestral comum do subgrupo ou ramo que pretendemos demonstrar. Essas últimas são a evidência crucial para uma classificação interna.

<sup>9</sup> Alguns autores argumentam que o modelo arbóreo não representa adequadamente a diversificação das línguas, principalmente em situações de contato intenso entre línguas geneticamente relacionadas, e propõem utilizar outros modelos, tais como o modelo de ondas ou de modelo de *linkages* (ver FRANÇOIS, 2014). Outros autores defendem a aplicabilidade do modelo arbóreo a esse tipo de situações (para uma discussão crítica, ver JACQUES, LIST, 2019). Além disso, o modelo arbóreo, por definição, não é aplicável a modalidades linguísticas de origem não genética, isto é, aos *pidgins* e às línguas crioulas (tais como o Karipúna, língua crioula de base lexical francesa falada no extremo norte do Amapá).

<sup>10</sup> Alguns autores têm apresentado soluções técnicas para os problemas identificados nas aplicações anteriores da glotocronologia. Entre outras coisas, tem-se atentado para a importância do discernimento entre as mudanças lexicais internas à língua (a chamada deriva) e aquelas ocasionadas pelo contato linguístico (ver STAROSTIN, 2013 para uma discussão).

<sup>11</sup> Preferimos reconstruir Proto-Mawé-Guarani e Proto-Aweti-Guarani \**ǝ* (em vez de \**r*, como fazem Meira e Drude (2015)) para a correspondência Sateré-Mawé *n* : Aweti *l* : Proto-Tupi-Guarani \**r*; os próprios autores apontam a algumas dificuldades em sua proposta, relacionadas à reconstrução do item com o significado ‘amargo’ (\**rǝp* segundo Meira e Drude, com uma nasalidade inexplicavelmente perdida em todas as línguas-filhas; \**ǝop* segundo a presente proposta).

<sup>12</sup> As línguas Palikur, Yaneshá’ (também chamada de Amuesha) e Chamicuro são sem sombras de dúvidas as de mais difícil classificação dentro da família. Uma combinação ainda pouco compreendida de mudanças sonoras erosivas somada ao impacto do contato com línguas não-Arawak torna bastante difícil a tarefa de identificar elementos cognatos entre essas línguas e os demais membros da família.

<sup>13</sup> Dois pontos merecem comentário. Em primeiro lugar, note que uma hipótese provável de parentesco funciona como *premissa* usual na aplicação do método, que, assim, não é aplicado a esmo às línguas, mas somente àquelas que já suspeitamos, com base em alguma evidência, terem uma origem comum. Esse fato é importante, uma vez que há uma opinião, ao nosso ver, errônea, porém popular, de que uma reconstrução fonológica extensa é *necessária* para que uma hipótese de parentesco entre línguas seja “levada a sério” (embora não haja dúvidas de que a existência de uma reconstrução adequada seja evidência *suficiente*). Em segundo lugar, o leitor perceptivo pode ter notado, mesmo com essa breve síntese, a similaridade entre o procedimento de reconstrução fonológica e os procedimentos de análise fonológica sincrônica. Correspondências em contraste indicam proto-fonemas distintos, enquanto correspondências em distribuição complementar são analisadas como desenvolvimentos contextuais de um único proto-fonema. As similaridades entre os dois procedimentos são enfatizadas nos clássicos trabalhos de Henry Hoenigswald, aos quais o leitor é referenciado (HOENIGSWALD, 1950, 1960).

<sup>14</sup> Rodrigues (2007) reconstrói a distinção entre \**c*(<sup>?</sup>) e \**ç*(<sup>?</sup>) a fim de explicar os reflexos distintos apenas em um subramo de uma família do tronco Tupi, o subramo Guarani. Para



uma crítica, ver Schleicher (1998) e Meira e Drude (2015). Quanto à reconstrução das consoantes glotalizadas, as evidências apresentadas por Rodrigues (2007) para substanciar a reconstrução de \*cʔ e \*çʔ provêm de conjuntos de cognatos que não instanciam a correspondência aqui examinada ('grosso', 'remédio', 'corda', 'olho', 'rio').

<sup>15</sup> Ao mesmo tempo, é necessário reconhecer que as correspondências identificadas por Rodrigues (2005, 2007) são, via de regra, corretamente identificadas; o problema principal de sua reconstrução jaz na interpretação diacrônica das mesmas.

<sup>16</sup> Outra maneira interessante de verificar as reconstruções obtidas através do método comparativo é a análise dos empréstimos antigos: como línguas diferentes sofrem mudanças sonoras diferentes, uma análise da evolução fonética de empréstimos pode revelar importantes fatos referentes à história fonológica tanto da língua-fonte, como da língua-recipiente. O contato entre fases anteriores de línguas indígenas brasileiras, entretanto, não tem sido suficientemente bem investigado, havendo apenas poucos trabalhos que tratam dessa questão (como, por exemplo, RODRIGUES, 1985, que discute as camadas cronológicas de contato linguístico entre as línguas das famílias Tupi-Guarani e Caribe; CARVALHO, 2017d sobre o contato entre línguas Tupi-Guarani e línguas Aruak meridionais e CARVALHO, 2018c sobre o contato entre o Terena, língua Aruak do Mato Grosso do Sul, e o Mbayá/Kadiwéu, língua da família Guaicuru).

<sup>17</sup> Não incluímos na Tabela 4 a reconstrução de \*n em sílabas pretônicas, a qual não é recuperável pelo método comparativo, uma vez que \*n foi perdido em ambos os dialetos nesse ambiente) e foi possível apenas graças à aplicação do método filológico aos dados de Steinen (1892) (ver MEIRA, 2005, p. 13, 15).

<sup>18</sup> Não discutimos aqui a importância desses materiais para as demais áreas da linguística, que é particularmente alta no caso de línguas que deixaram de ser faladas em algum momento após a data do registro. Limitamo-nos a mencionar o caso das línguas Kipeá e Dzubukuá (família Karirí), para as quais foi possível elaborar descrições fonológicas e morfossintáticas relativamente detalhadas (AZEVEDO, 1965; QUEIROZ, 2008, 2012) com base na gramática de Mamiani (1699) e nas catequeses de Mamiani (1698) e Nantes (1709), apesar de essas línguas não serem utilizadas na atualidade, e o caso da língua Patxohã, revitalizada pelo povo Pataxó a partir das listas de palavras registradas por viajantes no século XIX, entre outras fontes (BOMFIM, 2017).

<sup>19</sup> Lapiere (em preparação) considera que o processo de epêntese de *n*- é sincronicamente ativo em Panará.

<sup>20</sup> Os dados de Barbosa (1918) foram coletados por Alexandre de Souza Barbosa em 1911 na Aldeia da Água Vermelha, próximo de São Francisco de Sales (MG). As formas em (2) foram transcritas por ele como ⟨a(n)kió⟩, ⟨ampiô⟩, ⟨anhán⟩, ⟨ankiô⟩, respectivamente. Quanto à nossa interpretação da forma ⟨anhán⟩ como *añã*, baseamo-nos na seguinte nota de Barbosa: “*Nh* leia-se como em português, excepto em *anhán*, onde representa modificação vocal que em nossa língua não existe”.

<sup>21</sup> Embora a maioria dos pesquisadores estejam cientes de que os registros deste tipo frequentemente (mas não sempre) fazem uso de convenções de transcrição que replicam algumas particularidades da ortografia da língua materna do autor (normalmente alemão, português ou francês), infelizmente, há também trabalhos que chegam a conclusões errôneas por ignorarem essa possibilidade. Por exemplo, Braggio (2005), ao discutir as diferenças entre os dados do Xerente (língua Jê Central da família Jê, tronco Macro-Jê) registrados por Francis de Castelnau e publicados por Martius (1863, p. 139–141; Braggio erroneamente atribui a autoria desses dados ao próprio Martius), identifica equivocadamente as seguintes mudanças sonoras na história da língua Xerente:

- monotongação de *ai* → *e*, *ou* → *u* (⟨poucouanai⟩ → *ponkwanē* ‘dois’, ⟨croboudi⟩ → *krbu-di* ‘sedento’)
- apagamento de *g* em final de palavra (⟨kouacong⟩ → *wakō* ‘quati’);
- assimilação regressiva da nasalidade da consoante nasal alveolar em final de palavra pela vogal precedente e apagamento da consoante nasal alveolar (⟨kouacong⟩ → *wakō* ‘quati’, ⟨dicran⟩ → *da-krā* ‘cabeça’, ⟨noron⟩ → *nrō* ‘palmeira’), entre outras.

Está claro que nesses casos não se trata de mudanças diacrônicas, mas sim de um sistema de transcrição baseado na ortografia do francês. Este equívoco poderia ter sido evitado caso uma análise grafêmica tivesse sido aplicada a esse registro.

---

## Referências

---

ADAM, Lucien. Trois familles linguistiques des bassins de l’Amazone et de l’Orénoque. In: **Congrès International des Américanistes. Compte rendu de la septième session. Berlin 1888**. Berlin: W. H. Kuhl, 1890. p. 489–496.

\_\_\_\_\_. **Matériaux pour servir à l’établissement d’une grammaire comparée des dialectes de la famille Guaicurú**. Paris: J. Maisonneuve, 1899.

ADELAAR, Willem F. H. Propuesta de un nuevo vínculo genético entre dos grupos lingüísticos indígenas de la Amazonía occidental: Harakmbut y Katukina. In: MIRANDA, Luis (ed.). **I Congreso de lenguas indígenas de Sudamérica, v. 2**. Lima: Universidad Ricardo Palma, 2000. p. 219–236.

\_\_\_\_\_. Relações externas do Macro-Jê. O caso do Chiquitano. In: TELLES de A. P. LIMA, Stella Virgínia; Aldir S. de PAULA (orgs.). **Topicalizando Macro-Jê**. Recife: Nectar, 2008. p. 9–27.

AIKHENVALD, Alexandra. The Arawak language family. In: DIXON, R. M. W.; Alexandra AIKHENVALD (eds.). **The Amazonian Languages**. Cambridge: Cambridge University Press, 1999. p. 65–105.

ANDRADE, Rafael. As consoantes alvéolo-dentais do Proto-Tupará: revisão e reconstrução fonológica. In: OLIVEIRA, Christiane Cunha de (org.). **Memórias do II Encontro dos Americanistas no Cerrado**. Goiânia: Universidade Federal de Goiás, a sair.

ANGENOT, Jean-Pierre; Geralda de LIMA-ANGENOT. Sobre a reconstrução do Protochapakura. In: VOORT, Hein van der; Simon van de KERKE (eds.). **Ensaio sobre linguas indígenas de las tierras bajas de Sudamérica: Contribuciones al 49º Congreso Internacional de Americanistas en Quito, 1997**. Leiden: Research School of Asian, African and Amerindian Studies (CNWS), Universiteit Leiden, 2000. p. 53–70. (Lenguas Indígenas de América Latina (ILLA), v. 1.)

ANONBY, Stan. A historical comparative look at four Mondé languages. **SIL Electronic Working Papers**, v. 6, 2012.

AZEVEDO, Gilda Maria Corrêa de. **Língua Kiriri. Descrição do dialeto Kipeá**. Dissertação (Mestrado em Letras) – Universidade de Brasília, Brasília, 1965.

BARBOSA, Alexandre de Souza. **Cayapó e panará**. Uberaba (manuscrito), 1918.

BARÐDAL, Jóhanna; Spike GILDEA. Diachronic Construction Grammar: Epistemological context, basic assumptions and historical implications. In: BARÐDAL, Jóhanna; Elena

SMIRNOVA; Lotte SOMMERER; Spike GILDEA (eds.). **Diachronic Construction Grammar**. Amsterdã: John Benjamins, 2015. p. 1–50. (Constructional Approaches to Language, v. 18.)

BIRCHALL, Joshua. A comparison of verbal person marking across Tupian languages. **Boletim do Museu Paraense Emílio Goeldi. Ciências humanas**, Belém, v. 10, n. 2, p. 487–518, maio/ago. 2015.

BIRCHALL, Joshua; Michael DUNN; Simon GREENHILL. A combined comparative and phylogenetic analysis of the Chapacuran language family. **International Journal of American Linguistics**, Chicago, v. 82, n. 3, p. 255–284, jul. 2016.

BOMFIM, Anari Braz. **Patxohã, “língua de guerreiro”**: um estudo sobre o processo de retomada da língua Pataxó. Dissertação (Mestrado em Estudos Étnicos) – Universidade Federal da Bahia, Salvador, 2017.

BOSWOOD, Joan. Evidências para a inclusão do Aripaktsá no filo Macro-Jê. In: BRIDGEMAN, Loraine Irene (ed.). **Série Lingüística I**. Brasília: SIL, 1973. p. 67–78.

BRAGGIO, Silvia Lucia Bigonjal. Revisitando a fonética/fonologia da língua Xerente Akwê: uma visão comparativa dos dados de Martius (1866) a Maybury-Lewis (1965) com os de Braggio (2004). **Signótica**, Goiânia, v. 17, n. 2, p. 251–273, jul./dez. 2005.

BRANDÃO, Ana Paula; Fernando O. de CARVALHO; Everton PEREIRA. Estudo histórico-comparativo preliminar do subgrupo Juruena (Aruák). Trabalho apresentado em 6 de novembro de 2018 no **VI Congresso Internacional de Estudos Linguísticos e Literários na Amazônia** (VI CIELLA). Belém: Universidade Federal do Pará, 2018.

BRANDÃO, Ana Paula; Sidi FACUNDES. Estudos comparativos do léxico da fauna e flora Aruák. **Boletim do Museu Paraense Emílio Goeldi. Ciências humanas**, Belém, v. 2, n. 2, p. 133–168, maio/ago. 2007.

BRINTON, Daniel. **The American Race**. Nova Iorque: N. D. C. Hodges, 1891.

CAMARGOS, Lidiane S. **Consolidando uma proposta de família linguística Boróro**: contribuição aos estudos histórico-comparativos do tronco Macro-Jê. Tese (Doutorado em Linguística) – Universidade de Brasília, Brasília, 2013.

CAMPBELL, Lyle. Review of Language in the Americas, by Joseph Greenberg. **Language**, Baltimore, v. 64, n. 3, p. 591–615, set. 1988.

\_\_\_\_\_. **American Indian languages**: the historical linguistics of Native America. Oxford: Oxford University Press, 1997.

\_\_\_\_\_. Classification of the Indigenous Languages of South America. In: CAMPBELL, Lyle; Verónica GRONDONA (eds.). **The Indigenous Languages of South America**. A Comprehensive Guide. Berlim/Boston: De Gruyter Mouton, 2012. p. 59–166.

CAPTAIN, David. Proto Lokono-Guajiro. **Revista Latinoamericana de Estudios Etnolingüísticos**, Lima, v. 10, p. 137–172, 2005.

CARVALHO, Fernando O. de, 2016a. Obscure cognates and lexical reconstruction: notes on the diachrony of the Xinguan Arawak languages. **Boletim do Museu Paraense Emílio Goeldi. Ciências Humanas**, Belém, v. 11, n. 1, p. 277–294, jan./abr. 2016.

\_\_\_\_\_. Internal and comparative reconstruction in Yawalapiti: Palatalization and rule telescoping. **International Journal of American Linguistics**, Chicago, v. 82, n. 3, p. 285–316, jul. 2016b.

\_\_\_\_\_. On the Development of the Proto-Northern Jê Rhotic in Panará Historical Phonology. **Anthropological Linguistics**, Bloomington, v. 58, n. 1, p. 52–78, 2016c.

\_\_\_\_\_. The diachrony of person-number marking in the Lokono-Wayuunaiki subgroup of the Arawak family: Reconstruction, sound change and analogy. **Language Sciences**, Amsterdã, v. 55, p. 1–15, maio 2016d.

\_\_\_\_\_. Philological evidence for phonemic affricates and diachronic debuccalization in Early Terena (Arawak). **Boletim do Museu Paraense Emílio Goeldi**. Ciências Humanas, Belém, v. 12, n. 1, p. 157–171, jan./abr. 2017a.

\_\_\_\_\_. Fricative debuccalization and primary split in Terena (Arawak) historical phonology. **International Journal of American Linguistics**, Chicago, v. 83, n. 3, p. 509–537, jul. 2017b.

\_\_\_\_\_. Numerals in Lokono-Wayuunaiki: Reconstruction and Implications for Internal Classification. **Folia Linguistica Historica**, Berlim/Boston, v. 51, n. 38, 1–27, 2017c.

\_\_\_\_\_. Tupi-Guarani loanwords in southern Arawak: Taking contact etymologies seriously. **Revista Linguística**, Rio de Janeiro, v. 13, n. 3, p. 41–74, 2017d.

\_\_\_\_\_. On Terena (Arawakan) *-pâho*: Etymology and implications for internal classification. **Journal of Language Relationship**, Moscou, v. 15, n. 2, p. 69–86, 2017e.

\_\_\_\_\_. The historical phonology of Paunaka (Arawakan). **Boletim do Museu Paraense Emílio Goeldi**. Ciências humanas, Belém, v. 13, n. 2, p. 405–428, maio/ago. 2018a.

\_\_\_\_\_. Diachronic split and phoneme borrowing in Resígaro (Arawakan). **Canadian Journal of Linguistics**, Toronto, v. 63, n. 3, p. 1–20, mar. 2018b.

\_\_\_\_\_. Arawakan-Guaicuruan language contact in the South American Chaco. **International Journal of American Linguistics**, Chicago, v. 84, n. 2, p. 243–263, abr. 2018c.

\_\_\_\_\_. Terena (Arawakan) *-eúko* ‘uncle’ and *-ôko* ‘aunt’: Etymology and a kinship terminology puzzle. **Journal of Language Relationship**, Moscou, v. 16, n. 2, p. 79–92, 2018d.

\_\_\_\_\_. Demorphologization of the Proto-Arawakan Privative *\*ma-* in Terena. **LIAMES**, Campinas, no prelo.

\_\_\_\_\_. Revisitando o Proto-Jurúna: a reconstrução da série de oclusivas orais. In: OLIVEIRA, Edna dos Santos; Eduardo Alves VASCONCELOS; Romário Duarte SANCHES (orgs.). **Estudos Linguísticos na Amazônia**. Campinas: Editora Pontes, a sair.

CARVALHO, Fernando O. de; Ana Paula BRANDÃO. 2018. Estudo histórico-comparativo do subgrupo Juruena da família linguístico Arawak. Trabalho apresentado em 24 de maio de 2018 no **I Simpósio de Estudos Linguísticos na Amazônia**. Macapá: Universidade Federal do Amapá (UNIFAP) e Universidade do Estado do Amapá (UEAP), 2018.

CARVALHO, Fernando O. de; Françoise ROSE. Comparative reconstruction of Proto-Mojeño and the phonological diversification of Mojeño dialects. **LIAMES**, Campinas, v. 18, n. 1, p. 3–44, jan./jun. 2018.

CASTRO ALVES, Flávia de. Evolution of alignment in Timbira. **International Journal of American Linguistics**, Chicago, v. 76, n. 4, p. 439–475, out. 2010.

CHACON, Thiago Costa. A revised proposal of Proto-Tukanoan consonants and Tukanoan family classification. **International Journal of American Linguistics**, Chicago, v. 80, n. 3, p. 275–322, jul. 2014.

CHACON, Thiago Costa; Johann-Mattis LIST. Improved computational models of sound change shed light on the history of the Tukanoan languages. **Journal of Language Relationship**, Moscou, v. 13, n. 3, p. 177–203, 2015.

CORRÊA-DA-SILVA, Beatriz Carretta. **Mawé/Awetí/Tupí-Guaraní: relações linguísticas e implicações históricas.** Tese (Doutorado em Linguística) – Universidade de Brasília, Brasília, 2010.

COSTA, Januacele Francisca da; W. Leo M. WETZELS. **Proto-Nambikwara Sound Structure.** Amsterdã: Vrije Universiteit Amsterdam, 2008.

CRÉQUI-MONTFORT, Georges de; Paul RIVET. Linguistique bolivienne. La famille linguistique Čapakura. **Journal de la société des américanistes**, Paris, v. 10, p. 119–171, 1913.

CREVELS, Mily. Language Endangerment in South America: The Clock is Ticking. In: CAMPBELL, Lyle; Verónica GRONDONA (eds.). **The Indigenous Languages of South America. A Comprehensive Guide.** Berlim/Boston: De Gruyter Mouton, 2012. p. 167–234.

CROFT, William. **Radical Construction Grammar: Syntactic Theory in Typological Perspective.** Oxford: Oxford University Press, 2001.

DANIELSEN, Swintha; Michael DUNN; Pieter MUYSKEN. 2011. The spread of the Arawakan languages: A view from structural phylogenetics. In: Alf HORNBORG; Jonathan D. HILL (eds.). **Ethnicity in Ancient Amazonia: reconstructing past identities from archaeology, linguistics, and ethnohistory.** Boulder: University of Colorado Press, 2011. p. 173–196.

D'ANGELIS, Wilmar Rocha. **Pensar o Proto-Jê Meridional e revisitar o Proto-Jê, numa abordagem pragueana.** Relatório (Pós-Doutorado em Linguística Histórica). Brasília: Universidade de Brasília, 2008.

D'ANS, André Marcel. Reclasificación de las lenguas pano y datos glotocronológicos para la etnohistoria de la Amazonía peruana. **Revista del Museo Nacional**, Lima, v. 39, p. 349–369, 1973.

\_\_\_\_\_. Étude glottochronologique de neuf langues Pano. In: CERULLI, Ernesta; Gilda Della RAGIONE (eds.). **Atti Del XL Congresso Internazionale Degli Americanisti, Roma-Genova, 3–10 Settembre 1972**, v. 3. Gênova: Tilgher, 1975. p. 87–97.

DAVIS, Irvine. Comparative Jê phonology. **Estudos Lingüísticos: Revista Brasileira de Linguística Teórica e Aplicada**, São Paulo, v. 1, n. 2, p. 20–24, 1966.

\_\_\_\_\_. Some Macro-Jê Relationships. **International Journal of American Linguistics**, Chicago, v. 34, n. 1, p. 42–47, jan. 1968.

DE LA GRASSERIE, Raoul. De la famille linguistique Pano. In: **Congrès International des Américanistes. Compte rendu de la septième session. Berlin 1888.** Berlim: W. H. Kühl, 1890. p. 438–449.

DIENST, Stefan. The innovation of *s* in Kulina and Deni. **Anthropological Linguistics**, Bloomington, v. 47, n. 4, p. 424–441, 2005.

DIETRICH, Wolf. Correspondências fonológicas e lexicais entre Karitiána (Arikém, Tupí) e Tupí-Guaraní. **Revista Brasileira de Linguística Antropológica**, Brasília, v. 1, n. 2, p. 25–48, dez. 2009.

DIXON, R. M. W. Proto-Arawá phonology. **Anthropological Linguistics**, Bloomington, v. 46, n. 1, p. 1–83, 2004.

DRUDE, Sebastian. On the position of the Awetí language in the Tupí family. In: DIETRICH, Wolf; Haralambos SYMEONIDIS (eds.). **Guaraní y “Mawetí-Tupí-Guaraní”:** Estudios



históricos y descriptivos sobre una familia lingüística de América del Sur. Berlim: LIT Verlag, 2006. p. 47–68.

DURAND, Tom. **L'intransitivité scindée dans les langues arawak**. Tese (Doutorado em Linguística) – Université Sorbonne Paris Cité, Paris, 2016.

DYBO, Anna; George STAROSTIN. In Defense of the Comparative Method, or The End of the Vovin Controversy. In: STAROSTIN, George (ed.). **Aspekty komparativistiki III**. Moscou: Universidade Estatal Russa de Humanidades, 2008. p. 119–258. (Orientalia et Classica, v. 19.)

EPPS, Patience; Katherine BOLAÑOS. Reconsidering the “Makú” Language Family of Northwest Amazonia. **International Journal of American Linguistics**, Chicago, v. 83, n. 3, 467–507, jul. 2017.

FACUNDES, Sidney da Silva. **The language of the Apurinã people of Brazil (Arawak)**. Tese (Doutorado em Linguística) – University of New York at Buffalo, Buffalo, 2000.

\_\_\_\_\_. The comparative linguistic methodology and its contribution to improve the knowledge of Arawakan. In: HILL, Jonathan D.; Fernando SANTOS-GRANERO (eds.). **Comparative Arawakan histories**. Illinois: University of Illinois Press, 2002. p. 74–96.

FARGETTI, Cristina Martins; RODRIGUES, Carmen L. Consoantes do xipaya e do juruna: uma comparação em busca do proto-sistema. **Alfa: Revista de Linguística**, São José do Rio Preto, v. 52, n. 2, p. 535–563, 2008.

FRANÇOIS, Alexandre. Trees, Waves and Linkages: Models of Language Diversification. In: BOWERN, Claire; Bethwyn EVANS (eds.). **The Routledge Handbook of Historical Linguistics**. Nova Iorque: Routledge, 2014. p. 161–189.

GALUCIO, Ana Vilacy; Nilson GABAS JÚNIOR. Evidências de agrupamento genético Karo-Puruborá, tronco Tupi. In: **XVII Encontro Nacional da ANPOLL: boletim informativo nº 31**. Porto Alegre: Universidade Federal do Rio Grande do Sul, 2002. p. 163.

GALUCIO, Ana Vilacy; Sérgio MEIRA; Joshua BIRCHALL; Denny MOORE; Nilson GABAS JÚNIOR; Sebastian DRUDE; Luciana STORTO; Gessiane PICANÇO; Carmen Reis RODRIGUES. Genealogical relations and lexical distances within the Tupian linguistic family. **Boletim do Museu Paraense Emílio Goeldi**. Ciências humanas, Belém, v. 10, n. 2, p. 229–274, maio/ago. 2015.

GALUCIO, Ana Vilacy; Denny MOORE; Hein van der VOORT. O patrimônio linguístico do Brasil: novas perspectivas e abordagens no planejamento e gestão de uma política da diversidade linguística. **Revista do Patrimônio Histórico e Artístico Nacional**, Rio de Janeiro, v. 38, p. 195–219, 2018.

GALUCIO, Ana Vilacy; Antonia Fernanda NOGUEIRA. Comparative study of the Tupari branch of the Tupi family: contributions to understanding its historical development and internal classification. In: **Memorias del V Congreso de Idiomas Indígenas de Latinoamérica, 6–8 de octubre de 2011, Universidad de Texas en Austin**. Austin: University of Texas, 2011.

GILDEA, Spike. A comparative description of syllable reduction in the Cariban language family. **International Journal of American Linguistics**, Chicago, v. 61, n. 1, p. 62–102, jan. 1995.

\_\_\_\_\_. **On Reconstructing Grammar: Comparative Cariban Morphosyntax**. Nova Iorque, Oxford: Oxford University Press, 1998. (Studies in Anthropological Linguistics, v. 18.)

\_\_\_\_\_. Reconstructing Pre-Proto-Tupi-Guarani main clause grammar. In: **Atas do I Encontro Internacional do grupo de trabalho sobre línguas indígenas da ANPOLL**, v. 1. Belém: Universidade Federal do Pará, 2002. p. 315–326.

\_\_\_\_\_. Explaining similarities between main clauses and nominalized phrases. **Amerindia**, Paris, v. 32, p. 57–75, 2008.

GILIJ, Filippo Salvatore. **Saggio di Storia Americana, ossia Storia Naturale, Civile e Sacra dei Regni, e delle provincie Spagnole di Terraferma nell'America meridionale**. 4 v. Roma: Luigi Salvioni, 1780–1784.

GIRARD, Victor. **Proto-Carib phonology**. Tese (Doutorado em Linguística) – University of California, Berkeley, 1971a.

\_\_\_\_\_. **Proto-Takanan phonology**. Berkeley, Los Angeles: University of California Press, 1971b. (University of California Publications in Linguistics, 70.)

GOERGENS, Anne. **Stability and change in alignment systems: a study of agreement patterns in Arawakan languages**. Dissertação (Bacharelado em Linguística) – Lunds Universitet, Lund.

GUDSCHINSKY, Sarah C. Ofaié-Xavánte, a Jê language. In: GUDSCHINSKY, Sarah C. (ed.). **Estudos sobre línguas e culturas indígenas**. Brasília: SIL, 1971. p. 1–16.

GUÉRIOS, R. F. Mansur. A posição lingüística do Xetá. **Letras**, Curitiba, v. 10, p. 92–114, 1959.

GUIRARDELLO, Raquel. **A Reference Grammar of Trumai**. Tese (Doutorado em Linguística) – Rice University, Houston, 1999.

HEELAS, Richard Hosie. **The social organization of the Panara, a Gê tribe of Central Brazil**. Tese (Doutorado em Antropologia) – University of Oxford, Oxford, 1979.

HEITZMAN, Allene. **Reflexes of some Proto-Campa consonants in modern Campan languages and dialects**. Dissertação (Mestrado em Linguística) – University of Kansas, Lawrence, 1973.

HOENIGSWALD, Henry. The Principal Step in Comparative Grammar. **Language**, Baltimore, v. 26, n. 2, p. 357–364, 1950.

\_\_\_\_\_. **Language Change and Linguistic Reconstruction**. Chicago: University of Chicago Press, 1960.

JOLKESKY, Marcelo Pinho de Valhery. **Reconstrução fonológica e lexical do Proto-Jê Meridional**. Dissertação (Mestrado em Linguística) – Universidade Estadual de Campinas, Campinas, 2010.

\_\_\_\_\_. Uma reconstrução do Proto-Mamoré-Guaporé (família Arawak). **LIAMES**, Campinas, v. 16, n. 1, p. 7–37, jan./jun. 2016.

JACQUES, Guillaume; Johann-Mattis LIST. Why we need tree models in linguistic reconstruction (and when we should apply them). **Journal of Historical Linguistics**, Amsterdã, v. 9, no. 1, p. 128–167, 2019.

JENSEN, Cheryl. **O desenvolvimento histórico da língua Wayampí**. Campinas: Editora da Unicamp, 1989.

\_\_\_\_\_. Cross-referencing changes in some Tupí-Guaraní languages. In: PAYNE, Doris L. (org.). **Amazonian Linguistics: Studies in Lowland South American Languages**. Austin: University of Texas Press, 1990. p. 117–158.



\_\_\_\_\_. Comparative Tupí-Guaraní morphosyntax. In: DERBYSHIRE, Desmond C; Geoffrey K. PULLUM (orgs.). **Handbook of Amazonian Languages**, v. 4. Berlin: Mouton de Gruyter, 1998. p. 487-618.

KEY, Mary Ritchie. **Comparative Tacanan phonology. With Cavineña phonology and notes on Pano-Tacanan relationship**. Haia: Mouton, 1968. (Janua Linguarum Series Práctica, v. 50.)

LAFONE QUEVEDO, Samuel. **Idioma Mbaya**. Buenos Aires: Imprenta de Pablo Coni e Hijos, 1896.

LAPIERRE, Myriam. **A phonological analysis of Panãra**. Berkeley: University of California, em preparação.

LEMLE, Miriam. Internal classification of the Tupi-Guarani linguistic family. In: BENDOR-SAMUEL, David (ed.). **Tupi studies I**. Norman: Summer Institute of Linguistics of the University of Oklahoma, 1971. p. 107-129. (Summer Institute of Linguistics Publications in Linguistics and Related Fields, v. 29.)

LIMA-ANGENOT, Geralda de. Subsídios para a glotocronologia lexicostatística da família Chapakura. **CEPLA Working Papers in Amerindian Linguistics**, s/n. Guajará-Mirim: Universidade Federal de Rondônia, 1995.

\_\_\_\_\_. **Fonotática e fonologia do lexema Protochapakura**. Dissertação (Mestrado em Linguística) – Universidade Federal de Rondônia, Guajará-Mirim, 1997.

LOUKOTKA, Čestmír. Une tribu indienne peu connue dans l'état brésilien Paraná. **Acta Ethnographica Academiae Scientiarum Hungaricae**, Budapeste, v. 9, n. 3-4, p. 329-368, 1960.

\_\_\_\_\_. Documents et vocabulaires inédits de langues et de dialectes sud-américains. **Journal de la société des américanistes**, Paris, v. 52, p. 7-60, 1963.

\_\_\_\_\_. **Classification of South American Indian languages**. Los Angeles: University of California Press, 1968.

MALONE, T. **Proto-Tucanoan and Tucanoan genetic relationship**. Instituto Linguístico de Verano, Colômbia (manuscrito), 1986.

MAMIANI, Luiz Vincencio. **Catecismo da Doutrina Christã na Lingua Brasilica da Nação Kiriri**. Lisboa: Officina de Miguel Deslandes, 1698.

\_\_\_\_\_. **Arte de grammatica da lingua brasilica da naçam Kiriri**. Lisboa: Officina de Miguel Deslandes, 1699.

MASON, J. Alden. The Languages of South American Indians. In: STEWARD, Julian H. (ed.). **Handbook of South American Indians**. Vol. 6: Physical Anthropology, Linguistics, and Cultural Geography of South American Indians. Washington: Government Publishing Office, 1950. p. 157-317. (Smithsonian Institution, Bureau of American Ethnology, Bulletin 143.)

MARTINS, Andérbio Márcio Silva. **Uma avaliação da hipótese de relações genéticas entre o Guató e o tronco Macro-Jê**. Tese (Doutorado em Linguística) – Universidade de Brasília, Brasília, 2011.

MARTINS, Valteir. **Reconstrução fonológica do Protomaku Oriental**. Amsterdã: LOT, 2005.

MARTIUS, Carl Friedrich Philipp von. **Beiträge zur Ethnographie und Sprachenkunde Brasiliens. II. Zur Sprachenkunde**. Erlangen: Druck von Junge, 1863.

- MATISOFF, James. On Megalocomparison. **Language**, Baltimore, v. 66, n. 1, p. 106–120, 1990.
- MATTESON, Esther. Proto Arawakan. In: MATTESON, Esther; Alva WHEELER; Frances L. JACKSON; Nathan E. WALTZ; Diana R. CHRISTIAN (orgs.). **Comparative studies in Amerindian languages**. Haia: Mouton, 1972. p. 160–242.
- MEIRA, Sérgio S. C. O. **A Reconstruction of Proto-Taranoan**: Phonology and Inflectional Morphology. Tese (Doutorado em Linguística) – Rice University, Houston, 1998.
- \_\_\_\_\_. Reconstructing Pre-Bakairi segmental phonology. **Anthropological Linguistics**, Bloomington, v. 47, n. 3, p. 1–31, 2005.
- MEIRA, Sérgio; Sebastian DRUDE. A summary reconstruction of Proto-Maweti-Guarani segmental phonology. **Boletim do Museu Paraense Emílio Goeldi**. Ciências humanas, Belém, v. 10, n. 2, p. 275–296, maio/ago. 2015.
- MEIRA, Sérgio; Bruna FRANCHETTO. The Southern Cariban Languages and the Cariban Family. **International Journal of American Linguistics**, Chicago, v. 71, n. 2, 127–192, abr. 2005.
- MEIRA, Sérgio; Spike GILDEA; B. J. HOFF. On the origin of ablaut in the Cariban family. **International Journal of American Linguistics**, Chicago, v. 76, n. 4, p. 411–438, out. 2010.
- MELLO, Antônio Augusto Souza. **Estudo histórico da família lingüística Tupi-Guarani**: aspectos fonológicos e lexicais. Tese (Doutorado em Linguística) – Universidade Federal de Santa Catarina, Florianópolis, 2000.
- MENDES JUNIOR, Djalma Gomes. **Comparação fonológica do Kuruáya com o Mundurukú**. Dissertação (Mestrado em Linguística) – Universidade de Brasília, Brasília, 2007.
- MICHAEL, Lev. La reconstrucción y la clasificación interna de la rama Kampa de la familia Arawak. In: **Memorias del V Congreso de Idiomas Indígenas de Latinoamérica, 6–8 de octubre de 2011, Universidad de Texas en Austin**. Austin: University of Texas, 2011.
- MICHAEL, Lev; Robin ALCORN; Lisa FANN; Briana VAN EPPS; Mona ZARKA. Phonological reconstruction of the Kampan branch of Arawak. Trabalho apresentado em 1 de maio de 2010 no **13º Workshop on American Indigenous Languages**. Santa Barbara: University of California, 2010.
- MICHAEL, Lev; Natalia CHOUSOU-POLYDOURI; Keith BARTOLOMEI; Erin DONNELLY; Vivian WAUTERS; Sérgio MEIRA; Zachary O'HAGAN. A Bayesian Phylogenetic Classification of Tupí-Guaraní. **LIAMES**, Campinas, v. 15, n. 2, p. 193–221, jul./dez. 2015.
- MIGLIAZZA, Ernest C. **Yanomama Grammar and Intelligibility**. Tese (Doutorado em Linguística) – University of Indiana, Bloomington, 1972.
- MOORE, Denny. Classificação interna da família lingüística Mondé. **Estudos Linguísticos**, São Paulo, v. 34, p. 515–520, 2005.
- MOORE, Denny; Ana Vilacy GALUCIO. Reconstruction of Proto-Tupari consonants and vowels. In: LANGDON, Margaret (ed.). **Report 8. Survey of California and Other Indian Languages. Proceedings of the meeting of the Society for the Study of the Indigenous Languages (July 2–4, 1993) of the Americas and the Hokan-Penutian workshop (July 3, 1993)**. Columbus: Ohio State University, 1993. p. 119–137.

MOORE, Denny; Ana Vilacy GALUCIO; Nilson GABAS JÚNIOR. O Desafio de Documentar e Preservar as Línguas Amazônicas. **Scientific American Brasil**, São Paulo, v. 3, p. 36–43, 2008.

NANTES, Bernardo de. **Katecismo Indico da lingua Kariris, acrescentado de varias praticas doutrinaes, & moraes, adaptadas ao genio, & capacidade dos Indios do Brasil**. Lisboa: Officina de Valentim da Costa Deslandes, 1790.

NIKULIN, Andrey. Apofonia e sistema vocálico do Proto-Jê Meridional: contribuição para estudos comparativos das línguas Jê. **Revista Brasileira de Linguística Antropológica**, Brasília, v. 7, n. 2, 275–306, 2015.

\_\_\_\_\_. Historical phonology of Proto-Northern Jê. **Journal of Language Relationship**, Moscou, v. 14, n. 3, p. 165–185, 2016.

\_\_\_\_\_. A phonological reconstruction of Proto-Cerrado (Jê family). **Journal of Language Relationship**, Moscou, v. 15, n. 3, p. 147–180, 2017.

\_\_\_\_\_. A reconstruction of Proto-Jê phonology and lexicon. **Journal of Language Relationship**, Moscou, v. 17, n. 2, 2019.

NIKULIN, Andrey; Andrés Pablo SALANOVA. Northern Jê verb morphology and the reconstruction of finiteness alternations. **International Journal of American Linguistics**, Chicago, v. 85, n. 4, oct. 2019 (a sair).

NIKULIN, Andrey; Mário André Coelho da SILVA. As línguas Maxakalí e Krenák dentro do tronco Macro-Jê. **Cadernos de Etnolingüística (Estudos de Lingüística Sul-Americana)**, v. 7, 2019.

\_\_\_\_\_. A evolução do alinhamento na língua Maxakalí. In: OLIVEIRA, Christiane Cunha de (org.), **Memórias do II Encontro dos Americanistas no Cerrado**. Goiânia: Universidade Federal de Goiás, a sair.

NIMUENDAJÚ, Curt. **Cartas do Sertão de Curt Nimuendajú para Carlos Estevão de Oliveira** (ed. Thekla Hartmann). Lisboa: Museu Nacional de Etnologia and Assírio & Alvim, 2000 [1923–42].

OLIVEIRA, Gabriel Barros Viana de; Ana Suelly Arruda Câmara CABRAL. Mais evidências para a hipótese de Loukotka (1963, 1968). **Fragmentum**, Santa Maria, v. 46, p. 247–285, jul./dez. 2015.

OLIVEIRA, Sanderson Castro Soares de. **Contribuições para a reconstrução do Protopáno**. Tese (Doutorado em Linguística) – Universidade de Brasília, Brasília, 2014.

PAYNE, David L. A classification of Maipuran (Arawakan) languages based on shared lexical retentions. In: DERBYSHIRE, Desmond C.; Geoffrey K. PULLUM (eds.). **Handbook of Amazonian languages**, v. 3. Berlim: Mouton de Gruyter, 1991. p. 355–499.

PICANÇO, Gessiane Lobato. **Mundurukú: phonetics, phonology, synchrony, diachrony**. Tese (Doutorado em Linguística) – University of British Columbia, Vancouver, 2005.

\_\_\_\_\_. A reconstruction of nasal harmony in Proto-Mundurukú (Tupi). **International Journal of American Linguistics**, Chicago, v. 76, n. 4, p. 411–438, out. 2010.

\_\_\_\_\_. Optimality theory and historical phonology: an example from nasal harmony in Mundurukú. **Alfa: Revista de Linguística**, São José do Rio Preto, v. 57, n. 1, p. 257–274, 2013.

PRICE, P. David. The Nambiquara linguistic family. **Anthropological Linguistics**, Bloomington, v. 20, n. 1, p. 14–37, jan. 1978.

QUEIROZ, José Márcio Correia de. **Aspectos da fonologia Dzubukuá**. Dissertação (Mestrado em Linguística) – Universidade Federal de Pernambuco, Recife, 2008.

\_\_\_\_\_. **Um estudo gramatical da língua Dzubukuá, família Karirí**. Tese (Doutorado em Linguística) – Universidade Federal da Paraíba, João Pessoa, 2012.

RAMIREZ, Henri. **Línguas Arawak da Amazônia Setentrional**: comparação e descrição. Manaus: Editora da Universidade Federal do Amazonas, 2001.

RANKIN, Robert L. Review of Language in the Americas, by Joseph H. Greenberg. **International Journal of American Linguistics**, Chicago, v. 58, n. 3, p. 324–351, jul. 1992.

RIBEIRO, Eduardo Rivail. Ofayé, língua Macro-Jê: evidências adicionais. Trabalho apresentado em 28 de setembro de 2005 nos **Colóquios Linguísticos do Museu Antropológico**. Goiânia: Museu Antropológico (Universidade Federal de Goiás), 2005.

\_\_\_\_\_. **A grammar of Karajá**. Tese (Doutorado em Linguística) – University of Chicago, Chicago, 2012.

RIBEIRO, Eduardo Rivail; Hein van der VOORT. Nimuendajú was right: the inclusion of the Jabutí language family in the Macro-Jê stock. **International Journal of American Linguistics**, Chicago, v. 76, n. 4, p. 517–570, out. 2010.

RIBEIRO, Lincoln Almir Amarante. Uma proposta de classificação interna das línguas da família pano. **Investigações**, Recife, v. 19, n. 2, p. 157–188, 2006.

RIVET, Paul. Langues Américaines III: Langues de l'Amérique du Sud et des Antilles. In: MEILLET, Antoine; Marcel COHEN (eds.). **Les Langues du Monde**. Paris: Société de Linguistique de Paris, 1924. p. 639–712. (Coll. Ling. Paris, v. 16.)

RODRIGUES, Aryon Dall'Igna. Die Klassifikation des Tupi-Sprachstammes. In: **Proceedings of the Thirty-Second International Congress of Americanists, 1956**. Copenhagen: Munksgaard, 1958. p. 679–684.

\_\_\_\_\_. Classificação da língua dos Cinta-Larga. **Revista de Antropologia**, São Paulo, v. 14, p. 27–30, 1966.

\_\_\_\_\_. A língua dos índios Xetá como dialeto Guarani. **Cadernos de Estudos Linguísticos**, Campinas, v. 1, p. 7–11, 1978.

\_\_\_\_\_. Tupi-Guarani e Munduruku: evidências lexicais e fonológicas de parentesco genético. **Estudos Linguísticos**, São Paulo, v. 3, p. 194–209, 1980.

\_\_\_\_\_. Evidence for Tupi-Carib Relationships. In: KLEIN, Harriet E. Manelis; Louisa R. STARK (eds.). **South American Indian languages: retrospect and prospect**. Austin: University of Texas Press, 1985. p. 371–404.

\_\_\_\_\_. **Línguas Brasileiras**: para o conhecimento das línguas indígenas. São Paulo: Edições Loyola, 1986.

\_\_\_\_\_. Correspondências lexicais e fonológicas entre Tupí-Guaraní e Tuparí. In: CABRAL, Ana Suelly Arruda Câmara; Aryon Dall'Igna RODRIGUES (orgs.). **Línguas indígenas brasileiras: Fonologia, gramática e história. Atas do I Encontro Internacional do Grupo de Trabalho sobre Línguas Indígenas da ANPOLL**, v. 1. Belém: Editora Universitária UFPA, 2002. p. 288–297.

\_\_\_\_\_. As vogais orais do Proto-Tupi. In: RODRIGUES, Aryon Dall'Igna; Ana Suelly Arruda Câmara CABRAL (orgs.). **Novos estudos sobre línguas indígenas**. Brasília: Editora da UnB, 2005. p. 35–46.

\_\_\_\_\_. As consoantes do Proto-Tupi. In: CABRAL, Ana Suelly Arruda Câmara; Aryon Dall'Igna RODRIGUES (orgs.). **Línguas e culturas Tupí**. Campinas: Editora Curt Nimuendajú, 2007. p. 167–203.

SCHLEICHER, Charles O. **Comparative and internal reconstruction of Proto-Tupi-Guarani**. Tese (Doutorado em Linguística) – University of Wisconsin, Madison, 1998.

SCHMIDT, Max. Guaná. **Zeitschrift für Ethnologie**, Berlim, v. 35, n. 4, p. 560–604, 1903.

SEKI, Lucy. O Krenak (Botocudo/Borum) e as línguas Jê. In: SANTOS, Ludoviko dos; Ismael Pontes (orgs.). **Línguas Jê: Estudos Vários**. Londrina: EDUEL, 2002. p. 15–40.

SHELL, Olive A. **Pano Reconstruction**. Tese (Doutorado em Linguística) – University of Pennsylvania, Filadélfia, 1965.

STAROSTIN, George. Lexicostatistics as a basis for language classification: increasing the pros, reducing the cons. **Classification and Evolution in Biology, Linguistics and the History of Science**. In: FANGERAU, Heiner; Hans GEISLER; Thorsten HALLING; William MARTIN. **Concepts – Methods – Visualization**. Stuttgart: Franz Steiner Verlag, 2013. p. 125–146.

STEINEN, Karl von den. **Die Bakairí-Sprache**. Leipzig: K. F. Koehler's Antiquarium, 1892.

SWADESH, Morris. Lexicostatistic dating of prehistoric ethnic contacts. **Proceedings of the American Philosophical Society**, Philadelphia, v. 96, p. 452–63, 1952.

\_\_\_\_\_. Towards greater accuracy in lexicostatistic dating. **International Journal of American Linguistics**, Chicago, v. 21, p. 121–137, abr. 1955.

TAUNAY, Alfredo d'Escagnolle. Vocabulário da língua Guaná ou Chané (província de Matto Grosso). **Revista Trimensal do Instituto Histórico Geographico e Ethnographico do Brasil**, Rio de Janeiro, v. 38, p. 143–162, 1875.

URBAN, Greg. A história da cultura brasileira segundo as línguas nativas. In: CUNHA, Manuela Carneiro da (org.). **História dos índios no Brasil**. São Paulo: Companhia das Letras, Secretaria Municipal de Cultura, FAPESP, 1992. p. 87–102.

VASCONCELOS, Eduardo Alves. **Investigando a hipótese Cayapó do Sul-Panará**. Tese (Doutorado em Linguística) – Universidade Estadual de Campinas, Campinas, 2013.

VIEGAS BARROS, J. Pedro. **Proto-Guaicurú. Una reconstrucción fonológica, léxica y morfológica**. Munique: LINCOM Europa, 2013. (LINCOM studies in Native American linguistics, v. 69.)

VOORT, Hein van der. Proto-Jabutí: um primeiro passo na reconstrução da língua ancestral dos Arikapú e Djeoromitxi. **Boletim do Museu Paraense Emílio Goeldi**. Ciências humanas, Belém, v. 2, n. 2, p. 133–168, maio/ago. 2007.

WALKER, Robert S.; Lincoln A. RIBEIRO. Bayesian phylogeography of the Arawak expansion in lowland South America. **Proceedings of the Royal Society of London B: Biological Sciences**, Londres, v. 278, n. 1718, p. 2562–2567, 2011.

WALTZ, Nathan E.; Alva WHEELER. Proto-Tucanoan. In: MATTESON, Esther; Alva WHEELER; Frances L. JACKSON; Nathan E. WALTZ; Diana R. CHRISTIAN (orgs.). **Comparative studies in Amerindian languages**. Haia: Mouton, 1972. p. 19–49.

WIESEMANN, Ursula. Os dialetos da língua Kaingáng e o Xoklég. **Arquivos de Anatomia e Antropologia**, Rio de Janeiro, v. 3, p. 199–217, 1978.

ZAMPONI, Raoul. 2003. **Maipure**. Munique: LINCOM Europa. (Languages of the World, v. 192.)

---

## Para citar este artigo

---

NIKULIN, A., CARVALHO, F. O. de. Estudos diacrônicos de línguas indígenas brasileiras: um panorama. **MACABÉA – REVISTA ELETRÔNICA DO NETLLI**, CRATO, V. 8., N. 2., 2019, p. 255-305.

305

---

## Os Autores

---

**Andrey Nikulin** é especialista em Linguística Teórica e Aplicada pela Universidade Estatal de Moscou (2015) e doutorando pela Universidade de Brasília (UnB). Tem experiência nas áreas de linguística teórica (com ênfase em linguística histórico-comparativa e fonologia) e aplicada, atuando principalmente nas seguintes áreas: línguas Macro-Jê, língua Chiquitana, reconstrução fonológica, linguística histórica, linguística descritiva, fonologia areal, sintaxe diacrônica.

**Fernando O. de Carvalho** é pesquisador/professor da área de Linguística na Universidade Federal do Amapá (UNIFAP). Pós-doutor pelo Museu Nacional da Universidade Federal do Rio de Janeiro (Setor de Linguística – MN/UFRJ/2014–2016). Possui doutorado em Linguística pela Universidade Federal do Rio de Janeiro (2014), com estágio de doutorado no Departamento de Linguística do Max Planck Institut für Evolutionäre Anthropologie (MPI/EVA, Leipzig, Alemanha) e mestrado em Linguística pela Universidade de Brasília (UnB). Tem experiência na área de Linguística, com ênfase em linguística histórica, fonética/fonologia e morfologia, atuando principalmente nos seguintes temas: linguística histórica, línguas indígenas, línguas africanas, teoria linguística, fonologia/fonética e morfologia.