

A HISTÓRIA QUE CONTA O LÉXICO MÊBÊNGÔKRE

Andrés Pablo Salanova,¹
Andrey Nikulin²

RESUMO

O propósito deste artigo é o de identificar o léxico de origem não nativa na língua indígena Mêbêngôkre (família Jê, tronco Macro-Jê), falada pelos povos Kayapó e Xikrin no Pará, bem como pelos Kayapó Mêtýktire (Txucarramãe) no norte de Mato Grosso. São propostas novas etimologias de contato para diversas palavras Mêbêngôkre, incluindo *waxi* ‘linha’, *benorã* ‘tucunaré’, *awo* ‘tipo de árvore; barco de cortiça, ubá’, *ixe* ‘espelho’ (da língua Xambioá); *môtôbi’y* ‘amendoim’, *xãn* ‘gato’, *mokà* ‘mocó’, *xoko* ‘socó-boi’ (da Língua Geral Amazônica), *karaxu* ‘colher’ (da língua Yudjá), dentre outras; além disso, são identificados alguns empréstimos de origem Mêbêngôkre em outras línguas. O artigo conclui com uma proposta de síntese da história migratória dos povos Jê Setentrionais baseada tanto nos resultados deste estudo como em publicações anteriores.

Palavras-chave: Mêbêngôkre, contato de línguas, empréstimos lexicais, línguas Jê.

1. Introdução

Os povos Jê do Brasil central tiveram sempre uma postura de feroz independência diante do avanço da colonização portuguesa e neobrasileira, e isto fez com que eles fossem temidos e considerados guerreiros irredutíveis pelas populações vizinhas a eles. Para além desse temor, um preconceito persistente que se cristalizou na época colonial considerava que os povos não-Tupí do sertão seriam sociedades pequenas e muito móveis, com cultura material simples, e fechadas sobre si mesmas.³

¹ Professor do Departamento de Linguística da Universidade de Ottawa. Doutor em Linguística pelo Instituto de Tecnologia de Massachusetts. Email: <kaitire@uottawa.ca>.

² Pesquisador Colaborador do Programa de Pós-Graduação em Linguística da Universidade de Brasília (PPGL/UnB). Doutor em Linguística pela mesma universidade. Integrante do Núcleo de Tipologia e Línguas Indígenas (NTLI). Email: <andrey.nikulin@unb.br>.

³ A fama de ferozes defensores de seu território não anula o fato de que, nos casos mais bem documentados, os povos Jê tiveram em um primeiro momento trato amistoso com os colonizadores, voltando-se contra estes somente após terem sido hostilizados sem provocação. Para os Kayapó do Sul ou Bilreiros, vide Neme (1969), onde também se encontram interessantes observações sobre sua mobilidade e suposta simplicidade na agricultura e cultura material: em resumidas contas, os Kayapó do Sul seriam horticultores sedentários, morando em grandes aldeias, e sua aparente mobilidade e simplicidade

É certo que o movimento dêmico que vários povos Jê fizeram em direção ao oeste em tempos históricos, afastando-se das intrusões da sociedade neobrasileira, foi acompanhado de uma grande fragmentação de grupos, e de hostilidades tanto entre parcialidades como para com as populações que se encontravam em seu caminho, fossem eles outros grupos indígenas ou neobrasileiros. Porém, estas hostilidades documentadas e a reputação dos povos Jê durante a colônia e em tempos mais recentes não deveriam nos impedir de ver que em muitos casos houve intercâmbios culturais que resultavam de contatos pacíficos entre povos de línguas diferentes, mesmo à sombra do avanço da colonização. Já a ideia de que os povos Jê seriam fechados sobre si mesmos é tão distante da realidade etnográfica que sequer merece discussão: muito pelo contrário, a incorporação de valores do exterior, adquiridos por meios pacíficos ou predatórios, seria uma característica essencial da cosmologia de certos povos Jê (cf. GORDON, 2006, p. 95–103; LEA, 2012, p. 376 *et cap.* 9, *passim*, entre outros).

Neste artigo nos voltamos aos Xikrin e aos Kayapó, duas nações indígenas de fala Jê que compartilham uma mesma autodenominação — Mëbêngôkre — e uma mesma língua. Muitos dos povos indígenas que mais tiveram contato com os Mëbêngôkre em tempos recentes — os Iny/Karajá, os Yudjá/Juruna, os Apyãwa/Tapirapé, os Panará, os Arara e os povos Tupí-Guaraní da região entre o baixo Xingu e o baixo Araguaia — ainda guardam na memória um passado de hostilidades com os Mëbêngôkre. Houve, porém, momentos de trocas pacíficas com ao menos alguns deles. Neste artigo investigamos o reflexo destas trocas (e talvez da incorporação predatória de objetos) no léxico da língua Mëbêngôkre. O estudo do léxico, por sua vez, permite fazer novas suposições sobre as situações de contato anteriores à história documentada.

Os Mëbêngôkre e os povos Jê

Não há dados precisos recentes quanto ao número total de falantes de Mëbêngôkre; uma estimativa baseada nos dados do SIASI/SESAI (2014) para os Kayapó e da FUNASA (2010) para os Xikrin é de aproximadamente 13500 falantes. O Mëbêngôkre é língua cooficial no município de São Félix do Xingu (2019).

Costuma-se incluir o Mëbêngôkre no sub-ramo Setentrional da família linguística

tecnológica seria uma característica dos pequenos grupos expedicionários com os quais os colonizadores tinham contato mais frequente, e que se tornaram mais frequentes à medida que o grupo foi hostilizado pelos colonizadores.

Jê, juntamente às línguas Kĩsêdjê (Suyá), Tapayúna (Kajkwakhrattxi), Apinajé, e às múltiplas variedades do complexo linguístico Timbira, tais como Parkatêjê, Gavião-Pyhcopji (Pykobjê), Krĩkatí, Canela e Krahô. Na classificação mais recente (NIKULIN, 2020), as línguas Jê Setentrionais integram o sub-ramo chamado Jê de Goyaz, do ramo Cerratense, como se mostra na Figura 1.

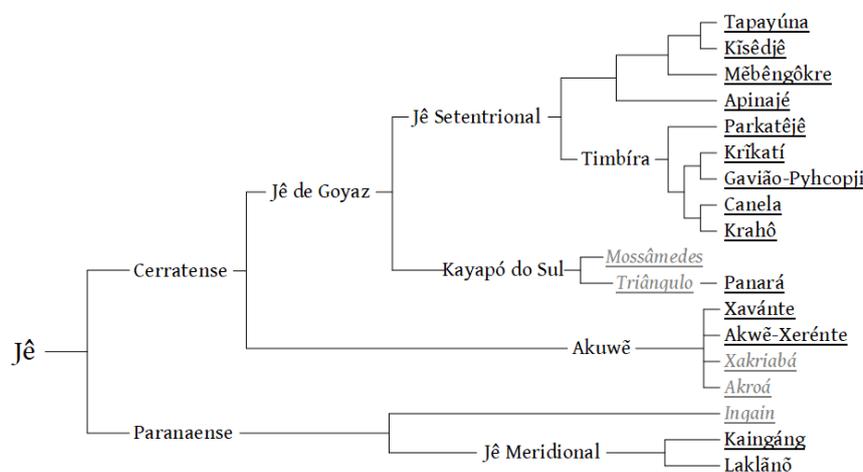


Figura 1. Cladograma da família Jê (adaptado de NIKULIN, 2020)⁴

Dentro do sub-ramo Jê Setentrional, o Mẽbêngôkre parece ser mais próximo às línguas Kĩsêdjê e Tapayúna, hoje faladas no Alto Xingu, porém originárias da bacia do Tapajós; o agrupamento constituído pelas línguas Mẽbêngôkre, Kĩsêdjê e Tapayúna pode ser rotulado de Trans-Araguaia. Essas três línguas compartilham 82–86% de cognatos na lista de 110 palavras de Swadesh (NIKULIN, 2019b). Por sua vez, as línguas Trans-Araguaia são mais próximas ao Apinajé, com o qual compartilham 74–83% de cognatos na mesma lista, do que qualquer uma delas é às variedades Timbira (69–80%). Dessa forma, o sub-ramo Jê Setentrional se subdivide em dois grandes agrupamentos: Timbira e o que poderíamos chamar de Trans-Tocantins (incluindo o Apinajé, o Mẽbêngôkre, o Kĩsêdjê e o Tapayúna). Essas relações encontram-se representadas na Figura 2.⁵

⁴ Nas Figuras 1 e 2, os rótulos sublinhados se referem a variedades linguísticas atestadas. Os demais rótulos dizem respeito a agrupamentos genéticos específicos. As variedades cujos rótulos estão grafados em cinza e itálico se encontram em desuso (extintas ou dormentes).

⁵ Embora não seja fácil situar no tempo os eventos de diversificação demica e linguística, é possível obter datações aproximadas por meio da aplicação da *glotocronologia*, uma ferramenta desenvolvida especialmente para este fim. Aplicando a fórmula elaborada por Vasilyev e Saenko (2017) às porcentagens de cognatos na lista de 110 palavras de Swadesh citadas no início deste artigo, temos que os ancestrais dos Mẽbêngôkre se separaram daqueles dos Kĩsêdjê e dos Tapayúna entre os séculos VIII e X, os Apinajé teriam se separado do grupo ancestral desses três povos entre os séculos V e IX, já a cisão dos falantes do Proto-Jê Setentrional em Timbira e Trans-Tocantins teria acontecido entre os séculos II e VII.

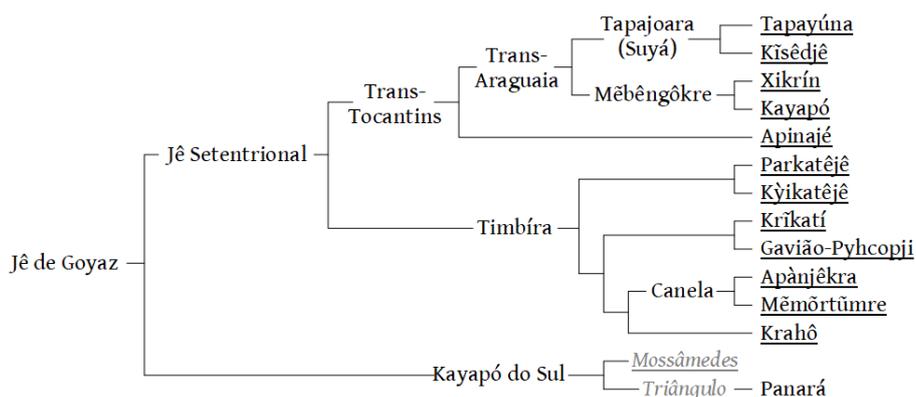


Figura 2. Cladograma do ramo Jê de Goyaz (adaptado de NIKULIN, 2020)

O território ancestral dos Jê Setentrionais

Os povos falantes de línguas Jê Setentrionais ocupam tradicionalmente os cerrados do Brasil central e algumas regiões adjacentes da floresta amazônica. Na história oral dos Mẽbêngôkre, mas também de outros povos Jê Setentrionais, a aldeia ancestral estaria localizada ao leste do território atual (os Mẽbêngôkre dizem às vezes “à beira do mar”, e, jocosa ou provocadoramente, “no Rio de Janeiro”). Logo depois de um conflito ou um evento traumático, os diferentes povos se dispersaram em várias direções, falando cada um sua língua.⁶

No que tange aos Mẽbêngôkre, se por um lado quase não há dúvidas de que o território ancestral dos povos Jê de Goyaz estivesse ao leste do rio Araguaia, a antiguidade da migração em direção ao oeste, e em particular a travessia do Araguaia pelos ancestrais dos Mẽbêngôkre, não é conhecida com exatidão, e de modo geral é pouco o que se sabe sobre a história desse povo, mesmo em tempos relativamente recentes.

Estas datações diferem profundamente em magnitude das datações propostas no estudo glotocronológico de Rumsey (1971), no qual se baseia Turner (1998). Rumsey indica uma separação de 320 anos entre o Mẽbêngôkre e o Tapayúna (Beijo-de-Pau), de 360–470 anos entre estas línguas e o Apinajé e de 500–610 anos entre as três e o Timbíra, coincidindo, portanto, com a estrutura da família que damos acima, apesar da menor profundidade temporal. Urban (1998, p. 90–1) oferece igualmente algumas datas para a separação dos diversos ramos Macro-Jê: a separação entre as línguas Jê de Goyaz e Akuwẽ teria ocorrido entre 1000 e 2000 anos atrás, a diversificação do ramo Jê de Goyaz teria 1000 anos de profundidade, e as diferentes variedades do Timbíra e do Mẽbêngôkre teriam 500 anos ou menos de profundidade. Urban não fornece maiores detalhes sobre sua visão da estrutura interna do ramo Jê Setentrional. Finalmente, existe também a glotocronologia preliminar da família Jê de Wilbert (1964), criticada lapidarmente por Leite (1966).

⁶ Na tradição oral Mẽbêngôkre, o evento é a derrubada coletiva de um gigantesco pé de milho, após o qual cada grupo arranca uma espiga e parte em uma direção diferente. Na tradição oral Timbíra, o que teria provocado a separação dos vários povos da atualidade fora um conflito entre crianças no interior da aldeia ancestral (cf. CROCKER, 1990, p. 305–6).

Das muitas perguntas que podem ser formuladas sobre a história dos Mëbêngôkre, algumas poucas podem achar resposta, mesmo que especulativa, nos resultados da pesquisa linguística. Por exemplo, quando é que os diferentes grupos linguísticos de origem comum se separaram? Com que grupos estiveram em contato? Em que território(s) habitavam? Alguns exemplos de como a linguística responde a estas perguntas serão discutidos na seção 2. Na seção 3 do artigo, fazemos um brevíssimo resumo da história conhecida dos Mëbêngôkre, tal como pode ser reconstruída a partir de fontes históricas externas e por meio de uma interpretação histórica dos relatos dos próprios Mëbêngôkre, feita principalmente por Verswijver (1992; vide TURNER, 1998 e LEA, 2012, seção 1.6 para um resumo acessível). A seção 4 apresenta a fonologia do Mëbêngôkre, com uma atenção especial às estruturas fonológicas atípicas dessa língua, indicadoras de uma potencial origem não nativa dos itens lexicais em que estas se encontram. Na seção 5, propomos novas etimologias de contato, as quais vinculam palavras Mëbêngôkre a fontes específicas em outras línguas originárias da região, pertencentes às famílias Iny/Karajá (5.1), Tupí-Guaraní (5.2), Jê (5.3) e Jurúna (5.4). As palavras que são prováveis empréstimos, mas cuja fonte não pôde ser determinada, são listadas na subseção 5.5. O contato entre o Mëbêngôkre e o português brasileiro será discutido na seção 6. Na seção 7 avançamos uma hipótese sobre as migrações históricas e os contatos exteriores dos Mëbêngôkre e de seus ancestrais, baseando-nos nos dados discutidos nas seções anteriores. A seção 8 conclui o artigo.

2. A linguística e a história

Não podemos nos deter aqui nos detalhes do método comparativo, que é o procedimento que permite identificar os grupos de línguas que possuem origem comum e reconstruir a língua falada pelo grupo originário (vide GABAS JR., 2001 para uma descrição geral do método comparativo em português e NIKULIN, CARVALHO, 2019 para uma discussão das aplicações desse procedimento às línguas indígenas brasileiras). Em vez disso, abordamos dois pontos que são relevantes para nós aqui, a saber, (i) as maneiras de detectar empréstimos e de demonstrar a origem alógena de uma determinada palavra, bem como (ii) as contribuições dos estudos etimológicos para o nosso entendimento das migrações (pré-)históricas e a reconstrução do lugar de origem de um determinado grupo.

A identificação de empréstimos

Ao investigar a origem de um determinado item lexical em uma determinada língua, o/a linguista coloca diante de si a tarefa de rastrear sua história (incluindo suas transformações formais, semânticas e de outros tipos) a partir do momento da criação da palavra (seja em um estágio anterior da mesma língua, seja em uma língua diferente) até o presente (ou, no caso das palavras em desuso, até sua última atestação). Essa história é conhecida sob o nome de **etimologia**. A situação em que é possível desvendar a totalidade da história de uma determinada palavra é relativamente incomum (nas famílias linguísticas que ainda não foram objeto de pesquisas etimológicas exaustivas), mas em muitos casos podemos rastrear pelo menos uma parte dessa história por meio da comparação com outras línguas. *Grosso modo*, as etimologias podem ser subdivididas em dois tipos: algumas palavras foram em algum momento emprestadas de outras línguas, ao passo que outras vieram sendo transmitidas, através de gerações, dentro de uma mesma tradição linguística desde o momento de sua criação (por “criação” entende-se a formação de um novo vocábulo a partir de elementos preexistentes na língua por meio de mecanismos morfológicos produtivos, assim como o surgimento de formações onomatopaicas). As etimologias do primeiro tipo (palavras “não nativas”) são justamente aquelas que nos interessam neste trabalho, e as respectivas premissas teóricas são discutidas em detalhe por Boček (2013) e Mailhammer (2013, 2015), dentre outros.

A tarefa de demonstrar a validade de uma etimologia de contato é frequente — embora equivocadamente — considerada trivial (cf. a discussão em BOČEK, 2013), porém diversos autores alertam à existência de determinados critérios rígidos aos quais qualquer proposta de uma etimologia de contato deve atender. Para Carvalho (2017, p. 44), esses critérios incluem:

- a. a ausência de uma etimologia interna (nativa) para o suposto empréstimo;
- b. a identificação da língua-fonte;
- c. a identificação da forma-fonte;
- d. a elaboração de uma explicação tão detalhada quanto for possível para qualquer desvio de forma ou significado entre o empréstimo e sua suposta fonte;
- e. a identificação da situação de contato.

Portanto, não é suficiente apenas apontar a semelhanças formais entre vocábulos de duas (ou mais) línguas para demonstrar que essas semelhanças são produto do contato de línguas: quanto menos das condições (a–e) forem cumpridas, mais preferível se torna a possibilidade de que a semelhança observada seja espúria, acidental ou genética. Por exemplo, embora a palavra para ‘pilão’ em Mëbêngôkre, *kawa* /kawa/, se assemelhe superficialmente a Javaé *kowo* /kowo/ ‘pilão’, não há qualquer motivo para atribuir essa semelhança ao contato de línguas, pois as condições (a) e (d) acima são violadas. Primeiramente, nas duas línguas, os respectivos termos têm etimologias internas: a palavra Mëbêngôkre vem de Proto-Jê Setentrional **kacû*, já a palavra Javaé vem de Proto-Karajá **/kowoku/*, com ambas as oclusivas velares preservadas no generoleto feminino dos dialetos Karajá e Xambioá (cf. RIBEIRO, 2012, p. 157). Quanto à condição (d), note que as vogais em Mëbêngôkre *kawa* e Javaé *kowo* são diferentes (ambas as línguas têm tanto /a/ como /o/ em seu inventário), o que constituiria um empecilho para uma eventual proposta etimológica que relacionaria essas duas formas uma à outra. Portanto, a semelhança entre Mëbêngôkre *kawa* ‘pilão’ e Javaé *kowo* ‘pilão’ deve ser vista como acidental.

A identificação de vocábulos emprestados torna-se ainda mais desafiadora quando se trata de um empréstimo de uma língua estreitamente relacionada, pois neste caso qualquer semelhança formal entre duas (ou mais) línguas pode ser atribuída, a princípio, a uma origem comum ou ao contato linguístico. Neste caso os desvios dos padrões de correspondências sonoras regulares entre as línguas podem nos fazer suspeitar da origem alógena de uma palavra. Por exemplo, embora os termos *karinhô* ‘fumo’ e *xururu* ‘bico-de-brasa’ da língua Mëbêngôkre se assemelhem a seus equivalentes nas demais línguas Jê Setentrionais (Proto-Jê Setentrional **karên* ‘fumo’, preservado em Apinajé, Kĩsêdjê, Tapayúna e Gavião-Pyhcopji, e **jôrôr* ‘bico-de-brasa’, preservado em Apinajé, Parkatêjê e Canela),⁷ as correspondências sonoras entre o Mëbêngôkre e outras línguas Jê Setentrionais são inteiramente irregulares nesses dois itens do léxico. Portanto, é razoável supor que as palavras *karinhô* e *xururu* não foram herdadas pelo Mëbêngôkre da protolíngua; ao invés disso, elas parecem ter sido emprestadas (direta ou indiretamente) de uma outra língua Jê, talvez não mais existente, em que as vogais **ê* e **ô* eram

⁷ As formas reconstruídas neste trabalho seguem a proposta reconstrutiva de Nikulin e Salanova (2019) com pequenas modificações e estão representadas no Alfabeto Macro-Jê (NIKULIN, 2020, p. 50–3). As formas das línguas contemporâneas se citam nas ortografias usadas pelos respectivos povos.

refletidas como as altas *i*, *u* e a consoante **ʃ* possuía um reflexo surdo (à semelhança do que acontece nas línguas dos Gavião-Pyhcopji, Krĩkatí, Akwê-Xerénte e Xavánte).

Em alguns casos, a própria forma fonológica de uma palavra pode ser indicadora de sua provável origem não nativa, mesmo que a fonte do suposto empréstimo seja desconhecida. Isto ocorre quando a palavra em questão contém fonemas, sequências de sons ou estruturas fonotáticas que não podem ser derivadas de maneira regular a partir de nenhuma forma reconstruída para um estágio anterior da mesma língua. Mesmo se a história fonológica de uma determinada língua não é bem compreendida, muitas vezes é possível identificar padrões fonológicos mais e menos “robustos”, sendo que as palavras que apresentam uma forma mais atípica podem ser consideradas “suspeitas” de serem emprestadas de outras línguas. Na seção 4, fazemos um resumo da fonologia do Mëbêngôkre e sintetizamos os padrões fonológicos menos comuns, indicadores da provável origem alógena das palavras que os instanciam, a partir do nosso conhecimento da fonologia tanto sincrônica quanto diacrônica do Mëbêngôkre.

O léxico e as rotas migratórias

Os estudos etimológicos do léxico de uma determinada língua podem ajudar a reconstruir os deslocamentos dêmicos de seus falantes ao longo do tempo. Pistas relevantes podem ser contidas nos mecanismos derivacionais empregados por uma língua para a formação de termos que denotam espécies de plantas ou animais. Moore e Storto (2002) discutem essa questão em algum detalhe e fornecem o seguinte exemplo do dialeto Xambioá da língua Iny/Karajá:

“A língua karajá fornece evidência lingüística de um outro tipo sobre a localização original do grupo. Os karajás são divididos em quatro grupos, que falam dialetos diferentes. Três dos grupos moram no cerrado e o outro, os xambioá, moram em uma região mais amazônica. Segundo Eduardo Ribeiro (comunicação pessoal) o dialeto dos xambioá deriva novas palavras para novas plantas da Amazônia na base de palavras para espécies do cerrado, por exemplo, o termo para ‘açai’, *kabirini*, é derivado do termo para ‘bacaba’, *kabiri*. Os quatro dialetos têm palavras parecidas para espécies do cerrado, e os mitos de origem, que são compartilhados pelos quatro grupos mencionam animais (seriema) e plantas (mangaba) típicas do cerrado. A inferência clara é que a localização original foi o cerrado e que os xambioá se deslocaram para Amazônia” (MOORE, STORTO, 2002, p. 73).

É interessante observar que a mesma lógica descrita por Moore e Storto (2002, p. 73) para os Xambioá pode ser aplicada, *mutatis mutandis*, para os Mëbêngôkre. Nesta língua, o termo para ‘açai’, *kamêrêkàk*, também é derivado do termo para ‘bacaba’, *kamêrê*, por meio da sufixação de *-kàk*, um morfema que poderia ser traduzido como ‘pseudo-’ (como veremos na seção 5.1, as próprias palavras *kabiri*, da língua Iny/Karajá, e *kamêrê*, do Mëbêngôkre, compartilham uma mesma origem). A língua ancestral do Mëbêngôkre, o Proto-Jê Setentrional, possuía termos próprios para ambas as espécies, **kambêr* ‘bacaba’ e **têr* ‘açai’. Podemos concluir, portanto, que o grupo que deu origem aos Mëbêngôkre permaneceu dentro da área de distribuição da bacabeira desde sua separação dos demais povos Jê Setentrionais, mas muito possivelmente chegou a abandonar a área de distribuição do açazeiro em algum momento (perdendo, assim, a denominação original **têr*); quando posteriormente os Mëbêngôkre chegaram a uma região mais amazônica onde novamente se encontraria o açai, um novo termo para o açai foi recriado a partir da antiga denominação da bacaba.

Nos exemplos acima, o Xambioá e o Mëbêngôkre fizeram uso de modelos morfológicos internos a suas respectivas línguas para derivar nomes de espécies que os antepassados dos falantes dessas línguas aparentemente desconheciam. Há ainda uma segunda possibilidade de preencher lacunas lexicais que emergem quando do deslocamento de um grupo a um novo bioma: é possível que os falantes adotem zoônimos, fitônimos e outros termos da(s) língua(s) falada(s) pelo(s) povo(s) que já se encontrava(m) neste novo bioma à época da chegada do grupo recém-migrado. Por exemplo, Carvalho (2018, a sair) identifica diversos empréstimos da língua Kadiwéu (família Guaicuru) na língua Terena (família Aruak) que designam espécies aquáticas e até mesmo formações hidrográficas, concluindo, a partir dessas evidências, que a rota migratória dos Terena incluiu uma passagem por uma região árida (o sopé dos Andes) antes de sua chegada ao Alto Paraguai, uma região extremamente úmida. Em Mëbêngôkre, ainda não foram identificados estratos lexicais desse tipo, mas há vários termos para flora e fauna, alguns especificamente de floresta amazônica, que não se encontram em línguas próximas. Alguns destes termos têm uma estrutura fonológica ligeiramente surpreendente, que leva a supor uma origem não nativa ou, em certos casos, onomatopaica: *’êtuwa* ‘calango’, *mokokti* ‘poraquê’, *mrê’êti* ‘maracajá’, *dyjdyj* ou *dujduj* ‘ave sp.’, *pārāpārā* ‘gafanhoto’.

O léxico e o território originário

De maneira inversa, por assim dizer, os itens do léxico que podem ser reconstruídos na proto-língua de uma família linguística informam sobre as características do ambiente em que vivia o povo ancestral dos falantes das línguas da família, além de fornecer informações sobre sua cultura (EPPS, 2015). Para o Proto-Jê Setentrional, é possível reconstruir um grande número de vocábulos para espécies de animais e plantas, incluindo os seguintes: tamanduá (*pət), veado-campeiro/galheiro (*mbo), guariba (*kubyt), capivara (*kūmtūm), tatu (*tôn), lontra (*ndê), veado-catingueiro (*karə), onça (*rop), rato (*apnĵô), morcego (*njêp), veado-mateiro (*aĵə-ʔti), veado (*njĵajĵy), raposa (*ĵo), tatu-canastra (*apĵêt), coelho (*kaj), quati (*cwakō), macaco (*kukôj), cutia (*kukêñ), anta (*kukryt), mucura/gambá (*kro-ʔti), papa-mel/irara (*krokrok), preá (*krô), paca (*ngra), porco-espinho/cuandu (*ngroj), caititu (*(a)nggrô), tatu-peba (*(a)nggrê); passarinho (*kuwêñ), ave grande/gavião (*cək), arara (*mbən), seriema (*mbrêk ~ *mbjêk), garça (*kambri), jaó (*aʔtor), caburé (*tôʔtôt), pomba (*tut), mutum (*mbutêk ~ *mbytêk), bico-de-brasa (*ĵôrôr), xexéu (*pêcə(-re/-ti)), curica (*côj-re), pica-pau (*njĵaj), urubu (*njôñ), sabiá (*njwâk), beija-flor (*juñ), perdiz (*bêkê), papagaio (*krwâj), periquito (*krê), tucano (*nrô); jabuti (*kaprăn), jacaré (*mĩ), cascavel (*(a)bat), coruja (*bã), camaleão (*wet), sucuri (*ro-ʔti), teju (*mbryĵūm), jiboia (*cəka), camaleão (*kōk), cobra (*kanĵ), peixe (*tep), piranha (*ambən), poraquê (*mbôp ~ *mbâp), arraia (*mbjê(C)tĵêt), jeju (*kundap), acará (*krăn), traíra (*krwât); caranguejo (*mbaj), escorpião (*mbak), formiga (*mbrūm), abelha-tubi (*cimbrê), borboleta (*wewe), abelha-manuel-de-abreu (*patom), minhoca (*kutō(j)), carrapato (*tê), caramujo (*ndwân), carapanã/mutuca (*pucu), aranha (*cê), marimbondo/caba (*njy, *apnĵy), gafanhoto (*ĵyĵyk ~ *njynĵyk), cigarra (*koʔkot), mosca (*kôp), centopeia (*kêkêk), abelha-arapuá (*kukrã(j)), cupim (*krir) e cupinzeiro (*rôrô ~ *rôr), abelha-tataíra (*ka(j)ngər), besouro (*kunĵn), piolho (*ngô), lagarta (*apngô); urucum (*py), bambu (*po), pequi (*prñ), feijão (*mbən-krwât), cará (*mbôp ~ *mbâp), jatobá (*mbôc), taturubá/cutite (*kambôk), bacaba (*kambêr), jenipapo (*mbrô-ti), capim-navalha/tiririca (*kaba) e sua semente (*angə), pimenta (*bər-cy), capim (*bô), milho (*bô-cy), mangaba (*(a)bêñ), patí (*wô), sororoca/banana-brava (*tyr), pau-d'arco (*tôk-re), embaúba (*atwâr), muruci (*kutê), açai (*têr), amescla-roeira (*rəm),

macaúba/tucum (*rõñ ~ *roñ), coco (*rõr), cará (*krerô), fumo (*karên), amendoim (*kacy), algodão (*kajæt), ingá (*kôkñõ- ~ *kôkñõ-), batata-doce (*jæt), macaxeira (*kwâr), erva-daninha (*a[?]kêt), axixá (*kagrã(j)), timbó (*a[?]krô), caju (*a[?]krwât), cumaru/castanha-de-burro (*krêp), buriti (*ηgrû ~ *ηgrwa). Sobrepondo a distribuição geográfica daquelas espécies da lista que têm uma distribuição mais limitada, seria, em princípio, possível ter uma ideia mais precisa de qual seria o território original do povo ancestral, muito embora a tarefa seja complicada pela inexistência de informações detalhadas acerca da distribuição das diferentes espécies no passado. No caso do Proto-Jê Setentrional, pode-se afirmar com um alto grau de certeza que seus falantes habitavam o norte do Cerrado, conforme evidenciado pela presença de espécies tais como o pequizeiro, o cumaru, o taturubá/cutite, a amescla-aroeira e outras.

3. Breve história dos Mëbêngôkre

Como já mencionamos acima, é relativamente consensual que a região de origem dos Mëbêngôkre era ao leste do Araguaia. No entanto, desde as primeiras referências claras (a partir de meados do século XIX), eles são localizados já ao oeste desse rio, divididos em quatro grupos principais: Porekry ou Purukarwýt (ancestrais dos Xikrin), Gorotire (ancestrais dos Kayapó), e dois grupos desaparecidos, os Irã'ãmrãnhre e os antigos Xikrin ou Djore, dos quais alguns sobreviventes se juntaram aos Gorotire e aos Purukarwýt, respectivamente (VIDAL, 1977, p. 28; VERSWIJVER, 1992, p. 86–7, fig. 12).

O momento em que os Mëbêngôkre atravessaram o Araguaia para se estabelecer em seu território atual é uma incógnita. As menções anteriores a 1850 têm referência ambígua: termos como Kradahô, Gradaú, Karaxa'u, dentre outros, foram em algum momento usados para designar grupos Mëbêngôkre, mas faltam evidências claras para associá-los aos Mëbêngôkre em tempos mais antigos. A tentativa mais bem argumentada de traçar a história dos Mëbêngôkre antes de 1850 é a de Verswijver (1992, cap. 2; cf. também TURNER, 1998). Para este autor, os Mëbêngôkre teriam atravessado o Araguaia após terem sido atacados por uma bandeira próximo à confluência com o Tocantins, entre 1810 e 1820.

O território em que os Mëbêngôkre se estabeleceram após atravessar o Araguaia teria sofrido uma forte depopulação nos séculos XVII e XVIII em razão de ataques

escravagistas provenientes de Belém e das descidas às missões jesuíticas. Isto leva a supor que a expansão dos Mëbêngôkre nesse território foi rápida, e que eles teriam tido poucos obstáculos para perambular por uma vasta área delimitada pelo Tapajós ao oeste, o canal do Amazonas ao norte, o Araguaia ao leste e o rio Tapirapé ao sul. Apesar da depopulação, no entanto, na região viviam ainda povos de vários grupos linguísticos com os quais os Mëbêngôkre tiveram contato.

As fontes etnográficas apontam a um contato frequente e prolongado, porém normalmente hostil, entre os Kayapó (principalmente Mëkrâknôti) e os Yudjá (Juruna) (cf. VERSWIJVER, 1982), por um lado, e a relações amistosas entre grupos Xikrin e os Iny/Karajá (mais particularmente, os Xambioá).⁸ Ambas situações de contato deixaram sua marca na cultura e no vocabulário dos Mëbêngôkre. Outros contatos, normalmente hostis, existiram entre os Xikrin e diversos grupos Tupí-Guaraní (Araweté, Suruí-Aikewára, Parakanã e outros) e os Arara (Carib), e entre os Kayapó e os Apyãwa (Tapirapé) e os Panará. Estes últimos foram obrigados, após o contato, a viver por uma geração em aldeias Mëbêngôkre do Parque Indígena do Xingu. Algo similar ocorreu com os Tapayúna, com os quais os Mëbêngôkre não tinham contato prévio, mas que permanecem estreitamente ligados a eles hoje em dia. Mesmo contatos hostis puderam deixar marcas na língua Mëbêngôkre, pois é frequentemente afirmado nas fontes etnográficas que mulheres raptadas de outros grupos eram fonte de informação sobre cantos e cultura material.

* * *

Além destes contatos, houve uma coabitação forçada de crianças e jovens Kayapó Irã'ãm'rânhre (e de uns pouquíssimos Gorotire) com crianças de outras etnias (Iny/Karajá, Apyãwa, Xavante, Akwê-Xerente, e talvez outras) nos internatos fundados na segunda metade do século XIX no Araguaia pelo governo goiano de Couto de Magalhães (o Colégio Isabel) e pelo religioso dominicano Gil Vilanova naquilo que logo viria a ser

⁸ Os mecanismos do contato podem ser muito complexos, e algo que apresentamos aqui de maneira simplificada como um contato direto pode ter sido em realidade intermediado. Tome-se, por exemplo, o seguinte dado sobre o contato entre os Xikrin e os Iny/Karajá, recolhido em Vidal (1977, p. 49): “O contato [dos Xikrin] com grupos não-Kayapó [i.e., não-Xikrin] parece ter sido de hostilidade. Krause (1911, RAM, LXX, p. 151), entretanto, menciona trocas de meninos entre Kayapó e Xambioá, para que aprendessem as respectivas línguas. O grupo do Kokorekre [sic], porém, manteve contatos pacíficos com os Karajá. Foi através dos remanescentes do Kokorekre que os Put-Karôt [ancestrais dos Xikrin atuais] receberam, mais tarde, a influência indireta daqueles índios do Araguaia. Os informantes dizem que muitos indivíduos desta aldeia sabiam falar Karajá. A natureza da relação colocava os Karajá, automaticamente, na categoria *ômbikwa* (parentes)”.

Conceição do Araguaia (GALLAIS, 1942). É difícil avaliar o impacto linguístico desta experiência, porém o impacto humano é claro: os Irã'ãmrãnhre foram poucos anos depois dizimados por doenças importadas, e desapareceram enquanto grupo étnico nos anos 1940.

O contato exterior mais importante para os Mëbêngôkre a partir da segunda metade do século XIX é, sem dúvida, a relação com os *kubê kryt*, hoje simplesmente *kubê*, os não-indígenas.⁹ Os vários grupos Mëbêngôkre estabeleceram contatos pacíficos com os *kubê kryt* em momentos diferentes: os Irã'ãmrãnhre na segunda metade do século XIX, os Gorotire na década de 1930, os Xikrin do Cateté e os Mëkrãknöti na década de 1950, os Kararaô na década de 1980. Em períodos anteriores ao contato pacífico com as frentes de atração, grupos Mëbêngôkre se aproximaram pacificamente em alguns casos, e em outros atacaram seringueiros e caçadores de peles à procura de armas, ferramentas, cativos e outro butim. Entretanto, os empréstimos do português em Mëbêngôkre (discutidos na seção 6) não são muito numerosos, e supomos que todos eles são relativamente recentes, i.e., posteriores ao contato pacífico permanente (mesmo assim, eles apresentam diversos graus de integração). Isto deve-se, em grande medida, à plasticidade da língua Mëbêngôkre para criar novos termos por composição.¹⁰

Tendo em vista a frequente cisão de aldeias e a grande mobilidade dos Mëbêngôkre no período imediatamente anterior ao contato com os não-indígenas, poderia se supor que os contatos com outros grupos seriam próprios a subgrupos, aldeias ou até mesmo facções. Isto é, sem dúvida, correto. O contato com os Arara (família Carib), por exemplo, é até onde sabemos exclusivo ao subgrupo Kararaô dos Kayapó e aos Xikrin. Porém, parece-nos que mesmo se o contato com um determinado grupo exterior fosse

⁹ A denotação do termo *kubê* é um tanto elástica (cf. GORDON, 2006, p. 120): a traduzimos aqui por 'não-indígena', porém antigamente ela era aplicada a qualquer grupo não-Mëbêngôkre. Isso deve ser lembrado ao considerar algumas traduções abaixo, tais como *kubê jaé* 'rede' ('ninho de *kubê*): a tradução literal não deve levar a supor que esse item foi introduzido por não-indígenas. Gordon e outros autores assinalam que o termo *kubê* se aplica também a subgrupos Mëbêngôkre (*Kubêkàkre*, *Kubêkrãkênh*), o qual sugere que a tradução de *kubê* seria mais propriamente 'externo ao grupo local'. No entanto, esse uso de *kubê*, ao nosso ver, não é como quando o termo se aplica a não-Mëbêngôkre: grupos políticos ou rituais entre os Mëbêngôkre assumem para si nomes diversos (de forma totêmica, talvez), incluindo nomes de outros grupos indígenas ou de criaturas imaginárias semi-humanas (também *kubê*). Trata-se aqui de uma *autodenominação*; os grupos assim nomeados são considerados *mëbêngôkre* pelos outros Mëbêngôkre.

¹⁰ Um(a) parecerista anônimo(a) indaga se a ideologia linguística não poderia ser responsável pela preferência dos falantes por neologismos em oposição a empréstimos do português. Carecemos de evidências que nos permitam comentar sobre essa interessante hipótese, pois é pouco o que temos ouvido dos Mëbêngôkre que diga respeito a uma ideologia linguística além de uma reverência um tanto vaga pela "fala dos antigos", reverência esta que não exclui um uso jocoso dessa modalidade de fala.

exclusivo de só uma parte dos Mëbêngôkre, a circulação de bens entre estes era tal que qualquer conceito que entrasse ao universo Mëbêngôkre por um ponto específico rapidamente se estenderia a todas ou quase todas as comunidades.¹¹ Podemos ilustrar isto de forma anedótica com duas festividades que os Mëbêngôkre identificam claramente como sendo importados de povos vizinhos: o *Bô*, baseado no Aruanã dos Iny/Karajá, foi provavelmente festejado primeiro pelos Xikrin, seus aliados, mas hoje é comum a todos os Mëbêngôkre; em sentido inverso, o *Kwÿrÿ kangô*, de origem Yudjá, teria entrado ao universo Mëbêngôkre através dos Kayapó do Xingu, que têm contato mais estreito com aqueles, mas é hoje praticado por todos os Mëbêngôkre. Há um número importante de cantos e de objetos rituais associados a essas festividades. Dessa forma, encaramos o atual estudo sem diferenciar entre variedades Mëbêngôkre além da cisão de primeira ordem entre Irã'âmrahre (SALA, 1920), Kayapó (que inclui Gorotire, Mëkrãknôti e Mëtyktire) e Xikrin, dialetos que, de fato, apresentam algumas diferenças de léxico.

4. Alguns elementos da fonologia e fonotática da língua Mëbêngôkre

O inventário de segmentos do Mëbêngôkre costuma ser dado como nos Quadros 1–2 (STOUT, THOMPSON, 1974; SALANOVA, 2001). Entre parênteses, citamos os equivalentes ortográficos de cada segmento, de acordo com o uso dos falantes da língua.

Quadro 1. As consoantes do Mëbêngôkre

	bilabiais	alveolares	palatais	velares	glotal
obstruintes surdas	p (p)	t (t)	tʃ (x)	k (k)	ʔ (ʔ)
obstruintes sonoras	b (b)	d (d)	dʒ (dj)	g (g)	
nasais	m (m)	n (n)	ɲ (nh)	ŋ (ng)	
soantes orais	w (w)	r (r)	j (j)		

Fonte: Elaborado pelos autores, 2020.

Quadro 2. As vogais do Mëbêngôkre

	anteriores	posteriores não-arredondadas	posteriores arredondadas	anteriores	posteriores não-arredondadas	posteriores arredondadas

¹¹ Vanessa Lea (UNICAMP, comunicação pessoal) informa que há casos de empréstimos recentes do português que se usam em algumas aldeias, mas não em outras. Esse é o caso, por exemplo, do termo para 'prego': enquanto em Kôkrajmôrô é usado o empréstimo *pregu*, os Mëtyktire utilizam o composto nativo *kukôj nhidjôt* (= 'pênis de macaco'). Supomos que essas discrepâncias têm a ver com as diferenças no grau de bilinguismo em português entre aldeias, e não se observam nos empréstimos mais antigos.

	orais			Nasais		
altas	i (i)	ɯ (y)	u (u)	ĩ (ĩ)	ũ (ỹ)	ũ (ũ)
médias-altas	e (ê)	ɾ (ỳ)	o (ô)	ẽ (ẽ)	ã (ã)	õ (õ)
médias-baixas	ɛ (e)	ʌ (à)	ɔ (o)			
baixas		a (a)			ã (ã)	

Fonte: Elaborado pelos autores, 2020.

A simetria aparente nestes quadros é um tanto enganosa. Por um lado, os fonemas /tʃ/ (quando ocorre em ataque silábico), /d/, /ũ/, /ũ/ e /ã/ são extremamente raros; por outro, /w j b g ŋ ʔ/ têm uma distribuição limitada. Resumimos algumas restrições que se aplicam aos temas simples.¹²

- A maioria esmagadora das ocorrências de /w/ diz respeito aos reflexos dos antigos núcleos complexos *wa, *wâ, *wỹ, *û do Proto-Jê Setentrional (= PJS), mesmo que sincronicamente /w/ seja analisado como um ataque quando ocorre diante de vogal: wakō ‘quati’, ngwỳnh ‘penugem’, kwỹrỹ ‘quebrar (singular)’, kruwa ‘flecha’ (< PJS *cwakō, *ngwâñ, *kwỹr, *krû). Segundo Nikulin e Salanova (2019), o Proto-Jê Setentrional apresentava o ataque *w (preservado em Mëbêngôkre), porém este ocorria em apenas quatro palavras de origem potencialmente não nativa: *wô ‘pati’, *wewe ‘borboleta’, *wet ‘camaleão’, *kuwêñ ‘passarinho’ > Mëbêngôkre wô, wewe, wet, kwênh. Em algumas palavras do Mëbêngôkre de origem nativa, ocorreu a inserção de /w/ a fim de resolver um hiato que emergiu em decorrência da elisão de *c (Mëbêngôkre kuwy ‘fogo’, kuwê ‘abcesso’ < PJS *kucy, *ku:cê). As demais palavras que apresentam /w/ (tais como awàrà ‘pau-de-cheiro’) são, portanto, prováveis empréstimos de outras línguas.
- /j/ ocorre com certa frequência (i) em coda, como em maj ‘caranguejo’, (ii) na sequência /je/ (reflexo do ditongo *jê do Proto-Jê Setentrional), como em mjên ‘marido’, (iii) diante de /a/ em sílaba átona, como em jaka ‘branco’ (nesse ambiente, há motivos para analisar o fonema /j/ como uma realização do morfofonema [dʒ]), bem como (iv) seguindo /m/ em coda na sílaba anterior, como em amjy (nesse ambiente, há motivos para analisá-lo como uma realização do morfofonema [ɲ]). Há também duas palavras em que /j/ foi inserido diacronicamente a fim de resolver um hiato que emergiu em decorrência da elisão de *c (Mëbêngôkre pējàti ‘xexéu’, ijô ‘tonsura’ < PJS *pēcā-ti, *cicô). De resto,

¹² Quando falamos da estrutura dos temas simples, excluímos os compostos, que são frequentes em Mëbêngôkre. Em princípio é possível identificar os compostos pela fonologia, pois estes mantêm uma acentuação secundária em todos os temas subordinados; por exemplo: rop ‘onça’ + no ‘olho’ → [ˌrɔmˈnɔ] ‘lanterna’, pari ‘pé’ + kà ‘casca’ → [ˌpatˈkʌ] (Xikrin) ~ [ˌpariˈkʌ] (Kayapó) ‘calçado’.

sabemos de apenas uma palavra de origem nativa com /j/ em Mëbêngôkre, *jàt* ‘batata-doce’ (< PJS **jət*). Logo, qualquer palavra que apresenta /j/ em um ambiente que não se encaixa na descrição acima é um provável empréstimo.

- /b/, /g/ e /ŋ/ normalmente ocorrem apenas em temas simples em início de sílaba tônica (final), com exceção de formas com reduplicação, tais como *nhōbajbat* ‘náusea’.¹³ Em outras palavras, no léxico nativo essas consoantes só ocorrem em ataque de sílaba final de morfema; palavras que não satisfazem essa restrição (por exemplo, Xikrin *abatàri*, Kayapó *abatành* ‘crescido’, *baitep* ‘não parente’) são, portanto, candidatas a serem empréstimos de uma fonte ainda desconhecida. Além disto, /g/ não ocorre em ataques complexos, ao passo que /b/ faz parte de um ataque complexo em uma única palavra, *bri* ‘rã’ (via metátese de **bir*).
- /ʔ/ ocorre apenas em início de sílaba tônica, exceto em palavras com reduplicação, e sempre como ataque simples.
- No léxico nativo, a única fonte de /tʃ/ em ataque silábico é a sequência **tj* (ex. *xêrê* ‘queimar’ < PJS **tjêr*, em que **jê* é analisado como um ditongo em PJS, e uma série de verbos anticausativos, em que o /tʃ/ advém da justaposição de um **t* do prefixo e um **j* do tema: *bixadjwÿrÿ* ‘descer’ < **bit-* + **jaĵwâr*). Portanto, a presença de /tʃ/ em posição de ataque é um forte indicador de provável origem não nativa de palavras como *pyxãnhre* ‘pau-brasil’ ou *jaxwe* ‘maluvido’.
- As vogais nasais /ũ/, /ũ/, /ã/¹⁴ ocorrem em algumas palavras, mas são raríssimos os casos em que não é possível identificar uma fonte externa para sua nasalização. Em palavras como *nhÿm* (Kayapó) ou *nhÿm* (Xikrin) ‘quem’, *nhÿnh* ‘onde’, *nhÿm* ‘turvo’, *jaÿm* ‘poeira’, *kunÿm* ‘capivara’, *tÿm* ‘antigo’, *kÿm* ‘fumaça’, *djÿn* ‘xingar’, *djãm* ‘estar em pé (não finito)’, *xãn* ‘gato doméstico’, entre outras, a vogal está sempre seguida por uma consoante nasal. Na palavra *ãkmere* ‘irmão (vocativo)’, ainda, o /k/ assimila a nasalidade da consoante seguinte, servindo de fonte provável para a nasalidade do /ã/: [ãŋ'mere]. Desde uma perspectiva diacrônica, /ũ/ é a única dessas três vogais que certamente existiu em Proto-Jê Setentrional, em Proto-Jê e mesmo em Proto-Macro-Jê, embora sempre com uma frequência limitada. Em Mëbêngôkre, ocorre em palavras como *nhÿ* ‘sentar’, *kwÿrÿ* ‘quebrar’, *nhÿ* ‘casco dos animais’, *ÿ* ‘sem nada’. Inversamente, quase todas as instâncias de /ã/ e /ũ/ em Mëbêngôkre são produto de dois processos recentes de nasalização, em que as vogais orais **a* e **u* historicamente adquiriram a nasalidade a partir de consoantes nasais adjacentes. O primeiro desses processos ocorreu em Proto-Jê de Goyaz e nasalizou as vogais **a*, **u*, **y* diante da coda *-*m*

¹³ Uma série de temas têm consoantes vozeadas devido a prefixos como *bi-* (anticausativo, não finito) e *g-* (um alomorfe do prefixo /a-/, que deriva termos triádicos de parentesco; SALANOVA, LEA, em prep.).

¹⁴ Note-se que na ortografia oficial, /ã/ e /ĩ/ são representados pelo mesmo grafema ⟨ã⟩. Neste parágrafo, a transcrição ortográfica representa sempre palavras com /ã/, enquanto que no resto do trabalho o grafema *ã* representa, na maioria das vezes, /ĩ/.

(NIKULIN, 2020, p. 88). O segundo operou apenas em Mëbêngôkre e nasalizou a vogal *a em algumas palavras seguindo um ataque nasal (como em *nhã* e *mjã* ‘morder’, *mrã* ‘caminhar’) ou espontaneamente, como em *nhÿrÿkwã* (Kayapó) ou *nhũnkwã* (Xikrin) ‘casa’, mas sempre em ambientes em que há outra fonte de nasalidade possível. Se desconsiderarmos esses ambientes e algumas interjeições (*mã* ‘vá!’) e onomatopéias (*’ã’ãre* ‘gaivota’), a ocorrência das vogais /ũ/ e /ã/ é indicadora de origem não nativa das palavras que as contêm.

- Por último, /d/ é o segmento mais raro do Mëbêngôkre, aparecendo só em empréstimos claros do português e nos seguintes itens: *dyjdyj* ou *dujduj* ‘ave sp.’, *jaduj* ‘curto’, *krwÿdy* ‘bico’. Neste último, que descende de PJS **krwât*, o /d/ alterna com /t/ em alguns compostos e o /u/ final se perde: *krwÿt kamrêk* ‘espécie de mutum’, plausivelmente derivado de ‘bico vermelho’. Etimologicamente, a palavra *màtkrwÿ’y* ‘feijão’ (< PJS **mban-krwât*, literalmente ‘grão de bico de arara’) contém essa mesma raiz, mas com queda da consoante e vogal final.

A estrutura silábica máxima em Mëbêngôkre é CCCVC. Os ataques estão sujeitos a uma série de restrições:

- os ataques respeitam uma escala de sonoridade em que as nasais e as obstruintes (oclusivas e africadas) vêm primeiro, o /r/ vem segundo, e as semivogais /w/ e /j/ vêm terceiro (etimologicamente, o /w/ nesses ataques sempre faz parte de um antigo ditongo; a vogal que o segue sincronicamente é quase sempre /a/ ou /ɜ/);
- consoantes que compartilham um articulador jamais coocorrem em um mesmo ataque; isto elimina ataques do tipo /pw/ (por conter duas consoantes labiais), /tr/ (por conter duas consoantes coronais), e assim sucessivamente;
- os ataques CCC são um caso especial da aplicação das duas restrições anteriores: são permitidos apenas os ataques /krw/ e /ŋrw/, únicos que respeitam ambas;
- as sílabas sem ataque são raras; aquelas que não decorrem da queda de *c estão limitadas à posição inicial de temas e sempre têm por núcleo as vogais /a/ ou /i/; palavras iniciadas com consoantes palatais ou com /pu/, /pu/ podem perder a consoante inicial quando flexionadas para a terceira pessoa:¹⁵ *pumũ* /pumũ/ ‘ver’ → *omũ* /ɔmũ/ ‘vê-lo’, *jano* /janɔ/ ‘enviar’ → *ano* /anɔ/ ‘enviá-lo’.

Quanto às codas, embora as consoantes /p/, /t/, /tʃ/, /k/, /m/, /n/, /ŋ/, /w/, /r/, /j/ possam ocorrer nessa posição, há uma série de restrições:

¹⁵ Algumas das palavras iniciadas com [p] sofrem ajustes na qualidade da vogal quando a consoante inicial cai; isto pode verse em [ɔmũ]. Para mais detalhes, ver Salanova (2011).

- as codas em sílabas átonas são pouco frequentes (exceto em casos de ressilabificação decorrente da síncope de uma vogal alta, como em *akno* /aknɔ/ ‘perder-se’, uma derivação anticausativa de *kuno* /kunɔ/ ‘perseguir’);
- o contraste entre /p/ e /m/, /t/ e /n/, e /tʃ/ e /ɲ/ em coda tem baixíssima carga funcional, e a escolha entre uma ou outra consoante pode na maioria dos casos ser adivinhada a partir da qualidade da vogal que antecede;
- as análises anteriores têm postulado /w/ e /j/ como codas, especialmente se precedidas por vogais altas homorgânicas (em umas poucas palavras, ocorreria /ow/ e /ɛj/), porém é possível analisar os respectivos fones como glides de transição que fazem parte dos ditongos decrescentes /uɔ/ [uwa] <uwa>, /iɔ/ [ija] <ija>, /ɛɔ/ [ɛje] <eje>, /oɔ/ [owa] <ôwa>, /ĩɔ/ [ĩã] <ĩã>;¹⁶ o <j> é empregado ainda para representar uma coda palatal em sílaba átona, contexto em que outras consoantes palatais não são representadas;
- ao ocorrer na posição de coda, o rótico /r/ é obrigatoriamente seguido na superfície por uma vogal eco átona, na maioria das vezes uma cópia da vogal que precede (ou, mais raramente, [i], mas só quando a vogal precedente é /a/, /ɔ/ ou /ʌ/).

Foneticamente, as codas costumam assimilar o modo de articulação da consoante que segue (/akno/ > [aŋ' nɔ] ‘perder-se’). Em Xikrin, o /r/ se neutraliza com /t/ diante de uma consoante surda: /par/ + /kʌ/ → [pat' kʌ];¹⁷ há uma certa variação dialetal entre codas /r/ e /j/ em vários lexemas: Xikrin /arne/ ‘satisfeito’, /abatʌr/ ‘crescido’ (pronunciados [ari' nɛ], [aba' tʌri]) ~ Kayapó /ajne/, /abatʌj/.

Quanto ao número de sílabas dos temas em Mëbêngôkre, os trissílabos são raros, e os dissílabos apresentam uma série de restrições. De fato, várias das restrições que dizem respeito a sílabas átonas acima poderiam ser reformuladas em termos daquilo que Nikulin e Salanova (2019) chamam de *estrutura sesquissilábica* dos temas:

- em um sesquissílabo, as sílabas pré-tônicas são escolhidas de uma lista restrita, que inclui /ka-/ , /ku-/ , /ja-/ , /pu-/ ~ /pu-/ (com harmonia vocálica de arredondamento), /dʒu-/ , /dʒʌ-/ , /ɲi-/ , /bi-/ , /aj-/ , /a-/ , /wa-/ , /i-/ , /e-/ , /to-/ , ou a reduplicação da sílaba tônica. Estas sílabas teriam origem em prefixos de mudança de valência (ver SALANOVA, 2011), e algumas delas são exclusivas dos verbos;

¹⁶ Nos materiais produzidos por missionários do SIL, esses ditongos são grafados <uw>, <ij>, <ej>, <ôw>, <ij>.

¹⁷ É possível que o contraste se mantenha como um contraste entre uma [ɽ] dental (</r/) e uma [t] alveolar ou pós-alveolar. Nunca realizamos um estudo detalhado desta questão.

- os temas simples são monossilábicos ou sesquissilábicos; em particular, todos os verbos transitivos compatíveis com a série acusativa dos índices de pessoa são monossilábicos, e quase todos os verbos transitivos que recebem a série absoluta dos índices de pessoa são sesquissilábicos.

Finalmente, o acento em Mëbêngôkre recai na última sílaba, com raríssimas exceções. Por um lado, o processo de epêntese vocálica mencionado acima produz um número limitado de sílabas pós-tônicas quando um vocábulo termina com /t/ subjacente. Este processo é obrigatório. Stout e Thomson (1974) afirmam que a epêntese de uma vogal pode ocorrer opcionalmente após outras consoantes finais. Este processo, que é amplamente atestado em Apinajé (OLIVEIRA, 2005, p. 78–9) e é obrigatório em posição final de frase em Kĩsêdjê (NONATO, 2014, p. 128–30) e Tapayúna (CAMARGO, 2010, p. 100–1), é para nós inteiramente desconhecido em Mëbêngôkre. Finalmente, nas palavras que terminam (em nossa análise) com ditongos decrescentes, o acento também obrigatoriamente recai, na superfície, na penúltima vogal: *ngija* ['ŋija] ‘mucura’, *abeje* [a'beje] ‘procurá-lo’, *ngrôwa* ['ŋrowa] ‘buriti’, *kruwa* ['kruwa] ‘flecha’.

Palavras fonológicas com acento não final podem resultar, ainda, do acréscimo de clíticos átonos. Entre estes, são poucos os que não mudam a função oracional da expressão: posposições e partículas predicativas enclíticas transformam um substantivo em sintagma posposicional ou em um predicado, por exemplo, e sua composição com o nome que tomam como complemento é perfeitamente transparente. Outros enclíticos, tais como *bit* ‘só’, *djwỳ* ‘também’ ou *'õ* ‘algum’, não mudam as propriedades sintáticas de um sintagma nominal, mas a modificação que fazem do substantivo não diz respeito a seu sentido e sim a sua função dentro da frase. Seu modo de composição é igualmente plenamente transparente, e eles não seriam em nenhum caso considerados como parte integrante de um lexema pelos falantes. Isto nos deixa com um único elemento final átono que pode ter um certo uso na criação lexical: o diminutivo =*re*, encontrado também com frequência junto ao aumentativo (tônico) *-ti* para produzir um sentido de coletividade, como em *Më tyk-ti=re*, *Goro-ti=re*, e outros nomes de grupos (Mëbêngôkre ou outros), em contextos em que outras línguas Jê Setentrionais comumente empregam um reflexo do enclítico **jê*, que, aliás, deu nome a essa família linguística.

É pouco o que precisamos dizer sobre a morfologia para o nosso propósito atual. O fato mais importante é que a maioria dos empréstimos não entram na língua como palavras plenamente flexionáveis. Os verbos transitivos emprestados, por exemplo,

normalmente não recebem flexão de objeto como acontece com os verbos nativos; ao invés disso, introduzem seus objetos por meio de uma posposição: *o prita < fritar*, em que o objeto é introduzido por *o* ‘com’, ou *(ku)mã pra < comprar*, em que a sílaba inicial do vocábulo português foi analisada como um índice de terceira pessoa seguido da posposição *mã* (comumente abreviada em *m*). Existem algumas exceções ao anterior (*grava*), porém não conhecemos verbos emprestados que tenham formas finitas e não finitas distintas. Algo similar ocorre com os nomes: quase todos os substantivos emprestados são alienáveis, e em consequência o possuidor não se expressa neles por meio de flexão de pessoa e sim por meio de um sintagma adposicional.

5. Contato com outras línguas indígenas

Nas seções 5.1–5, apresentamos algumas evidências linguísticas (quase exclusivamente de cunho lexical) que apontam a situações de contato que envolveram, por um lado, o Mëbêngôkre e, por outro lado, outras línguas indígenas das famílias Iny/Karajá (5.1), Tupí-Guaraní (5.2), Jê (5.3) e Jurúna (5.4), além de prováveis empréstimos de fonte desconhecida (5.5).

5.1 Iny/Karajá

O contato interétnico que resultou no maior número de empréstimos lexicais em Mëbêngôkre foi, sem dúvida, aquele entre os Xikrin e o subgrupo Xambioá do povo Iny, mais conhecido em português como *Karajá* e em Xikrin como *Ngôjrê* ‘canoeiros’ (VIDAL, 1977, p. 49 transcreve *Ngoire*); o termo é aplicado também pelos Kayapó aos Yudjá. A primeira menção do contato entre as línguas desses povos da qual temos conhecimento encontra-se no trabalho de Krause (1911, p. 388), que afirma que a designação do cachimbo em Mëbêngôkre (Xikrin *watkoko*, Kayapó *warikoko* /warkɔkɔ/; *wălikôkô* na transcrição de Krause) provém da língua Iny.¹⁸ Da mesma forma, Sala (1920, p. 404) observa, em relação ao léxico Mëbêngôkre:

¹⁸ Neste trabalho optamos por usar o termo *Iny* para nos referirmos ao conjunto das variedades linguísticas conhecidos na literatura como Karajá do Norte, Karajá do Sul, Javaé e Xambioá (cf. RIBEIRO, 2001/2002, 2012). O termo *Karajá* será utilizado para denotar as primeiras duas dessas variedades, em oposição ao Javaé e ao Xambioá.

“Encontram-se poucos nomes estrangeiros. Da lingua Carajá só encontramos *warukoko** cachimbo, *ibim*, mau.

*A respeito desse nome nao se póde saber bem se é de origem Carajá ou Kaiapó”.

Se, por um lado, não pudemos identificar a palavra do Mëbêngôkre transcrita por Sala (1920) como “*ibim*” (e nem seu possível étimo Iny), a origem Iny de *warikoko/watkoko* é confirmada por Gordon (2006, p. 121) e Ribeiro (2012, p. 13, 148), dentre outros autores. O equivalente desse termo é amplamente atestado na literatura sobre os Iny como *weri(k)ò(k)ò* /wɛɾɪ(k)ɔ(k)ɔ/ (cf. RODRIGUES, 2008, p. 55; KARAJÁ *et al.*, 2013, dentre outros); no generoleto masculino, o primeiro (e, opcionalmente, o segundo) *k* pode ser elidido (ver RIBEIRO, 2012, p. 130–55 sobre a elisão de /k/ na fala dos homens Iny).¹⁹

Em trabalhos mais recentes, podem ser encontradas maiores indicações de empréstimos linguísticos de origem Iny em Mëbêngôkre. Por exemplo, Toral (1992, p. 21, citando comunicação pessoal com Terence Turner) e Gordon (2006, p. 121) mencionam, além de *warikoko*, os termos *rorirori* /ɾɔɾɪɾɔɾɪ/ ‘tipo de cocar, coifa plumária’ e *warabaê* /warabae/ ‘cesto estojiforme’, ambos claramente emprestados da língua Iny. O étimo do primeiro é *lòrilòri* /lɔɾɪlɔɾɪ/ (KARAJÁ *et al.*, 2013; TAPIRAPÉ, 2019; NUNES, 2016, p. 577), com um significado idêntico (cf. também LEA, 2012, p. 477). Já o provável étimo de *warabaê* é atestado em diversas fontes com vogais finais diferentes: *wrabahu* (KARAJÁ *et al.*, 2013), *wrabahi* (N. LIMA, 2019b, 124) ou *wrabahy* (E. R. Ribeiro, mensagem postada no fórum da Etnolinguística em 5 de janeiro de 2011).²⁰ Nesta última mensagem, Ribeiro aponta ao fato de a sequência *-ba-* ser extremamente atípica da fonologia Iny, pois a vogal *a* normalmente nasaliza as oclusivas /b/ e /d/ que a precedem, resultando em [ma], [na] (por esse motivo, neste trabalho utilizamos /ã/ e não /a/ nas representações fonológicas da língua Iny). Portanto, é possível especular que tanto Xikrin *warabaê* como Karajá *wrabahu* (-i, -y) (e suas variantes dialetais em Javaé e

¹⁹ Em algumas fontes, encontramos formas com uma vogal diferente na primeira sílaba: *wèri(k)ò(k)ò* /wɛɾɪ(k)ɔ(k)ɔ/ (LOURENÇO, 2009, p. 194, 355; RIBEIRO, 2012, p. 148) ou *wari(k)ò(k)ò* /wãɾɪ(k)ɔ(k)ɔ/ (RIBEIRO, 2012, p. 13). Desconhecemos se se trata de variação dialetal.

²⁰ Cópia arquivada em <<http://listserv.linguistlist.org/pipermail/etnolinguistica/2011-January/005223.html>>.

Xambioá, com *wara...* substituindo *wra...*) sejam empréstimos de uma terceira língua, até o presente não identificada.²¹

Diversos outros empréstimos na mesma direção (isto é, do Iny para o Mëbêngôkre) foram identificados por Ribeiro (2012, p. 13), incluindo *bikwa* /bikwa/ ‘parente’, *bero* /berɔ/ ‘puba’, *rara* /rara/ ‘cesto bolsiforme’ e *wiwi* /wiwi/ ‘canto’, emprestados, respectivamente, de *bi(k)òwa* /bɪ(k)ɔwã/, *bèrò* /berɔ/, *lala* /lãlã/, *wii* /wi/, com significados idênticos (o último item ocorre apenas na fala masculina nos dialetos Xambioá e Karajá do Norte; em Karajá do Sul e Javaé, o generoleto masculino possui a forma *wiu* /wiu/, já no generoleto feminino de todos os dialetos ocorre *wiku* /wiku/). Ainda segundo Ribeiro (2012, p. 13–4), há um ou dois prováveis empréstimos em direção oposta, isto é, do Mëbêngôkre para o Iny: o nome pessoal *Krumarè* /kru'bãre/ (< Mëbêngôkre *Krômare* /krɔma-rɛ/) e o termo para ‘bacaba’, *kabiri* /kabiri/ (< Mëbêngôkre *kamêrè* /kameɾ/, de PJS **kambêr*). Em relação a este último par de termos, Ribeiro não descarta a possibilidade de se tratar de termos cognatos, herdados do Proto-Macro-Jê (e não de empréstimos), porém isto nos parece pouco provável: segundo a proposta reconstrutiva de Nikulin (2020), o cognato esperado de PJS **kambêr* em Iny não seria *kabiri* e sim **kabrè* (para a correspondência entre Mëbêngôkre *ê* ~ Iny *è*, compare Mëbêngôkre *mê* ‘gota’ e Iny *bèè* ‘água’, ambos de Proto-Macro-Jê **mbi₁ n^o* ‘água’). Portanto, é altamente provável que *kabiri* seja um vocábulo emprestado na língua Iny, de origem Jê Setentrional. A julgar pelo padrão de substituição vocálica, *kabiri* não pode ser oriundo do Mëbêngôkre, mas sim de uma outra variedade Jê Setentrional (talvez não mais existente) em que PJS **ê* possui o reflexo /i/ (como em Gavião-Pyhcopji e em Krĩkatĩ).

A seguir, apresentamos quatro etimologias adicionais (até onde sabemos, inéditas) que envolvem prováveis empréstimos do Iny (mais especificamente, do dialeto Xamboá) em Mëbêngôkre (principalmente no dialeto Xikrin). As palavras etimologizadas são: *waxi* ‘linha’, *benorã* ‘tucunaré’, *awo* ‘tipo de árvore; ubá’, *ixe(re)* ‘espelho’.

Xikrin *waxi* /watʃi/ ‘linha’ < Xambioá *waxi* /wãθi/ ‘anzol’ (atestado em Karajá e Javaé por OLIVEIRA, 2009, p. 97, n. 2; SILVA, 2010, p. 34; KARAJÁ *et al.*, 2013; NUNES, 2016, p. 185, nota 58; N. LIMA, 2019a, p. 39; JAVAÉ, RODRIGUES, 2020, p.

²¹ Segundo Gordon (*loc. cit.*, nota 16), os Xikrin informam que o item foi apropriado dos *kubê kamrêk* (grupos de língua Tupí-Guaraní do baixo interflúvio Tocantins-Xingu).

29). Este empréstimo é restrito ao dialeto Xikrin, pois os Kayapó utilizam um termo de origem nativa, *êdjà* /edʒʌ/ (analisável como Ø-ê-djà 3-amarrar-INSTR). A direcionalidade do empréstimo proposta é confirmada não apenas pelo fato de o termo Xikrin carecer de cognatos óbvios em outras línguas Jê, mas também pela ocorrência da consoante *x* /tʃ/ em um ambiente extremamente atípico para o Mëbêngôkre (em ataque silábico precedendo uma vogal diferente de /e/, ver seção 4). Quanto à aparente discrepância semântica, note que a linha e o anzol são conceitualizados como uma unidade em Xikrin; para especificar o anzol propriamente dito, utiliza-se o termo *waxidjwa* (← *waxi* ‘linha’ + *djwa* ‘dente; objeto pontudo’). Um mecanismo semelhante vê-se em *po* (propriamente, ‘taquara’) e *podjwa*, ambos significando ‘flecha’. Além disso, muito embora o termo *waxi* da língua Iny seja traduzido como ‘anzol’ em todas as fontes por nós consultadas, não é impensável que o mesmo tenha apresentado a semântica de ‘linha’ num momento anterior. Um provável derivado de *waxi* que nos faz pensar nessa possibilidade é *waxiwahatè* /wãθiwãhãðé/ ‘arco’ (KARAJÁ *et al.*, 2013, dentre outros), cujo significado se aproxima semanticamente de ‘linha’ (em Mëbêngôkre, por exemplo, *djudjê* ‘arco’ é derivado de *djê* ‘amarra’). Entretanto, ainda não pudemos etimologizar o elemento *-wahatè* /-wãhãðé/.

Xikrin *benorã* /bɛnɔrɔ/ ‘tucunaré’ < Xambioá *bènora* /bɛdõrã/ ‘tucunaré’ (R. P. de OLIVEIRA, 2018, p. 31; atestado em Karajá por NUNES, 2016, p. 220; KARAJÁ, 2019, p. 45).²² Este empréstimo é restrito ao dialeto Xikrin, pois os Kayapó utilizam um termo de origem nativa, *tepi kôt* /tepi kɔt/ (de *tep* /tep/ ‘peixe’ e *ikôt* /ikɔt/ ‘alça’). Novamente, a direcionalidade do empréstimo não levanta dúvidas: Xikrin *benorã* não possui cognatos conhecidos em outras línguas Jê e nem pode ser decomposto em morfemas de forma plausível;²³ além disso, tanto a trissilabidade desse item como a ocorrência de /b/ em uma sílaba não final de um morfema são características indicadoras de origem alógena da palavra. Em contraste, Xambioá *bènora* não viola nenhuma

²² C. P. de Oliveira (2009, p. 97, n. 2) fornece a tradução ‘pirarucu’, porém trata-se de um aparente equívoco. Segundo Nunes (2016, p. 31, 131, 402, 430, 516, nota 11) e Karajá *et al.* (2013), as palavras para ‘pirarucu’ em Karajá são: *bdòlèkè* (♀), *bdòlèè* (♂) (dialeto do Norte); *bdòkujkè* (dialeto do Sul); *watxibirè* (palavra antiga); em uma narrativa, ocorreu ainda *bodolèkè* (NUNES, 2016, p. 521).

²³ É verdade que *no rã* significa ‘olho murcho’, e que o tucunaré se caracteriza por ter um desenho na cauda que lembra um olho. Porém *be-* não tem significado independente, e a conexão semântica é um tanto remota.

restrição fonotática da língua Iny e ocorre pelo menos em dois dialetos (Xambioá e Karajá), tornando plausível a hipótese de que *bènora* seja um vocábulo nativo em Iny.²⁴

Xikrin *awo* /awɔ/ ‘tipo de árvore’, Irã’ãmrãnhre <auó> ‘barco de cortiça, ubá’ (SALA, 1920, p. 410) < **Xambioá *awò* ‘canoas.♂’** (atestado no generoleto feminino como *awòkò* por RIBEIRO, 2001/2002, p. 96, nota 25; em outras variedades a sílaba inicial é *hã-*: Karajá *hãw(k)ò* /hãw(ək)ɔ/, Javaé *hãwò* /hãwɔ/; OLIVEIRA, 2010, p. 113; RIBEIRO, 2012, p. 157; KARAJÁ *et al.*, 2013; NUNES, 2016, p. 555). Em Xikrin, *awo* denota uma espécie de árvore que não pudemos identificar; segundo Ikrô Kayapó (comunicação pessoal), ancestralmente era utilizado para a confecção de canoas. É natural supor que o significado atestado por Sala (1920) para o dialeto Irã’ãmrãnhre (‘barco de cortiça, ubá’) seja o original; dessa forma, *awo* teria sofrido uma mudança semântica metonímica em Xikrin. A direcionalidade do empréstimo neste caso não pode ser contestada pelos seguintes motivos: (i) a palavra Mëbêngôkre contém /w/ em um ambiente atípico (ver seção 4); (ii) a palavra Mëbêngôkre carece de cognatos conhecidos em outras línguas Jê; (iii) a palavra Xambioá no generoleto feminino apresenta uma oclusiva velar /k/, cuja ocorrência seria inexplicável caso se assumisse uma direção inversa de empréstimo para esta etimologia; (iv) a palavra Xambioá possui cognatos em todos os dialetos do Iny (protoforma **hãwko*); (v) a confecção e o uso de canoas são, em geral, característicos dos Iny, mas não dos povos Jê Setentrionais.

Kayapó *ixe* /itʃɛ/, Xikrin *ixere* /itʃɛ/ ou /itʃɛ=rɛ/ ‘espelho’ < Xambioá *itxèrèna* /itʃɛrɛdã/ ‘espelho’ (atestado em Karajá por RIBEIRO, 2012, p. 254). Esta é a mais controversa das quatro etimologias Mëbêngôkre-Xambioá propostas nesta subseção, pois a palavra Mëbêngôkre não apresenta nenhuma correspondência para a sílaba final do vocábulo Iny (e, no caso do dialeto Kayapó, para as duas sílabas finais). De forma provisória, propomos que Xambioá *itxèrèna* foi interpretado pelos falantes do Xikrin como a oração em 1 (DIM = diminutivo, PRED = predicativo).

²⁴ Não descartamos a possibilidade de se tratar de um derivado opaco de *bèdò* /bɛdɔ/ ‘filhote (*peixe*)’ (NUNES, 2016, p. 555). Muito embora as alternâncias entre vogais orais e nasais (tais como /ɔ/ e /õ/) não fazem parte da (morfo)fonologia Iny, é interessante observar que tanto /ɔ/ como /õ/ podem continuar a vogal **ũ* do Proto-Macro-Jê, sem que o condicionamento da cisão seja conhecido (NIKULIN, 2020, p. 154). Portanto, tanto *bèdò* ‘filhote’ como *bènora* ‘tucunaré’ poderiam, em princípio, continuar Proto-Macro-Jê **mbi(C)nũ(C) ~ *mbi(C)ñũ(C)* ‘peixe’ (*op. cit.*, p. 380; compare Rikbáktsa *piknu* ‘peixe’, Djeoromitxí *mĩnõ* ‘peixe’, Arikapú *mĩnũ* ‘peixe’); o desenvolvimento diferenciado de Proto-Macro-Jê **ũ* poderia estar relacionado de alguma forma com a presença do elemento *-ra* em *bènora*.

- (1) *ixe=re na*
espelho=DIM PRED
'É um espelho.'

Note que o elemento grafado como *-re* em Xikrin pode ser analisado tanto como uma coda rótica /r/ seguida de uma vogal eco (ver seção 4) ou como um clítico diminutivo átono /=rɛ/ (se esta análise for adotada, a ocorrência do clítico deverá ser considerada lexicalizada nessa variedade, pois o tema não derivado **ixe* não existe em Xikrin, à semelhança do que ocorre em português em palavras como *mindinho*, que carecem de uma contraparte não derivada). Em contraste, no dialeto Kayapó a reanálise alcançou tal ponto que a ocorrência de *-re* /=rɛ/ não é mais obrigatória nem mesmo comum.

A hipótese de que o termo *ixe(re)* do Mëbêngôkre é de origem não nativa é reforçada pelo fato de não haver cognatos conhecidos dessa palavra em outras línguas Jê, bem como pela ocorrência da africada *x* /tʃ/ em um ambiente atípico (em ataque silábico diante de uma vogal que não seja *ê* /e/). Já em Iny *itxèrèna* é um termo facilmente analisável, derivado do verbo *-tèhè* /-dʒhɛ/ por meio do acréscimo de três morfemas: o prefixo nominalizador *i-* /i-/ (que desencadeia a palatalização de /d/ para /tʃ/), o infixos nominalizador *-r-* /-r-/ (que apaga o /h/ precedente) e o sufixo *-na* /-dã/ de *nomina instrumenti* ou *nomina loci* (RIBEIRO, 2012, p. 118, 211–2). Portanto, Xambioá *itxèrèna* é um termo nativo, ao passo que Mëbêngôkre *ixe* é um provável empréstimo.

5.2 Tupí-Guaraní

Exceto pelos Xikrin do Bacajá e os Kayapó Kararaô, cuja presença ali é relativamente recente, e de um grupo Carib, os Arara, a região entre o baixo curso dos rios Tocantins e Xingu é um território ocupado por vários grupos Tupí. Esta população se compõe de grupos de fala Tupí-Guaraní, hoje pequenos, mas sem dúvida mais numerosos outrora, e, no oeste desse território, de grupos Tupí de outras famílias: os Yudjá (Jurúna), canoieiros do Rio Xingu e, já nos afluentes ocidentais do Xingu, os Xipáya e Kuruáya. Segundo as duas principais propostas para as migrações deste grupo etnolinguístico (cf. MÉTRAUX, 1927; BROCHADO, 1984), a ocupação Tupí-Guaraní na região não seria

de grande antiguidade,²⁵ mas já é atestada nas primeiras crônicas europeias, no século XVII.

Os etnônimos registrados para os grupos Tupí-Guaraní desta região são muitos, e sua denotação ambígua. Ao lado de Pacajá e Tacunyapé, mencionados já no século XVII e hoje desaparecidos, encontram-se em tempos mais recentes pelo menos os seguintes: Parakanã, Kupê-rop ou Cupelobos, Amanayé, Asuriní, Anambé, Suruí-Aikewára e Araweté. Também estavam presentes nesta região no século XVII os Wajãpí, que em tempos posteriores atravessaram o Rio Amazonas e estabeleceram-se no Amapá (GALLOIS, 1980, pp. 55–9 *apud* VIVEIROS DE CASTRO, 1986, p. 139). Os Kupê-rop, mencionados nos séculos XIX e XX (cf. NIMUENDAJÚ, 2017), não correspondem a nenhum grupo identificado na atualidade. O etnônimo Asuriní se refere a dois grupos diferentes que falam variedades linguísticas relativamente distantes.²⁶ A filiação linguística de vários grupos desaparecidos ou assimilados a outros não é inteiramente clara.²⁷ Mais ao sul do lado do Araguaia, e separados de todos os anteriores pelo território Mëbêngôkre, encontram-se os Apyãwa (Tapirapé). Finalmente há um grupo Tupí-Guaraní, os Kamayurá, na região dos formadores do Rio Xingu (o mal-chamado “Alto Xingu”). O contato entre os Mëbêngôkre e estes, assim como com dois outros grupos Tupí da Terra Indígena do Xingu (Awetí e Kawaiweté) data, até onde podemos saber, do tempo posterior ao contato, motivo pelo qual não falaremos deles aqui.²⁸

Em qualquer caso, segundo as classificações mais estabelecidas (ver síntese em MICHAEL *et al.*, 2015), os grupos Tupí-Guaraní desta região representariam três ramos diferentes da família. Por um lado, os Parakanã, os Suruí-Aikewára e os Asuriní do Tocantins formariam um agrupamento cuja língua é relativamente próxima do Apyãwa,

²⁵ Com matizes, pois para Métraux os Tupí-Guaraní dessa região seriam a última avançada de uma expansão pela costa em direção ao norte a partir de São Paulo, enquanto que para Brochado o ponto de dispersão dos Tupí da costa seria o baixo Madeira, e a região de que tratamos aqui seria um dos primeiros pontos ocupados durante sua expansão. Em qualquer caso, a diferença entre as datações arqueológicas em pontos extremos da costa é pequena (FAUSTO, 1998, p. 382), permitindo-nos afirmar que a expansão Tupí-Guaraní ocorreu em um período extremamente curto, há pouco mais de 1000 anos.

²⁶ Para uma discussão detalhada do que se sabe sobre os grupos Tupí-Guaraní desta região, ver Viveiros de Castro (1986, p. 136–41).

²⁷ Um caso excepcional por ter sido noticiado, mas provavelmente representativo da realidade brutal do contato na região, é o de dois indivíduos sobreviventes de ataques dos colonos que foram contactados em proximidades de Marabá na década de 1980, aos quais foram dados os nomes de Aurê e Aurá, falantes de uma língua Tupí-Guaraní não atestada em outras fontes (MELLO, 1996).

²⁸ O contato dos Kîsêdjê com os povos alto-xinguanos é, se bem tradicionalmente hostil, de mais longa data, e resultou em uma série de empréstimos. Lembramos que a interação dos Mëbêngôkre Mëtyktire com os Kîsêdjê e os Tapayúna tem sido intensa após o contato, e poderia ser o veículo de transferências indiretas de elementos alto-xinguanos.

falado mais ao sul. Por outro, o Araweté formaria um agrupamento com o Anambé e o Amanayé (o Asuriní do Xingu também é próximo dessas línguas). Mais distantes do ponto de vista linguístico (e faladas não propriamente nesta região e sim em regiões adjacentes) são o Ka'apor e o Guajá, além da língua dos desaparecidos Tacunyapé, bem como o Wajãpí, cujos falantes, como dissemos, deixaram a região para se estabelecer mais ao norte. É sabido também que alguns grupos Tupinambá vindos de regiões costeiras mais ao leste se estabeleceram na região, fugindo da opressão colonial (FAUSTO, 1998, p. 382; GRENAND, 1982, p. 150–3).

As missões jesuíticas, partindo de Belém, estabeleceram-se a partir da primeira metade do século XVII em vários pontos da margem direita do Amazonas e no baixo curso dos rios Tocantins, Xingu e Tapajós e ao longo do Madeira. Nestas missões — às quais indivíduos de grupos étnicos diversos eram levados à força ou iam em busca de proteção das expedições escravagistas — a língua empregada teria sido, desde o início, a Língua Geral Amazônica (LGA), uma variedade linguística que se formou a partir do Tupí da costa entre a população mestiça do Maranhão e do Pará nos primeiros tempos da colônia. A LGA continuou a ser falada na região após a expulsão dos jesuítas e o desmantelamento de suas missões na segunda metade do século XVIII, e teria sido a língua da população ribeirinha da região até o influxo massivo de imigrantes de outras regiões do Brasil no final do século XIX. Conhecida hoje pelo nome de Nheengatú, ela é falada na atualidade apenas no Alto Rio Negro.

A Língua Geral Paulista fora falada no que é hoje a região Centro-Oeste até pelo menos o século XVIII. Embora seus falantes não chegaram a se estabelecer na região de que tratamos aqui, houve expedições militares levadas a cabo por bandeirantes paulistas, falantes da Língua Geral, na região da Ilha do Bananal (PALACIN, 1972, p. 16–7 *apud* RIBEIRO, 2001/2002). Ribeiro (*op. cit.*, p. 93–4, nota 15) fala de um ataque a traição à principal aldeia Iny por uma destas expedições em meados do século XVIII, relatado por Fonseca (1868 [1775]), e da posterior participação como intérprete, quando do primeiro contato pacífico com o grupo vinte anos depois, de uma cativa Iny dessa expedição.

Quanto às relações entre os Mëbêngôkre e os grupos falantes de línguas Tupí-Guaraní, a evidência histórica aponta para contatos sempre hostis entre os Kayapó Mëkrãknöti e os Apyãwa por um lado, e por outro entre os Xikrin e vários grupos Tupí-Guaraní do baixo interflúvio, incluindo pelo menos os Suruí-Aikewára (*Mydjêtire* em

Xikrin), os Araweté, os Parakanã (*Akàkakôre* para os Xikrin), os Asuriní do Xingú (*Kubê kamrêkti* ou *Krãjakàrà*, sendo este último termo usado pelos Kayapó Mêtÿktire para os Panará), os Apyãwa e os Asuriní do Tocantins.²⁹ Isto é, haveria representantes dos chamados sub-ramos IV e V da família Tupí-Guaraní. As notícias que se têm de contatos dos Mëbêngôkre com ribeirinhos nesta região são de contatos hostis na época em que se intensifica a exploração da borracha, isto é, já em tempos em que a língua empregada seria o português. Porém, é bastante provável, tendo em vista as notícias de supostos ataques dos Mëbêngôkre na região do médio Xingu no século XVIII (cf. TURNER, 1998, p. 314), que tenha havido contatos com ribeirinhos em tempos em que estes ainda eram falantes da LGA. No sudeste do território Mëbêngôkre os contatos com colonos teriam sido, ao contrário, com uma frente pastoril proveniente do Nordeste e de Goiás já no século XIX, e, portanto, falante de português desde os primeiros tempos. No entanto, devemos notar que em Xambioá existe pelo menos um empréstimo oriundo da LGA, *mabèra* ‘papel’ (RIBEIRO, 2001/2002, p. 81–2).

As evidências do contato linguístico Mëbêngôkre–Tupí-Guaraní são escassas, limitando-se a alguns poucos empréstimos lexicais, que ainda não foram identificados ou discutidos na literatura anterior da qual tenhamos conhecimento.³⁰ Abaixo apresentamos as etimologias de alguns claros empréstimos de línguas Tupí-Guaraní em Mëbêngôkre (*môtôbi’y* ‘amendoim’, *xãn* ‘gato doméstico’, *mokà* ‘mocó’, *xoko* ‘socó’, todos aparentemente oriundos da LGA), além de alguns casos mais duvidosos.

Xikrin *môtôbi[y]* /motobi[ʔu]/ ‘amendoim’ (-y [-ʔu] = ‘semente’) < Língua Geral Amazônica <mundui ~ munduby> ‘amendoim’ (STRADELLI, 1929, p. 551). Este empréstimo é restrito ao dialeto Xikrin, pois os Kayapó utilizam um termo de origem nativa, *kayre[y]* /kau=rɛ[-ʔu]/ (de PJS **kacy*). A direcionalidade do empréstimo é evidente, pois o Xikrin é a única variedade Jê Setentrional que não faz uso de um reflexo de PJS **kacy*. Em contraste, o vocábulo da LGA é de origem claramente nativa, pois reflexos de Proto-Tupí-Guaraní **mũnduβi* */munuβi/ ‘amendoim’ encontram-se atestados nas mais diversas línguas da família, incluindo Apyãwa *monowi*, Tenetehára,

²⁹ A referência dos etnônimos usados pelos Mëbêngôkre provém de Vidal (1977, p. 28, n. 43) e Müller e Silva (2018, para os *Krãjakàrà*; para Vidal, este etnônimo se aplica aos Arara).

³⁰ Sala (1920, p. 404), ao discutir os escassos empréstimos linguísticos em Mëbêngôkre, diz: “[d]a lingua guarany encontramos *atorore* especie de *jahò* (*crypturus sui*)”. Desconhecemos o étimo Guaraní pretendido por Sala (1920) para Mëbêngôkre *atoro* /atɔr/ ‘jaó’; na realidade é um reflexo regular de PJS **aʔtor* ‘jaó’.

Apiaká e Kawaiweté/Kayabí *munuwi*, Ka'apor *mundui*, Tapieté *mundúí*, Guarasugwe *munúwi* (DOBSON, 1997; GONZÁLEZ, 2008, p. 29; PAULA, 2012, p. 253; HARRISON, HARRISON, 2013; MORIMÃ, 1984 *apud* SOUSA, 2017; RAMIREZ *et al.*, 2017), dentre outras, tais como o Zo'é, o Wajãpi e o Émérillon Teko (SOUSA, 2013, p. 80).³¹ A LGA foi escolhida como a fonte mais provável do empréstimo, pois parece ser a única variedade Tupí-Guaraní em que o reflexo de **mũnduβi* possui (i) uma consoante foneticamente pós-oralizada (e não nasal) continuando **nd* e (ii) uma consoante bilabial (e não labiovelar) continuando **β*. Se o étimo de *môtôbi* [ʔy] possuísse a forma próxima a *munuwi* ou *monowi* (como é o caso nas línguas mais próximas geograficamente ao Mëbêngôkre), esperaríamos que o empréstimo fosse adaptado como **mônôwi* (ou **munuwi*), visto que o Mëbêngôkre possui em seu inventário as consoantes /n/ e /w/. Já a adaptação de <mundui ~ munduby> da LGA (possivelmente representando *mũnduβi*)³² como *môtôbi*- não é problemática: o Mëbêngôkre carece de /β/ e impõe fortes restrições à ocorrência de /d/,³³ que parece ter se fonologizado apenas recentemente na língua (ver seção 4), fazendo com que a substituição de *β* e *d* da língua-fonte por /b/ e /t/ na língua-recipientes seja compreensível. Quanto à substituição da vogal alta /u/ da LGA pela média-alta /o/ em Xikrin, Cruz (2011, p. 109–10) mostra que embora o contraste entre /o/ e /u/ tenha historicamente existido na LGA, entre os séculos XVIII e XIX o mesmo se perdeu em sílabas átonas (e estava em vias de desaparecimento em sílabas tônicas no

³¹ Mello (2000, p. 177) reconstrói **/manuβi/* (presume-se, **mãnduβi* na superfície) para o Proto-Tupí-Guaraní, porém os reflexos dessa forma (com **a*) na primeira sílaba limitam-se a algumas línguas do sub-ramo Guaraní (Mbyá, Nhandéva, Nhandéwa, Kaiowá, Guaraní Paraguaio *manduvi*; IVO, 2018, p. 97; PEREZ FELIPIM, QUEDA, 2005; CENTURIÓN SERVIN, DAVALOS ARCE, 2009) e ao Tupí Antigo (*mandubi* /*manuβi*/; NAVARRO, 2015). Uma vez que formas com /u/ são encontradas até mesmo no sub-ramo Guaraní (Tapieté), acreditamos que a reconstrução correta da forma Proto-Tupí-Guaraní seja **mũnduβi*. A mudança irregular **mũnduβi* > **mãnduβi* teria ocorrido em um período relativamente recente em algumas línguas do sub-ramo Guaraní e em algumas variedades do Tupí Antigo (não naquela que deu origem à LGA).

³² Cruz (2011, p. 94–9), ao discutir a fonologia do Nheengatú (uma variedade regional da LGA falada no Alto Rio Negro), demonstra que nessa variedade os antigos alofones pós-oralizados das nasais subjacentes precedidos de uma vogal posicionalmente nasalizada (como em **nãmbi* **/nami/* ‘orelha’, **pĩnda* **/pina/* ‘anzol’) foram reanalisados, em sílabas tônicas, como oclusivas não especificadas para o traço [±voz] precedidas de uma vogal subjacentemente nasal (/nãPi/ ou /pĩTa/); os segmentos em questão continuam sendo foneticamente vozeados, porém o vozeamento é previsível seguindo vogais nasais. Estendendo a proposta de Cruz (2011) para a variedade da LGA documentada por Stradelli (1929), seria possível derivar a forma <mundui ~ munduby> de uma representação próxima a /mũTuβi/ (neste caso, a neutralização de [±voz] e nasalidade se daria também em sílabas átonas e não apenas nas tônicas, como em Nheengatú). A correta representação do vozeamento e da nasalidade na LGA é irrelevante para o nosso propósito aqui.

³³ Na fala de alguns Mëbêngôkre, o /w/ pode ser produzido como [β] em algumas palavras, mas trata-se de um fenômeno relativamente infrequente que não registramos entre os Xikrin.

século XIX; o Nheengatú do Alto Rio Negro, tal como falado na atualidade, possui apenas /u/. Se assumirmos que o vocábulo da LGA entrou no léxico do Xikrin num momento posterior ao século XVIII, o padrão de substituição ora em discussão não é de se estranhar, pois é altamente provável que o fonema /u/ da LGA, por ser a única vogal arredondada disponível em sílabas átonas, possuía também alofones mais baixos.

Mëbêngôkre *xãn* /tʃã̃n/ ‘gato doméstico’ < Língua Geral Amazônica <pixána> ‘gato doméstico’ (STRADELLI, 1929, p. 611). É notório que o termo para ‘gato doméstico’ da LGA (registrado por Stradelli como <pixána>, fonologicamente /piʃã̃na/), proveniente por sua vez do português *bichano* (SANTOS, 2020, p. 69), acabou entrando no léxico de muitas línguas indígenas das terras baixas sul-americanas, tais como Tariána, Desáno, Wapixána, Makuxí, Pemón e Kákua (cf. EPPS, 2020). Acreditamos essa ser a origem também de *xãn*, usado em Mëbêngôkre. A supressão da vogal final átona não é muito problemática, pois ocorre ocasionalmente nos empréstimos do português (ver seção 6). É pouco provável que *xãn* provenha diretamente do português, pois a realização fonética da vogal tônica em *bichano* é consideravelmente mais próxima da vogal /ã̃/ (e não /ã/) do Mëbêngôkre. Em contraste, se assumirmos que a fonte do empréstimo tenha sido /piʃã̃na/, da LGA, a adaptação da vogal torna-se trivial. O único obstáculo para a etimologia é a ausência, em Mëbêngôkre, de qualquer correspondência para a sílaba /pi-/ da LGA. Não sabemos se o truncamento da sílaba inicial ocorreu em Mëbêngôkre ou na língua-fonte, mas de qualquer maneira trata-se de um fenômeno irregular, talvez de natureza afetiva.³⁴

Mëbêngôkre *mokà* /mɔkɔ/ ‘bolsa de tucum’ < Língua Geral Amazônica (?) #mboko /moko/ ‘tipo de bolsa’. Esta etimologia enfrenta duas dificuldades. Em primeiro lugar, o respectivo étimo não foi atestado na LGA (pelo menos, nas fontes por nós consultadas) e nem nas demais línguas da família Tupí-Guaraní faladas no interflúvio Tocantins-Xingu. A segunda dificuldade diz respeito à adaptação da vogal arredondada

³⁴ Notamos que o étimo dessa palavra em português também foi alvo de um processo irregular de truncamento afetivo, porém o resultado da aplicação desse processo sofreu uma considerável alteração semântica via metáfora: *bichana* ‘gata’ → *xana* ‘vagina’ (ORSI, 2009, p. 109). Embora seja possível que a forma *xana* tenha preservado a semântica original (i.e., ‘gata’) em alguma variedade do português em um estágio anterior, podendo servir de fonte para o empréstimo em Mëbêngôkre, tal possibilidade permanece especulativa; portanto, deixamos a questão em aberto.

de seu étimo (ver abaixo sobre as evidências que apontam a sua existência) como /ʌ/ em Mëbêngôkre.

Hipotetizamos que a LGA possuía uma forma próxima a *#mboko* ‘tipo de bolsa’ (o jogo da velha simboliza a incerteza quanto à forma exata do vocábulo). Isto pode ser inferido a partir da existência de cognatos no sub-ramo Guaraní, incluindo Mbyá e Paĩ Tavyterã *mboko* ‘bolsa feita ancestralmente de fibra de algodão silvestre’ (Celeste Escobar, comunicação pessoal) e Guaraní Boliviano *mboko* ‘tipo de bolsa’ (SAARESRANTA, 2011, p. 50),³⁵ bem como de um provável empréstimo da LGA em português, *mocó* (mais raramente, *bocó*) ‘bolsa de tiracolo feita de couro, capanga’ (cf. FERREIRA, 1986, p. 267, 1145).³⁶ Como este último termo é restrito a variedades do português brasileiro utilizadas nas regiões Norte e Nordeste, é provável que este tenha por fonte a LGA e não uma outra língua Tupí-Guaraní.

Quanto à adaptação fonológica, acreditamos que o empréstimo possui a forma *mokà* /mɔkʌ/ (e não **moko* */mɔkɔ/, como seria esperado) em decorrência de um efeito conhecido como *etimologia popular*: sincronicamente, é possível decompor *mokà* em *mo* ‘veado’ + *kà* ‘pele, couro, casca’ (literalmente, ‘couro de veado’), embora de um ponto de vista semântico tal composto não pareça ser justificado, visto que o *mokà* não é feito de couro de veado e sim de fibra de tucum. As permutações desse tipo são comuns no vocabulário não nativo de várias línguas: por exemplo, a palavra *cheeseburger* do inglês (derivada de *cheese* /tʃi:z/ ‘queijo’) foi emprestada ao português como *X-búrguer*, com a forma influenciada pelo vocábulo preexistente *chis* ‘letra X’.

É interessante notar que um termo semelhante existe em duas outras variedades Jê Setentrionais, Canela e Krahô (*mãco* /mãkɔ/ ‘mocó’; ALBUQUERQUE, 2012, p. 201). Canela/Krahô *mãco* não corresponde regularmente a Mëbêngôkre *mokà*, pois o cognato esperado de *mãco* em Mëbêngôkre seria **mãgo* */mãgɔ/ ou **mãngo* */mãŋɔ/ (e o cognato esperado de Mëbêngôkre *mokà*, caso fosse uma palavra nativa, seria Canela/Krahô **pohkà* */pɔʔkʰɔ/). É mais provável, portanto, que a palavra Canela/Krahô tenha sido

³⁵ O termo *mboko* do Guaraní Boliviano deu origem, via empréstimo, a espanhol *bocó* (regionalismo do oriente boliviano) e a Chiquitano *bokoj* /βo'ko-xi/ (no dialeto de San Miguel de Velasco; NIKULIN, 2019a).

³⁶ Segundo o dicionário Aurélio (FERREIRA, 1986, p. 1145), o termo *mocó* provém de “tupi *mo'kó*”, porém não pudemos comprovar em outras fontes a suposta existência do respectivo termo em Tupí Antigo ou em qualquer uma das Línguas Gerais.

emprestada, de forma independente, do português (ou talvez de uma língua Tupí-Guaraní não identificada).

Mëbêngôkre *xoko* /tʃɔkɔ/ ‘socó-boi’ < Língua Geral Amazônica <socó> ‘socó’ (STRADELLI, 1929, p. 646). A própria forma do termo *xoko* ‘socó-boi’ denuncia um provável empréstimo na língua, já que a africada *x* /tʃ/ ocorre nele em um ambiente atípico (em ataque silábico precedendo uma vogal diferente de /e/, ver seção 4). É evidente que esse vocábulo provém, em última instância, de uma variedade Tupí-Guaraní, pois reflexos de Proto-Tupí-Guaraní **tsoko* são encontrados em múltiplas línguas dessa família. Entretanto, na maioria das línguas Tupí-Guaraní que apresentam um reflexo de **tsoko* a africada **ts* do Proto-Tupí-Guaraní foi debucalizada (cf. *hoko* no Guaraní Paraguaio ou em Tenetehára) ou elidida, como em Wajãpi *oko* (MELLO, 2000, p. 184),³⁷ resultando em formas que não poderiam ter sido adaptadas em Mëbêngôkre como *xoko* /tʃɔkɔ/. As únicas línguas listadas por Mello (2000) que apresentam reflexos de **tsoko* que começam com uma africada ou fricativa surda coronal são o Ka’apor, o Guaráyu, o Sirionó e a LGA/Nheengatú.³⁸ Por motivos geográficos, as primeiras três línguas não podem ter sido servido de fonte para Mëbêngôkre *xoko*, deixando-nos com a LGA como a única opção plausível. Julgamos improvável que o empréstimo tenha sido intermediado pelo português, em que *socó* também é um empréstimo de uma das Línguas Gerais: os empréstimos do português estão geralmente ausentes na fala dos mais velhos, e são claramente identificados como tais pelos falantes, algo que não acontece com *xoko*.

(?) **Mëbêngôkre ’òk[ti] /ʔok[ti]/, ’òk[re] /ʔok[=rɛ]/ ‘espécie de batata’ < Tupí-Guaraní (língua não identificada) #ʔok ‘tubérculo’**. Se esta etimologia estiver correta, o suposto étimo de *’òkti* (com o sufixo aumentativo) ou *’òkre* (com o clítico diminutivo) seria um reflexo de Proto-Tupí-Guaraní **ʔok* ‘tubérculo’ (reconstrução de MELLO, 2000, p. 210). Mello (2000) acompanha a forma reconstruída de dois pontos de interrogação e cita apenas dois reflexos, Tembé *i-a-kwer* (com o sufixo de tempo passado) e Guayakí

³⁷ Mello (2000) fornece a reconstrução **oko* acompanhada de um ponto de interrogação. Aparentemente, trata-se de um erro de digitação, já que os reflexos citados pelo autor apresentam uma consoante inicial na maioria das línguas (e aquelas que não apresentam uma consoante, como Wajãpi *oko*, são as línguas em que a africada **ts* regularmente se elidiu).

³⁸ Poderíamos acrescentar à lista o Suruí-Aikewára (*soko* /sɔ’kɔ/, LOPES, 2014, p. 489), porém neste caso trata-se não de um reflexo de **tsoko* (cuja forma esperada em Suruí-Aikewára seria **oko*) e sim de um empréstimo (do português ou talvez da LGA). A consoante /s/ do Suruí-Aikewára pode continuar regularmente apenas a aproximante **j* do Proto-Tupí-Guaraní.

krape i'a ‘espécie de tubérculo’. A mesma reconstrução é dada por Corrêa-da-Silva (2010, p. 404), que cita ainda Awetí e Sateré-Mawé *ʔok* ‘tubérculo’; entretanto, não pudemos verificar a existência das formas supracitadas nas fontes de referência sobre o Awetí e o Sateré-Mawé. Embora a **raiz** em questão seja amplamente documentada nas línguas Tupí-Guaraní (e naquelas de outras famílias do tronco Tupí), a maioria esmagadora de suas atestações diz respeito ao composto com o significado ‘mandioca’ (Proto-Tupí-Guaraní **māndiʔok* **/maniʔok/*, Proto-Mundurukú **māsik*; MELLO, 2000, p. 176–7; PIKANÇO, 2019, p. 140). Sabemos de uma única língua Tupí em que o cognato de Proto-Tupí-Guaraní **-ʔok* foi atestado fora de compostos: trata-se da língua Tuparí, em que foi atestado o vocábulo *ek* ‘tubérculo’ (ALVES, 2004, p. 164). Esta forma não pode ter servido de fonte para Mëbêngôkre *’ôkti/’ôkre* tanto por motivos geográficos como por razões fonológicas (a diferença entre as vogais), mas é possível que em alguma língua Tupí-Guaraní (ou talvez o Awetí ou o Sateré-Mawé) o cognato de Tuparí *ek*, cuja forma esperada seria próxima a *#ʔok*, ainda ocorra fora de compostos e tenha sido emprestado para o Mëbêngôkre para designar um tipo de batata. Em síntese, trata-se de uma etimologia problemática, que poderia ganhar em credibilidade se futura pesquisa em documentação das línguas Tupí-Guaraní comprovar a possibilidade de ocorrência “solta” de reflexos de **ʔok*. Uma possibilidade alternativa, que nos parece pouco provável por motivos semânticos, é que a palavra *’ôkti/’ôkre* ‘batata’ esteja relacionada com *’ôk* ‘seiva; pintar’ (< Proto-Jê Setentrional **côk* ‘seiva; pintar’). Neste caso a extensão semântica (‘seiva’/‘pintar’ → ‘batata’) teria ocorrido somente em Mëbêngôkre, pois nas demais línguas Jê Setentrionais os reflexos de **côk* significam apenas ‘seiva’ e ‘pintar’.

(?) Mëbêngôkre *jaduj* /*jaduj*/ ‘curto’ < Tupí-Guaraní (língua não identificada) *#jatu-ʔi* ‘curto (diminutivo)’. A forma do Mëbêngôkre tem sido comparada com Apinajé *o* ‘*atuj* /*ɔ=ʔ-atuj*/ ‘logo’, Canela *jatuj* /*jatuj*/ ‘curto’ e Gavião-Pyhcopji/Krĩkatí *jatohj* /*jatoj*/ ‘dar meia volta, encurtar’, *jatohjre* /*jatoj-re*/ ‘curto’ (NIKULIN, 2020, p. 498). Entretanto, este é o único exemplo em que o Mëbêngôkre apresenta a consoante /d/ em ataque silábico correspondendo a /t/ nas demais línguas (ver seção 4 sobre a raridade de /d/ na língua), sugerindo que talvez se trate de empréstimos que entraram nas línguas Jê Setentrionais já diversificadas de uma terceira fonte. De fato, há palavras semelhantes (e semanticamente idênticas) em línguas Tupí-Guaraní dos chamados sub-ramos IV e V, que apresentam reflexos de **-atut* (por exemplo, Anambé *jatu* ‘ĩ, Apyãwa *ãtot*, Suruí-

Aikewára *iatuatur*; cf. JULIÃO, 2005, p. 82; ALMEIDA *et al.*, 1983, p. 80; LOPES, 2014, p. 464). Uma forma como aquela encontrada em Anambé (isto é, com a elisão da consoante final e o acréscimo de um sufixo diminutivo) é suficientemente próxima daquelas encontradas nas línguas Jê Setentrionais, mas mesmo se assumirmos que *jaduj* é um empréstimo das línguas Tupí-Guaraní, a ocorrência de uma oclusiva vozeada /d/ em Mëbêngôkre permanece sem explicação.

Finalmente, mencionamos duas palavras não etimologizadas do Mëbêngôkre que se assemelham remotamente a seus equivalentes na língua de Aurê e Aurá: **Mëbêngôkre** *ijêk* /*ijek*/ ‘macaco cuxiú’ e *akranhiti* /*akraɲi[ti]*/ ‘abacaxi’ ~ Aurê e Aurá *kʷitfe* ‘macaco cuxiú’, *akatfi* ‘pʷi’ ‘abacaxi’ (MELLO, 2000, p. 136–7). Observe que o conhecimento que se tem da fonologia da língua de Aurê e Aurá é apenas fragmentário, e as formas supracitadas podem ser imprecisamente atestadas. Acreditamos tratar-se de semelhanças espúrias (e não particularmente fortes), mas as deixamos registradas, pois atualmente não podemos propor etimologias melhores para *ijêk* e *akranhiti*.

5.3 Jê

Em subseções anteriores, vimos exemplos de vocábulos emprestados pelo Mëbêngôkre de línguas de outras famílias. Nesta subseção, em que discutimos o contato entre o Mëbêngôkre e algumas línguas estreitamente relacionadas da família Jê, apresentaremos exemplos tanto de prováveis empréstimos de outras línguas em Mëbêngôkre como de elementos lexicais e até mesmo fonológicos difundidos na direção contrária (isto é, do Mëbêngôkre para outras línguas).

Os povos de fala Jê que vivem em territórios adjacentes àquele ocupado pelos falantes de Mëbêngôkre são os Kîsêdjê (*Krwatire* para os Kayapó) e, mais recentemente, Panará (*Krãjakàrà* para os Kayapó) e Tapayúna. Não podemos descartar ainda a possibilidade de que alguns povos Jê, hoje desaparecidos, tenham estado presentes na área no passado.

Dois exemplos de possíveis empréstimos lexicais oriundos de uma outra língua Jê, talvez não mais existente, já foram mencionados na seção 2. Trata-se de *karinhô* /*karipo*/ ‘fumo’ e *xururu* /*tʃurur*/ ‘bico-de-brasa’, palavras estas obviamente relacionadas a Proto-Jê Setentrional **karên* ‘fumo’ e **jôrôr* ‘bico-de-brasa’. Entretanto, não há qualquer regularidade na correspondência entre as formas Mëbêngôkre e aquelas

das demais línguas Jê Setentrionais: os reflexos esperados em Mëbêngôkre seriam **karên* e **djôrôrô*, respectivamente. Dessa forma, é possível que as palavras *karinhô* e *xururu* tenham sido emprestadas pelo Mëbêngôkre de uma outra língua Jê, em que as vogais **ê/*ô* historicamente se converteram em vogais altas (*i/u*) e a palatal vozeada **j* se ensurdeceu. Os dois tipos de mudança sonora ocorreram em duas línguas do complexo dialetal Timbira (Gavião-Pyhcopji e Krĩkatí), bem como nas línguas do ramo Akuwẽ (Akwẽ-Xerente e Xavante); entretanto, é pouco provável que alguma dessas línguas tenha estado em contato com o Mëbêngôkre por motivos geográficos. Podemos hipotetizar que antigamente existia uma variedade Jê hoje desconhecida, possivelmente falada na região do Araguaia, em que os reflexos de **karên* e **jôrôr* tenham sido próximos a *#karijo* e *#jfurur*; essa mesma língua poderia ter sido a fonte de Iny *kabiri* ‘bacaba’ (compare PJS **kambêr*), discutido em 5.1. Após ter emprestado as palavras *karinhô* e *xururu* dessa língua hipotética, o Mëbêngôkre, por sua vez, teria sido a fonte de Panará *karijô* ‘fumo’ (DOURADO, 2001, p. 100).³⁹ Este último termo em Panará substituiu o vocábulo nativo *arên* (< **karên*), atestado (como <arená>, <arêna>, <aréne>) em fontes que documentam um estágio mais antigo do Panará, quando seus falantes ainda ocupavam um território mais ao sul e eram conhecidos como os “Kayapó do Sul” (VASCONCELOS, 2013, p. 146).

O povo Tapayúna/Kajkwakhrattxi, originário da região do Rio Arinos (bacia do Tapajós), passou a conviver com falantes do Mëbêngôkre — mais precisamente, com o grupo Mëtyktire da nação Kayapó — em decorrência de uma sequência de eventos que marcaram a história recente desse povo. Nos meados do século XX, o avanço das frentes coloniais no território ancestral dos Tapayúna, localizado entre as cidades de Juruena e Aripuanã, resultou em um etnocídio desse povo, cujos 41 sobreviventes foram transferidos entre 1969 e 1970 para o Parque Indígena do Xingu e inicialmente alocados com os Kĩsêdjê, falantes de uma língua estreitamente relacionada; logo após a transferência, mais dez Tapayúna faleceram. Os Tapayúna permaneceram com os Kĩsêdjê até 1988, ano em que muitos Tapayúna abandonaram a aldeia Kĩsêdjê e passaram a residir com os Mëtyktire, na T.I. Capoto-Jarina (CAMARGO, 2015, p. 34–43). Mesmo

³⁹ Steinen (1886, p. 359) registra Kĩsêdjê <kalinsó> ‘cigarro’, ao lado de <kalinę> ‘fumo’. A última palavra é atestada na língua moderna como *karêni* /kalen/ [ka'Jeni] (NONATO *et al.*, 2012, p. 12), com uma vogal média-alta /e/. Quanto à primeira palavra, poderia se tratar de uma transcrição imprecisa de *karên sô* (literalmente, ‘folha de tabaco’) ou, se assumirmos que Steinen (1886) registrou corretamente uma vogal alta, de um empréstimo da mesma fonte que Mëbêngôkre *karinhô*. Note que o alçamento de vogais médias-altas não faz parte da história fonológica do Kĩsêdjê.

após terem fundado em 2009 sua própria aldeia, Kawêrêtxikô, a língua Mëbêngôkre permanece dominante (D. LIMA, 2019, p. 281). A convivência dos Tapayúna com falantes do Kayapó nas últimas três décadas tem resultado em um caso interessante de uma situação de contato de línguas em que a transmissão horizontal ocorreu não somente no léxico, mas também na fonologia. Camargo (2010, p. 34–9; 2015, p. 52–4) discute algumas instâncias de troca de vocábulos originalmente Tapayúna por seus cognatos em Mëbêngôkre; segundo a autora, os Tapayúna mais propensos à influência dessa última língua pertenciam a uma mesma faixa etária e tinham cerca de 40 anos em 2010 (ou seja, são os Tapayúna nascidos aproximadamente à época de transferência para o Parque Indígena do Xingu). Exemplos dessas trocas lexicais, mencionados por Camargo (*op. cit.*), incluem *kumrēj* ‘muito’, *mry* ‘bicho’ (substituindo os vocábulos Tapayúna *kuwěj* /kuwěj/, *nry* /ři/). Na fonologia, destaca-se a opcionalidade da fase oral na realização das nasais subjacentes em ambientes orais: por exemplo, Camargo (2010, p. 72–3) documenta realizações tais como [ˈnda] ‘chuva’, [ˈnda] ‘olho’, [ˈnde] ‘ariranha’, [ˈdɿ] ‘lança’, [ˈndiwi] ‘novo’ em variação livre com [ˈna], [ˈnɔ], [ˈne], [ˈnɿ], [ˈniwi] (representações subjacentes: /na/, /nɔ/, /ne/, /nɿ/, /niwi/). As realizações com a pós-oralização são, evidentemente, mais conservadoras, pois o fenômeno é obrigatório em outras línguas Jê Setentrionais (tais como Kĩsêdjê e Apinajé) e é reconstruível também para o Proto-Jê Setentrional. Em contraste, a ocorrência de fones nasais em ambientes orais é característica apenas do Mëbêngôkre, em que a ausência de pós-oralização constitui uma inovação. É, portanto, provável que a influência do Mëbêngôkre subjaz a introdução da variação livre entre [ˈnd] e [ˈn] em Tapayúna reportada por Camargo (2010).

Já em Kĩsêdjê — uma língua estreitamente relacionada ao Tapayúna — não foram encontrados claros casos de empréstimos Mëbêngôkre, com a possível exceção de *khwě* /k^hwě/ ‘indígena de outra etnia’ (NONATO *et al.*, 2012, p. 16; observe também *khwětê* /k^hwětê/ ‘rede’, literalmente “o pano de *khwě*”), que poderia ser emprestado de Mëbêngôkre *kubě* ‘estrangeiro’. Note que o Kĩsêdjê não possui oclusivas vozeadas, tais como /b/, em seu inventário fonológico, sendo concebível a adaptação de /b/ de outra língua como /w/. Kĩsêdjê *khwě* não pode ser um cognato de Mëbêngôkre *kubě*, pois esta última palavra corresponde regularmente a Kĩsêdjê *khupě* /k^hupě/ ‘indígena’ (ambas de Proto-Jê Setentrional **kubě*); observe também os derivados *khupě khá* /k^hupě-k^hɜ/ ‘roupa’ (Proto-Jê Setentrional **kubě-ʔkə*) e *khupěkhátxi* /k^hupě-k^hɜ-tʃi/ ‘não-indígena’ (NONATO

et al., 2012, p. 15). Como o reflexo regular de Proto-Jê Setentrional **b* em sílaba tônica (final) em Kĩsêdjê é /p/, somos obrigados a considerar *khupẽ* uma palavra nativa (herdada da protolíngua); já *khwẽ* deve ser interpretado como um empréstimo de uma língua relacionada. Fonologicamente, seria possível derivar o vocábulo Kĩsêdjê de Tapayúna *khuwẽ* /k^huwẽ/ ‘pessoa de outra etnia’ (D. LIMA, p. 200, nota 36), porém os registros de Steinen (1886, p. 358) mostram que a palavra *khwêê* ‘rede’ (derivada de *khwẽ*; na transcrição de Steinen, <kuetê>) existia em Kĩsêdjê já no século XIX, ou seja, cerca de cem anos antes do reencontro dos Kĩsêdjê com os Tapayúna.

5.4 Yudjá

Os Yudjá (ou Jurúna) já foram mencionados acima: esta nação de canoieiros, logo de ser dominante no médio e baixo Xingu nos séculos XVII a XIX, procurou refúgio no alto Xingu fugindo de ataques dos Měbêngôkre-Kayapó Měkrākñōti. Sua língua, juntamente à língua Xipáya, pertence à família Jurúna do tronco Tupí. Encontramos um caso claro e mais um caso duvidoso de transmissão horizontal de léxico do Yudjá para o Měbêngôkre.

Měbêngôkre *karaxu* /karatʃu/ ‘colher’ < Yudjá *karaxu* /karafu/ ‘colher’ (LIMA, 2014, p. 26). A origem não nativa de Měbêngôkre *karaxu* ‘colher’ pode ser inferida com base na ocorrência da africada *x* /tʃ/ em um ambiente atípico (em ataque silábico precedendo uma vogal diferente de /e/, ver seção 4) e é comprovada pela ausência de cognatos em outras línguas Jê. Em contraste, a palavra Yudjá é de uma maior antiguidade, pois ela corresponde regularmente a Xipáya /kaʃa'fu/ ‘colher’ (FARGETTI, RODRIGUES, 2008, p. 542), podendo continuar Proto-Jurúna **karafu*. A adaptação de Yudjá /f/ como Měbêngôkre /tʃ/ é natural, pois o Měbêngôkre carece de fricativas em seu inventário fonológico (observe que caso se assumisse um empréstimo na direção inversa, seria muito difícil explicar por que a africada /tʃ/ teria sido adaptada como /f/ em Yudjá, já que a última língua possui tanto /f/ como /tʃ/).

Palavras quase idênticas são encontradas em mais duas línguas da região. Em Arara/Ugrongmo (família Carib), o termo para ‘colher’ é *karatju* /karatʃu/ (SOUZA, 2010, p. 64), identificado por Souza (2010, p. 7) como um empréstimo do Kayapó (entretanto, nenhum argumento para sustentar a direcionalidade é apresentado).

Acreditamos que não seria menos plausível postular, nesse caso, um empréstimo direto do Yudjá para o Arara, principalmente considerando que o único grupo Kayapó com o qual os Arara estão em contato — os Kararaô — chegaram ao baixo Iriri apenas no início da década de 1980. Notamos também que, segundo Farias (2017, p. 6), Xipáya /kaʃa'ʃu/ foi emprestado para a língua Kuruáya, da família Mundurukú, em que é atestada a forma /káraʃo/ ‘colher’; não é provável que se trate de cognatos herdados do Proto-Tupí, pois a ocorrência de /ʃ/ em Kuruáya (e em Proto-Mundurukú) é restrita ao ambiente *_/a,e/* no vocabulário nativo, com a exceção de *wáʃo* ‘sagui’ (PICANÇO, 2019, p. 103). Dessa forma, reflexos de Proto-Jurúna **karaʃu* teriam sido fontes (diretas ou indiretas) de termos para ‘colher’ em pelo menos três línguas não relacionadas: Mëbêngôkre (Jê), Kuruáya (Mundurukú) e Arara (Carib).

(?) Mëbêngôkre *awa* /awa/ ‘ave sp. (*Troglodytidae*)’ < Yudjá *uxixi auahanu* ‘garrinchão (*Campylorhynchus turdinus*)’ (BERTO, 2013, p. 119). O termo Mëbêngôkre não possui uma etimologia Jê conhecida; a primeira parte do composto Yudjá é *uxixi* /ùʃiʃi/ é um termo genérico para ‘pássaro’. Note que o garrinchão faz parte da família *Troglodytidae*, fazendo com que a comparação seja semanticamente plausível. Entretanto, a mesma é obviamente dificultada pela ausência de qualquer correspondência para as sílabas *-hanu* do Yudjá em Mëbêngôkre. Uma possibilidade que consideramos é que o termo *auahanu* da língua Yudjá possa ter sido equivocadamente percebido por falantes não nativos (isto é, pelos Mëbêngôkre) como uma construção de predicação não-verbal no aspecto imperfectivo não-progressivo, que envolve, em Yudjá, o sufixo predicativizador *-ha* e a partícula de aspecto *anu*, como em *Wetaki peruha anu* ‘Wetag é Kĩsêdjê’ (de *peru* ‘Kĩsêdjê’), *pitxaha anu* ‘é peixe’ (de *pitxa* ‘peixe’), *yudjaha anu* ‘é Yudjá’ (FARGETTI, 2001, p. 110, 198). Tal possibilidade, contudo, permanece altamente especulativa, e por ora hesitamos em afirmar que Mëbêngôkre *awa* seja um empréstimo do Yudjá (ou vice-versa).

Um topônimo Mëbêngôkre que poderia remontar à língua Yudjá é *Pixaxa* /pitʃatʃa/ ‘rio Curuaés’, cujas primeiras duas sílabas se assemelham a Yudjá *pitxa* /pitʃà/ ‘peixe’ (FARGETTI, 2001, p. 71); contudo, se o étimo de *Pixaxa* for mesmo um composto Yudjá, ainda não sabemos qual poderia ser seu segundo elemento.

5.5 Outras

Há muitas outras palavras Mëbêngôkre que não têm cognatos nas línguas mais próximas, mas para as quais não pudemos encontrar um étimo plausível em outras línguas faladas na região. Muitas destas palavras têm características fonológicas que violam os padrões fonotáticos mais robustos do Mëbêngôkre. Segue uma lista não exaustiva: *étuwa* /etu/ ‘calango’, *pyxanhre* /puɬʃaŋ=ɾe/ ‘pau-brasil’, *mokokti* /møkøk-ti/ ‘poraquê’, *ibê* /ibe/ ‘surubim pintado’, *mrê’êti* /mreʔe-ti/ ‘maracajá’, *tey* /teu/ ‘ferrão, cauda de ave’, *koneh* /kɔŋɛɲ/ ‘centro, metade’ (e seu provável derivado *konetã* /kɔŋɛtã/ ‘filho do meio’), *nhigo* /ɲigo/ ‘ninar’, *onĭja* /ɔ=ni/ ou *nĭjari* /ni=ri/ ‘longe’, *ngy* /ɲu/ ‘rastejar, avançar como uma lagarta’, *jênh* /jeɲ/ ‘carga’, *janhy* /jaɲu/ ‘restos’, *kamjôrôti* /kamjôr-ti/ ‘espécie de bagre’, *kràk* /kɾak/ ‘chamar’, *panhõp* /pa-ɲõp/ ‘cotovelo’ (de *pa* ‘braço’), *pê-nàt* /pe-nat/ ‘lado’ (de *nàt* ‘cintura’), *kangrõnh* /kaɲrõɲ/ ‘ameaçar’, *ràx* /ɾatʃ/ ‘arame’ (e seu provável derivado *kàxiràx* /katʃiratʃ/ ‘prego’), *jabatàn* /jabatɲ/ (Kayapó) ou *jabatàri* /jabatɾ/ (Xikrin) ‘crescido’, *jaxwe* /jatʃwe/ ‘maluvido’, *kaxêt* /katʃet/ ‘pontilhado’, *awàrà* /awɾɾ/ ‘pau-de-cheiro’, *koxek* /kɔʃɛk/ ‘abrir as pernas’, *xiprô(ti)re* /tʃipro(-ti)=ɾe/ ‘andorinha’, *xyryxyry* /tʃur~tʃur/ ‘sal; árvore *sp.*’, *xãrixãri* /tʃãɾ~tʃãɾ/ ‘guizo’, *xôkrã* /tʃokɾã/ ‘pau-d’arco’, *xoxot* /tʃɔʃɔt/ ‘gralha’, *wajapyxô* /wajaputʃo/ ‘camarão’, *tàkàk xôti* /tɾakɾk tʃo-ti/ ‘timbó’, *pêkã(re)* /pekã(=ɾe)/ ‘alma-de-gato’, *dyjdyj* /duj~duj/ ou *dujduj* /duj~duj/ ‘ave *sp.*’, *kawe* /kawe/ ‘esquina’, *wapĭja* /wapĩ/ ‘gavião-de-anta’.

Além destas palavras, há vários nomes próprios que têm uma estrutura fonotática um tanto surpreendente.⁴⁰ O fonema /tʃ/ é inusualmente frequente nos nomes próprios, estando presente nos nomes de dois personagens mitológicos, *Xwÿnhpêkrô* e *Xakawãpÿ*, e em alguns nomes comuns, tais como *Xikiri*, *Xwak(re)* e *Xikatô*. Além destes, são possivelmente de origem externa os nomes *Kamajôra*, *Wajwaj*, *Wanaô*, entre outros.

⁴⁰ De fato, a maioria dos nomes próprios Mëbêngôkre têm uma etimologia transparente, tanto os *nomes bonitos*, que começam com radicais não-analisáveis associados a um ritual de nomeação (p. ex., *Ngrej-bôrô* ‘Ngrej assada’), quanto os nomes ordinários, que são formados a partir de nomes comuns modificados (como *Rop-ni* ‘onça fêmea’). Qualquer nome que não pode ser interpretado composicionalmente é um empréstimo em potencial. Não sobra lembrar que os nomes são considerados um valor (LEA, 2012, cap. 7), e sua incorporação à sociedade Mëbêngôkre é equivalente à incorporação de um item de cultura material.

Finalmente, devemos lembrar que os etnônimos Kayapó e Xikrin que designam as duas principais subdivisões dos Mëbêngôkre são de origem incerta.⁴¹ Em Mëbêngôkre são termos de uso relativamente recente, introduzidos a partir do português.

6. Contato com o português

Por limitações de espaço, não podemos examinar em detalhe as diferentes estratégias de adaptação dos numerosos empréstimos oriundos do português na língua Mëbêngôkre atual. A marca do português no léxico Mëbêngôkre tem características que o diferenciam das línguas discutidas na seção 5: por um lado, a duração e a intensidade do contato entre os falantes do Mëbêngôkre e do português é maior, permitindo-nos distinguir entre camadas cronológicas de empréstimos, que exibem graus variados de nativização fonológica e/ou morfológica; por outro, o número de falantes bilíngues é hoje considerável e em constante crescimento, o que resulta em níveis diferentes de adaptação de empréstimos segundo o nível de bilinguismo do/da falante, bem como na possibilidade de empréstimos ocasionais, que não chegam a se estabilizar na língua. Inclusive, poderia se falar em um constante “reempréstimo” de certos termos, na medida em que os falantes se tornam melhores conhecedores do português. Por esses motivos, o estudo dos empréstimos do português deveria levar em consideração pelo menos as diferenças de faixa etária, região geográfica, e gênero dos falantes. Finalmente, no estudo dos empréstimos do português seria possível e desejável diferenciá-los segundo a fonte, entre os muitos grupos de *kubê kryt* com os quais os Mëbêngôkre estiveram em contato (ribeirinhos, frentes de atração, missionários, funcionários do estado, meios de comunicação, etc.), pois estes teriam características fonéticas diferentes que poderiam resultar em adaptações diferentes. Tudo isso está muito além do que podemos fazer aqui. Limitamo-nos, então, a fazer um breve sobrevoo de alguns empréstimos mais bem estabelecidos que ilustram diferentes tipos de adaptação fonológica.

Mëbêngôkre *kape* [ka'pɛ] ‘café’. Este exemplo ilustra o tipo mais simples de adaptação, em que um fone do português ([f]) que carece de um equivalente em Mëbêngôkre é substituído pelo fonema nativo mais próximo (/p/).

⁴¹ O termo *Kayapó*, de suposta origem Tupí-Guaraní e empregado originalmente para os Kayapó do Sul, é etimologizado em algumas fontes como ‘feito macaco’ (LEA, *op. cit.*, p. 60). Consideramos que esta etimologia é implausível: embora *ka'i apo* possa significar ‘fazer macaquite’ ou ‘fabricar macacos’ em línguas como o Guaraní Paraguaio, o verbo *apo* não tem a leitura de ‘feito, igual a’ por si só.

Mëbêngôkre *kaĩ* [ka'ĩ] ~ [ka'hĩ] ‘carrinho de mão’. O fone [h] existe na fala cuidadosa em Mëbêngôkre como um reflexo de *c em posição inicial e intervocálica, porém na análise fonológica comumente adotada corresponde à ausência de um fonema consonantal. Algo que é contrastivo em português é transferido como um segmento em variação livre com zero. Não constatamos que o [h] seja mais estável neste caso que no léxico nativo. A apócope da sílaba final talvez não exija uma explicação muito complexa: a pronúncia portuguesa desse item é com frequência bissilábica: [ka'hĩõ] ~ [ka'hĩ].

Mëbêngôkre *kratô* [kra'to] ‘trator’. Embora todos os fones presentes em [tra'to] existam em Mëbêngôkre, a combinação de /t/ e /r/ é proibida em ataque silábico, pois ambos os segmentos compartilham o mesmo ponto de articulação (ver seção 4 e SALANOVA, 2001). Portanto, a sequência [tr] do português é adaptada aqui como /kr/. Este é um exemplo relativamente estável, mas na fala de pessoas monolíngues estas adaptações encontram-se com facilidade (por ex., palavras como *antropólogo* são adaptadas a *kroporo* [krɔ'pɔrɔ]).

Mëbêngôkre *kubêta* [kubē'ta] ‘cobertor’. A fonte deste empréstimo é ‘coberta’ e não ‘cobertor’. Há uma mudança de acento à última sílaba, e uma nasalização espontânea do [ē]. Tanto a nasalização como a mudança de acento poderiam ser atribuídas à analogia: a palavra *kubêjaê* ‘rede’, analisada transparentemente como /kubē+jæ/ ‘não-indígena’ + ‘ninho’, seria um modelo para reanalisar as duas primeiras sílabas de *kubêta* como ‘não-indígena’, mesmo se o complemento /ta/ não tem significado independente.⁴² A mudança da posição do acento segue o padrão dominante da língua, o oxítono.

Mëbêngôkre *xibôro* [tʃi'boro] ‘cebola’. O problema da acentuação paroxítona é resolvido por meio da eliminação da vogal final da representação fonológica: a vogal [o] final é sincronicamente uma vogal eco átona inserida de forma obrigatória após /r/. A adaptação da primeira vogal como /i/ ao invés de /e/ é ligeiramente surpreendente.

Mëbêngôkre *mukutêru* [muku'teru] ‘mosquiteiro’. O acento recai em uma sílaba onde normalmente não ocorreria em Mëbêngôkre. A transformação de [i] em [u] na segunda sílaba poderia obedecer à tendência de as sílabas pretônicas serem

⁴² Existe a raiz verbal *ta* ‘arrancar’, mas trata-se de uma coincidência: toda derivação de nomes deverbais em Mëbêngôkre dá-se a partir da forma não finita dos verbos, que no caso do verbo *ta* ‘arrancar’ é 'yr.

comumente escolhidas a partir de um pequeno inventário, do qual [ku] (mas não [ki]) faz parte. A queda da coda na primeira sílaba poderia se dever ao mesmo princípio.

Mébêngôkre *mikrôni* [mi'kroni] 'micro-ônibus'. Embora a acentuação paroxítona seja por vezes preservada nos empréstimos, a acentuação proparoxítona é evitada. Neste caso, isto se obtém por meio de apócope da última sílaba.

Curiosamente, muitos conceitos de introdução recente não são designados por empréstimos do português, como é o caso em muitas outras línguas do Brasil. Isto se deve, sem dúvida, à grande facilidade da língua para cunhar novas palavras a partir de recursos nativos. De um ponto de vista sintático, as estratégias empregadas são inteiramente produtivas na língua. As limitações de espaço não permitem que nos adentremos nesta questão aqui, mas damos alguns exemplos que mostram como são cunhados alguns destes termos. Muitos deles são sintagmas genitivos: *rop no* 'lanterna' (lit. 'olho da onça'), *kàj puroro* 'enxada' (lit. 'faca achatada'), *ngôj krã* 'copo' (lit. 'cabeça da panela', em que *krã* 'cabeça' é empregado como um termo de classe para objetos redondos), *kàj djwa ngri* 'canivete' (lit. 'dente de faca pequeno'), *kà mrãnh tỳx* 'automóvel' (lit. 'canoa de caminhar forte'), *krã kam ngôj* 'soldado' (lit. 'panela na cabeça'), *tep pynê djà* 'tarrafa' (lit. 'instrumento/lugar de prender peixe'), *pi'ók jakre djwỳnh* 'professor' (lit. 'mostrador de escrita'), *kubẽ kà kayry pro* 'roupa' (lit. 'cobertura de pano (= pele/casca de estrangeiro) costurado').⁴³ Os núcleos de sintagmas nominais podem tomar por complemento não apenas nomes ou verbos não finitos, mas também sintagmas posposicionais: *kape mã ukanga* 'café instantâneo' (lit. 'preguiça para café'), *krã kam ngôj* 'soldado' (lit. 'panela na cabeça').

Apesar desta facilidade para criar léxico para novos conceitos, é interessante identificar dois domínios em que são usados vários termos de origem portuguesa: por um lado, empregam-se alguns vocativos para termos de parentesco (*mãmãj, papaj, wôwô, wowo, xiw, xija* < *mamãe, papai, vovô, vovó, tio, tia*); por outro, alguns nomes de peixes em português coexistem e às vezes deslocam os termos nativos; é o caso de *paku* /*paku*/

⁴³ Note que em todos os casos o núcleo **sintático** desses sintagmas nominais é a última palavra na sequência (*no* 'olho', *puroro* 'raso, chato', *krã* 'cabeça', *ngri* 'pequenez', *tỳx* 'dureza', *djà* 'instrumento/lugar', *djwỳnh* 'dono', *pro* 'cobertura'), respeitando o padrão geral de núcleo final da língua; em contraste, o núcleo **semântico** não é necessariamente final (nos exemplos acima, não o é em *kàj puroro* 'enxada' e *kà mrãnh tỳx* 'automóvel', cujos núcleos semânticos são, respectivamente, *kàj* 'faca' e *kà* 'canoa, casca/pele, invólucro'. Um tratamento detalhado da sintaxe do sintagma nominal é oferecido em Salanova (2017).

‘pacu’ entre alguns grupos Kayapó, substituindo *djurorotire* /dzurɔɾ-ti=ɾɛ/, que é ainda usado pelo menos pelos Xikrin e Mětyktire. A ocorrência destes empréstimos poderia parecer surpreendente, dada a centralidade das relações de parentesco e a familiaridade dos falantes com a fauna ictícola. Porém, é possível oferecer, mesmo que de maneira especulativa, uma explicação para cada um desses casos. Os termos de parentesco Měbêngôkre não denotam o mesmo que denotam os termos portugueses: *nã* /nã/, que costuma dar-se como equivalente de *mãe*, inclui também as irmãs da mãe, as esposas dos irmãos do pai, e as filhas do irmão da mãe; *kwatyj* /kwatuj/, que costuma traduzir-se por *avó*, compreende as irmãs do pai, entre outras. O uso de termos de parentesco portugueses poderia estar associado a mudanças recentes nos padrões de residência, que enfatizam a família nuclear. Quanto ao vocabulário para os peixes, trata-se de um conhecimento especializado, cujo aprendizado completo é algo valorizado como marca de sabedoria: os ictiônimos em português são conhecidos por muitas pessoas que, em geral, têm pouca proficiência nessa língua. Não é impossível que alguns tenham substituído os termos nativos.

Existem também nomes próprios tomados do português, alguns em aparência bastante antigos, a julgar por seu grau de nativização (por exemplo, *Jobaw* < *Iovaldo*).

7. O que dizem os dados linguísticos

Nas seções anteriores vimos que o Měbêngôkre possui diversos empréstimos lexicais de outras línguas, tais como o Xambioá (*waxi* ‘linha’, *benorã* ‘tucunaré’, *awo* ‘tipo de árvore; ubá’, *ixe* ‘espelho’, *warikoko/watkoko* ‘cachimbo’, *rorirori* ‘tipo de cocar, coifa plumária’, *warabaê* ‘cesto estojiforme’, *bikwa* ‘parente’, *bero* ‘puba’, *rara* ‘cesto bolsiforme’, *wiwi* ‘canto’), a Língua Geral Amazônica (*môtôbi* ‘y’ ‘amendoim’, *xãn* ‘gato’, *mokà* ‘mocó’, *xoko* ‘socó-boi’) e o português, bem como alguns empréstimos menos numerosos da língua Yudjá (*karaxu* ‘colher’, ? *awa* ‘garrinção’) e pelo menos dois empréstimos de uma provável língua Jê extinta (*karinhô* ‘fumo’, *xururu* ‘bico-de-brasa’). Vimos também que empréstimos de origem Měbêngôkre são encontrados em Panará (*karijô* ‘fumo’), Tapayúna (*mry* ‘bicho’, *kumrêj* ‘muito’) e talvez Kîsêdjê (*khwê* ‘índigena de outra etnia’) e Arara/Ugrongmo (*karatju* ‘colher’).

Algumas das situações de contato que subjazem os empréstimos supracitados já haviam sido identificadas em estudos etnolinguísticos (cf. RIBEIRO, 2012 para o contato

Xikrin–Xambioá); outras se dão no presente (Kayapó–Panárá, Kayapó–Tapayúna, Kayapó–Kĩsêdjê, Mëbêngôkre–português). A existência de empréstimos Yudjá em Mëbêngôkre não havia sido descrita em trabalhos anteriores, mas também não é particularmente surpreendente, dado o histórico de contato entre os dois povos. Entretanto, a existência de empréstimos da Língua Geral Amazônica é um dado novo e surpreendente, sugestivo de contatos diretos ou indiretos que não haviam sido reconhecidos antes na literatura. Além disso, há evidências linguísticas mais especulativas de prováveis contatos com falantes de uma ou várias línguas Tupí-Guaraní não identificadas. Ainda mais hipotética, embora plausível, é a suposta existência de um povo Jê não identificado no Araguaia, cuja língua teria sido a fonte de Mëbêngôkre *karinhô* ‘fumo’ e *xururu* ‘bico-de-brasa’ e de Iny/Karajá *kabiri* ‘bacaba’. Não é impossível que outros grupos Jê Setentrionais com os quais os ancestrais dos Mëbêngôkre tiveram um contato mais ou menos intenso tenham desaparecido ou tenham se fusionado com eles após um declínio populacional importante, tal como documenta Nimuendajú (1946, p. 30 *et passim*) a respeito de alguns grupos Timbira.

Juntamos estas informações ao que já sabemos sobre as relações internas do ramo Jê Setentrional, bem como às informações sobre a distribuição geográfica das espécies cuja designação se reconstrói para o Proto-Jê Setentrional,⁴⁴ para propor um quadro geral para a história deste agrupamento etnolinguístico.⁴⁵

1. Os falantes do Proto-Jê de Goyaz habitavam uma região ainda não identificada do cerrado, onde estiveram em contato com alguma língua do norte amazônico, da qual tomaram os empréstimos **kukôj* ‘macaco’ e **ndý* ‘novo’ (NIKULIN, CARVALHO, 2018, p. 557). Os ancestrais dos Panará se separam destes entre os séculos VI e IV a.C., e mais tarde estabelecem-se mais ao sul.

⁴⁴ A nossa fonte para determinar a distribuição das espécies é o Herbário Virtual da Flora e dos Fungos, elaborado pelo Instituto Nacional de Ciência e Tecnologia (INCT-HVFF) com apoio financeiro do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), disponível *online* em <<http://inct.splink.org.br>> e consultado entre julho e agosto de 2020. O INCT-HVFF não indica diretamente a distribuição das espécies e sim a origem geográfica de espécimes individuais depositados em coleções zoológicas em museus e centros de pesquisa do Brasil e do resto do mundo. Por esse motivo, as informações muitas vezes não são um bom indicador da extensão máxima de uma espécie. Além disso, as localizações referem-se a tempos relativamente recentes, e poderiam não coincidir com aquelas válidas nos períodos aos quais fazemos referência em nossa reconstrução.

⁴⁵ Apesar de serem muitas as reservas da comunidade científica a respeito do método glotocronológico que nos permite datar as diferentes etapas descritas nesta síntese, é o único recurso de que dispomos para fazer isto na base de dados estritamente linguísticos. As datações são obtidas aplicando escrupulosamente a fórmula proposta por Vasilyev e Saenko (2017, ver nota 3).

2. Os falantes do Proto-Jê Setentrional estabelecem-se provavelmente no centro-norte do Cerrado, ao leste da localização atual dos povos Jê Setentrionais, mas em um bioma semelhante (talvez no planalto da bacia do Parnaíba).
3. Os ancestrais dos Timbira se separam do resto da família aproximadamente nessa região entre os séculos II e VII; os ancestrais dos demais grupos migram ao oeste, ao interflúvio Tocantins-Araguaia.
4. Os Apinajé se separam do resto da família entre os séculos V e IX e permanecem no interflúvio Tocantins-Araguaia; já os ancestrais dos Mëbêngôkre, dos Kĩsêdjê e dos Tapayúna atravessam o Araguaia em uma única migração.
5. Os ancestrais dos Mëbêngôkre e dos Kĩsêdjê/Tapayúna se separam ao oeste do Araguaia entre os séculos VIII e X, com a migração dos ancestrais dos Kĩsêdjê e dos Tapayúna para a bacia do Tapajós.
6. Os Mëbêngôkre se estendem em direção norte a partir dos campos do Araguaia. A divisão entre Kayapó, ao sul, e Xikrin, ao norte, data de cerca de 1800 (VRESVIJWER, *op. cit.*, p. 87–8). As altas margens de erro do cálculo glotocronológico nessa profundidade temporal impossibilitam a datação da cisão com base em dados linguísticos. Foi nesse período que os empréstimos de origem Xambioá, Yudjá e LGA entraram no léxico da língua.
7. Durante o século XIX, os Mëbêngôkre circulam entre o Araguaia e o Xingu, migrando ao oeste em tempos relativamente recentes. A separação entre os Kayapó Irã'ãmrahre e os Kayapó Gorotire, e entre estes e os Mëkrãknôti são datadas por Verswijver como ocorrendo respectivamente em torno a 1840 e em 1905–6. Os empréstimos do português são ainda mais recentes, podendo ser datados da segunda metade do século XX.

A novidade nesta síntese é a proposta de que os ancestrais dos Kĩsêdjê/Tapayúna e dos Mëbêngôkre atravessaram o Araguaia como um único grupo, *pace* Verswijver (1992) e Turner (1998), que hipotetizam que os ancestrais dos segundos teriam se separado dos Apinajé há relativamente pouco tempo e permanecido no interflúvio Tocantins-Araguaia até o início do século XIX.

Será possível elaborar um quadro mais detalhado da localização geográfica dos diversos grupos pré-históricos com um avanço nos estudos do léxico de flora e fauna das diversas línguas Trans-Araguaia; resta, entretanto, o empecilho de conhecer a distribuição precisa das espécies animais e vegetais em tempos anteriores ao presente.

8. Considerações finais

O presente trabalho é uma primeira tentativa de sistematização do léxico Mëbêngôkre de origem não nativa. Restam muitas palavras ainda não etimologizadas que devem ser comparadas com o vocabulário de línguas faladas em regiões próximas. É de se esperar que este se torne mais bem conhecido à medida que a pesquisa sobre línguas indígenas se intensifica. Notamos, porém, com uma certa contrição que a pesquisa sobre o léxico continua sendo protelada, seja porque a maioria do/as linguistas profissionais se interessam por outras áreas, seja porque poucos têm a disposição de passar longos períodos em campo e conhecer as línguas em profundidade. Os missionários protestantes da atualidade, que, de fato, chegam a ter esse conhecimento, têm pouco interesse no vocabulário cultural e de flora e fauna, e na maioria dos casos publicam dicionários de baixa qualidade, em que as formas são, via de regra, corretamente atestadas, mas suas nuances semânticas e suas propriedades morfossintáticas não são discutidas em suficiente detalhe. A grande promessa nesta área de conhecimento é a formação de falantes nativos para a pesquisa linguística. O conhecimento sobre o léxico, facilmente acessível aos falantes nativos, aguarda a sistematização que não apenas serviria à comunidade científica, mas também seria uma das aplicações mais relevantes da linguística para o uso dentro das comunidades.

Referências

ALBUQUERQUE, Francisco Edviges (org.). *Më ipê Krahô catêjê te amjĩ ton xà: Arte e cultura do povo Krahô*. Belo Horizonte: Universidade Federal de Minas Gerais, 2012.

ALMEIDA, Antônio; JESUS, Irmãzinhas de; PAULA, Luiz Gouvêa de. *A língua Tapirapé*. Rio de Janeiro: Biblioteca Reprográfica Xerox, 1983.

ALVES, Poliana Maria. *O léxico do Tuparí: proposta de um dicionário bilíngüe*. 2004. viii + 286 f. Tese (Doutorado em Linguística e Língua Portuguesa) – Faculdade de Ciências e Letras (campus de Araraquara), Universidade Estadual Paulista “Júlio de Mesquita Filho”. 2004.

BERTO, Flávia de Freitas. *Kania ipewapewa: estudo do léxico juruna sobre a avifauna*. 2013. xxi + 136 f. Dissertação (Mestrado em Linguística e Língua Portuguesa) – Faculdade de Ciências e Letras (campus de Araraquara), Universidade Estadual Paulista “Júlio de Mesquita Filho”. 2013.

BOČEK, Vít. Etymology and language contact studies: some notes on mutual relationships. In: VYKYPĚL, Bohumil; BOČEK, Vít (eds.). *Etymology: An Old Discipline in New Contexts*, p. 13–21. Praga: Lidové noviny, 2013. (Studia Etymologica Brunensia 16).

BROCHADO, José J. J. Proenza. *An ecological model of the spread of pottery and agriculture into Eastern South America*. 1984. 574 f. Tese (Doutorado em Antropologia) – Department of Anthropology, University of Illinois at Urbana-Champaign. 1984.

CAMARGO, Nayara da Silva. *Língua Tapayúna: aspectos sociolingüísticos e uma análise fonológica preliminar*. 2010. 117 f. Dissertação (Mestrado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas. 2010.

CAMARGO, Nayara da Silva. *Tapayuna (Jê): aspectos morfossintáticos, históricos e sociolingüísticos*. 2015. 215 f. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas. 2015.

CARVALHO, Fernando O. de. Tupi-Guarani loanwords in Southern Arawak: taking contact etymologies seriously. *Linguística*, Rio de Janeiro, v. 13, n. 3, 41–74, 2017.

CARVALHO, Fernando O. de. Arawakan-Guaicuruan language contact in the South American Chaco. *International Journal of American Linguistics*, Chicago, v. 84, n. 2, p. 243–63, abr. 2018.

CARVALHO, Fernando O. de. Etymology meets ethnohistory: Linguistic evidence for the pre-historic origin of the Guaná-Chané in the Northwestern Chaco. *Anthropological Linguistics*, Bloomington, a sair.

CENTURIÓN SERVIN, Celsa; DAVALOS ARCE, Juana (revs.). *Milibri pocket. Castellano-guaraní, guaraní-castellano: grafía tradicional y científica*. Assunção: Milibri Ediciones, 2009.

CROCKER, William. *The Canela (Eastern Timbira), I. An ethnographic introduction*. Smithsonian Contributions to Anthropology, 33. Washington, EUA: Smithsonian Institution Press, 1990.

CRUZ, Aline da. *Fonologia e gramática do Nheengatú: a língua geral falada pelos povos Baré, Warekena e Baniwa*. Utrecht: LOT, 2011.

DOBSON, Rose. *Gramática prática com exercícios da língua Kayabi*. Cuiabá: Sociedade Internacional de Lingüística, 1997.

DOURADO, Luciana Gonçalves. *Aspectos morfossintáticos da língua Panará (Jê)*. 2001. xv + 240 f. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas. 2001.

EPPS, Patience. Historical linguistics and socio-cultural reconstruction. In: Bower, Claire; Bethwyn Evans (eds.). *The Routledge Handbook of Historical Linguistics*, p. 579–97. Abingdon/Nova Iorque: Routledge, 2015.

EPPS, Patience. South American Languages. In: Bower, Claire; Patience Epps; Jane Hill; Patrick McConvell (comps.). *Languages of Hunter-Gatherers and their Neighbors: Database*, 2020. <<https://huntergatherer.la.utexas.edu>>, acesso em 11/08/2020.

FARGETTI, Cristina Martins. *Estudo fonológico e morfossintático da língua Juruna*. 2001. 317 f. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas. 2001.

FARGETTI, Cristina Martins; RODRIGUES, Carmen L. Reis. Consoantes do xipaya e do juruna – uma comparação em busca do proto-sistema. *Alfa*, São Paulo, v. 52, n. 2, p. 535–63, 2008.

FARIAS, Gabriel Soares. *Kuruaya e Xipaya: influência de contato ou pura relação genética?* 2017. 14 f. Relatório (Iniciação Científica) – Universidade Federal do Pará. 2017. <<https://docplayer.com.br/57111282-Universidade-federal-do-para-pro-reitoria-de-pesquisa-e-pos-graduacao-departamento-de-pesquisa.html>>, acesso em 15/08/2020.

FAUSTO, Carlos. Fragmentos de história e cultura tupinambá: da etnologia como instrumento crítico de conhecimento etno-histórico. In: CARNEIRO DA CUNHA, Manuela (org.). *História dos índios no Brasil*. 2ª ed. (pp. 381–96). São Paulo: Companhia das Letras, 1998.

FERREIRA, Aurélio Buarque de Hollanda. *Novo Dicionário da Língua Portuguesa*. 2ª ed. Rio de Janeiro: Nova Fronteira, 1986.

FONSECA, José Pinto da. Cópia da carta que o alferes José Pinto da Fonseca escreveu ao Exm. General de Goyazes, dando-lhe conta do descobrimento de duas nações de índios, dirigida do sítio onde portou. *Revista Trimensal de Historia e Geographia, ou Jornal do Instituto Historico e Geographico Brasileiro*, Rio de Janeiro, v. 8, p. 376–9, 1867 [1775].

GABAS JR., Nilson. Lingüística histórica. In: MUSSALIM, Fernanda; BENTES, Anna Christina (orgs.). *Introdução à lingüística: domínios e fronteiras*, v. 1 (p. 77–103). São Paulo: Cortez, 2001.

GALLAIS, Étienne-Marie. *O apóstolo do Araguaia. Frei Gil Vilanova, missionário dominicano*. Conceição do Araguaia: Prelazia Dominicana de Conceição do Araguaia, 1942.

GALLOIS, Dominique. *Contribuição ao estudo do povoamento indígena da Guiana brasileira, um caso específico: os Waiãpi*. 1980. Dissertação (Mestrado em Antropologia Social) – Departamento de Antropologia, Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, 1980.

GONZÁLEZ, Hebe. Una aproximación a la fonología del tapiete (Tupí-Guaraní). *LIAMES*, Campinas, v. 8, p. 7–43, 2008.

GORDON, César. *Economia selvagem: ritual e mercadoria entre os índios Xikrin-Mebêngokre*. São Paulo: UNESP, ISA; Rio de Janeiro: NUTI, 2006.

GRENAND, Pierre. *Ainsi parlaient nos ancêtres: essai d'ethno-histoire wayãpi*. Paris: ORSTOM, 1982.

HARRISON, Carl; HARRISON, Carola. *Dicionário Guajajara-Português*. Anápolis: Associação Internacional de Linguística SIL – Brasil, 2013.

IVO, Ivana Pereira. *Características fonéticas e fonologia do Guarani no Brasil*. 2018. 293 f. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas. 2018.

JAVAÉ, Ricardo Têwaxi; RODRIGUES, Patrícia de Mendonça. Navegando pelo Rio Javaés: uma apreciação a dois sobre a toponímia histórica javaé. *Hawò*, Goiânia, v. 1, p. 2–43, 2020.

KAKUMASU, James Y.; KAKUMASU, Kiyoko. *Dicionário por tópicos Kaapor-português*. Cuiabá: Associação Internacional de Linguística SIL – Brasil, 2007.

KARAJÁ, Hatawaki; KARAJÁ, Juanahu; KARAJÁ, Leandro; KARAJÁ, Teribre; KARAJÁ, Wadoi; KARAJÁ, Woubedu; OLIVEIRA, Cristiane; WHAN, Chang. *Dicionário enciclopédico Inyrybè/Karajá-português brasileiro*. Rio de Janeiro: Museu do Índio, 2013. <http://prodoclin.museudoindio.gov.br/images/conteudo/Karaja/produtos_pesquisadores/Karaj%C3%A1_Iny_Lexicon.pdf>, acesso em 18/08/2020.

KARAJÁ, Loiwa Damazia. Timybo iny rarybèmyhÿre tasÿ relykymyhyre Iny Kyri. In: LIMA, Nei Clara de; LEITÃO, Rosani Moreira (orgs.). *Iny Tkylysinamy Rybèna: arte Iny Karajá, patrimônio cultural do Brasil*, p. 41–6. Goiânia: IPHAN-GO, 2019.

KRAUSE, Fritz. *In den Wildnissen Brasiliens: Bericht und Ergebnisse der Leipziger Araguay-Expedition 1908*. Leipzig: R. Voigtländer's Verlag, 1911.

LEA, Vanessa. *Riquezas intangíveis de pessoas partiveis. Os Mëbêngókre (Kayapó) do Brasil Central*. São Paulo: Editora da Universidade de São Paulo, 2012.

LEITE, Yonne. Resenha de *Material lingüístico ye*, de Johannes Wilbert. *Estudos Lingüísticos (Revista Brasileira de Lingüística Teórica e Aplicada)*, São Paulo, v. 1, n. 2, p. 82–4, 1966.

LIMA, Daniela Batista de. *Transformações, xamanismo e guerra entre os Kajkwakratxi (Tapayuna)*. 2019. 516 f. Tese (Doutorado em Antropologia Social) – Departamento de Antropologia, Instituto de Ciências Sociais, Universidade de Brasília. 2019.

LIMA, Nei Clara de. Iny bðèdkÿnana: raritxamy ramyhÿre Bèrakuhekÿ-di. In: LIMA, Nei Clara de; LEITÃO, Rosani Moreira (orgs.). *Iny Tkylysinamy Rybèna: arte Iny Karajá, patrimônio cultural do Brasil*, p. 29–39. Goiânia: IPHAN-GO, 2019a.

LIMA, Nei Clara de. Wèriri boho. In: LIMA, Nei Clara de; LEITÃO, Rosani Moreira (orgs.). *Iny Tkylysinamy Rybèna: arte Iny Karajá, patrimônio cultural do Brasil*, p. 123–4. Goiânia: IPHAN-GO, 2019b.

LIMA, Suzi Oliveira de. *The grammar of individuation and counting*. 2014. xvi + 229 f. Tese (Doutorado em Linguística) – Department of Linguistics, University of Massachusetts Amherst. 2014.

LOPES, Jorge Domingues. *Uma interface da documentação linguística e modelos lexicográficos para línguas indígenas brasileiras: uma proposta para o Suruí-Aikewára*. 2014. 599 f. Tese (Doutorado em Linguística) – Departamento de Linguística, Português e Línguas Clássicas, Instituto de Letras, Universidade de Brasília. 2014.

LOURENÇO, Sonia Regina. *Brincadeiras de Aruanã: performances, mito, música e dança entre os Javaé da ilha do Bananal (TO)*. 2009. 518 f. Tese (Doutorado em Antropologia Social) – Universidade Federal de Santa Catarina. 2009.

MAILHAMMER, Robert. Towards a framework of contact etymology. In: MAILHAMMER, Robert (ed.). *Lexical and structural etymology*, p. 9–31. Berlin: Mouton de Gruyter, 2013.

MAILHAMMER, Robert. Etymology. In: BOWERN, Claire; EVANS, Bethwyn (eds.). *The Routledge Handbook of Historical Linguistics*, p. 423–41. Abingdon/Nova Iorque: Routledge, 2015.

MELLO, Antonio Augusto Souza. Genetic affiliation of the language of the Indians Auré and Aurá. *Opción*, Maracaibo, v. 19, p. 67–81, 1996.

MELLO, Antonio Augusto Souza. *Estudo histórico da família lingüística Tupi-Guarani: aspectos fonológicos e lexicais*. 2000. vii + 286 f. Tese (Doutorado em Linguística) – Departamento de Lingüística e Língua Vernácula, Instituto de Comunicação e Expressão, Universidade Federal de Santa Catarina. 2000.

MÉTRAUX, Alfred. Migrations historiques de Tupi-Guarani. *Journal de la Société des américanistes de Paris, nouvelle série*, Paris, v. 29, n. 1–45, 1927.

MOORE, Denny; STORTO, Luciana. As línguas indígenas e a pré-história. In: PENA, Sérgio D. J. (org.). *Homo brasilis*, p. 63–82. São Paulo: FUNPEC, 2002.

MORIMÃ, Álvaro. *Vocabulário Apiaká*. Nova Esperança, manuscrito, 1984.

MÜLLER, Regina Polo; SILVA, Fabíola Andréa. Asurini do Xingu. In: *Enciclopédia dos Povos Indígenas do Brasil*. São Paulo: Instituto Socioambiental, 2018. <https://pib.socioambiental.org/pt/Povo:Asurini_do_Xingu>, acesso em 17/08/2020.

NAVARRO, Eduardo de Almeida. *Dicionário tupi antigo: a língua indígena clássica do Brasil*. Vocabulário português-tupi e dicionário tupi-português, tupinismos no português do Brasil, etimologias de topônimos e antropônimos de origem tupi. São Paulo: Global, 2015.

NEME, Mário. *Dados para a história dos índios Caiapó*. Anais do Museu Paulista, São Paulo, v. 23, p. 101–47, 1969.

NIKULIN, Andrey. Contacto de lenguas en la Chiquitanía. *Revista Brasileira de Línguas Indígenas*, Macapá, v. 2, n. 2, p. 5–30, 2019a.

NIKULIN, Andrey. Annotated Swadesh wordlists for the Jê group (Macro-Jê family). 2019b. In: STAROSTIN, George (ed.) *The Global Lexicostatistical Database*. Moscou: Higher School of Economics; Santa Fé: Santa Fe Institute, 2011–2019. <<https://starling.rinet.ru/cgi-bin/response.cgi?root=new100&morpho=0&basename=new100\zhejee>>, acesso em 11/08/2020.

NIKULIN, Andrey. *Proto-Macro-Jê: um estudo reconstrutivo*. 2020. xxiv + 571 f. Tese (Doutorado em Linguística) – Departamento de Linguística, Português e Línguas Clássicas, Instituto de Letras, Universidade de Brasília. 2020.

NIKULIN, Andrey; CARVALHO, Fernando O. de. Prehistoria de las lenguas y familias lingüísticas del Gran Chaco, de la meseta brasileña y cercanías: propuesta de base de datos léxicos y resultados preliminares. In: REGÚNAGA, María Alejandra; SPINELLI, Silvia Andrea; ORDEN, María Emilia (comps.). *IV Encuentro de Lenguas Indígenas Americanas (ELIA)*: libro de actas (p. 545–60). Sanra Rosa: Universidad Nacional de La Pampa, 2018.

NIKULIN, Andrey; CARVALHO, Fernando O. de. Estudos diacrônicos de línguas indígenas brasileiras: um panorama. *Macabéa – Revista Eletrônica do Netlli*, Crato, v. 8, n. 2, p. 255–305, jul./dez. 2019.

NIKULIN, Andrey; SALANOVA, Andrés Pablo. Northern Jê verb morphology and the reconstruction of finiteness alternations. *International Journal of American Linguistics*, Chicago, v. 85, n. 4, p. 533–67, 2019.

NIMUENDAJÚ, Curt. *The Eastern Timbira*. Berkeley e Los Angeles: University of California Press, 1946.

NIMUENDAJÚ, Curt. *Mapa etno-histórico do Brasil e regiões adjacentes*. Instituto do Patrimônio Histórico e Artístico Nacional, Instituto Brasileiro de Geografia e Estatística, 2ª ed. <<http://portal.iphan.gov.br/uploads/publicacao/MapaEtnoHistorico2ed2017.pdf>>. Brasília: IPHAN/IBGE, 2017.

NONATO, Rafael. *Clause chaining, switch reference and coordination*. 2014. 152 f. Tese (Doutorado em Linguística) – Department of Linguistics and Philosophy, Massachusetts Institute of Technology. 2014.

NONATO, Rafael; SUYÁ, Jamtô; SUYÁ, Kawiri. *Dicionário Kĩsêdjê-Português*. Rio de Janeiro: Museu do Índio, 2012. <prododim.museudoindio.gov.br/images/contudo/kisedje/dicionario_kisedje1.pdf>, acesso em 18/08/2020.

NUNES, Eduardo Soares. *Transformações Karajá: os “antigos” e o “pessoal de hoje” no mundo dos brancos*. 2016. vi + 609 f. Tese (Doutorado em Antropologia Social) – Departamento de Antropologia, Universidade de Brasília. 2016.

OLIVEIRA, Caroline Pereira de. Educação indígena: ensino de língua étnica, metodologia intercultural e algumas reflexões. *Diadorim*, Rio de Janeiro, v. 6, p. 87–101, 2009.

OLIVEIRA, Christiane Cunha de. *The language of the Apinajé people of Central Brazil*. 2005. xiv + 430 f. Tese (Doutorado em Linguística) – Department of Linguistics, University of Oregon. 2005.

OLIVEIRA, Raquel Palmeira de. *Aspectos fonéticos e fonológicos da língua Karajá-Xambioá*. 2018. 35 f. Monografia (Licenciatura em Letras) – Universidade Federal do Tocantins (*campus* de Araguaína). 2018.

ORSI, Viviane. *Metáforas do universo lexical português e italiano das zonas erógenas: ânus, nádegas, pênis, seios, testículos e vulva*. 2009. 225 f. Tese (Doutorado em Estudos Linguísticos) – Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista (*campus* de São José do Rio Preto). 2009.

PALACÍN, Pe. Luiz. *Goiás (1722/1822): estrutura e conjuntura numa capitania de Minas*. Goiânia: Departamento Estadual de Cultura, 1972.

PAULA, Eunice Dias de. *Eventos de fala entre os Apyãwa (Tapirapé) na perspectiva da etnossintaxe: singularidades em textos orais e escritos*. 2012. 268 f. Tese (Doutorado em Letras e Linguística) – Faculdade de Letras, Universidade Federal de Goiás.

PEREZ FELIPIM, Adriana; QUEDA, Oriowaldo. O sistema agrícola Guarani Mbyá e seus cultivares de milho: um estudo de caso. *Interciência*, Caracas, v. 30, n. 3, p. 143–50, mar. 2005.

PICANÇO, Gessiane Lobato. *A fonologia diacrônica do Proto-Mundurukú (Tupí)*. Curitiba: Appris, 2019.

RAMIREZ, Henri; VEGINI, Valdir; FRANÇA, Maria Cristina Victorino de. O warázu do Guaporé (tupi-guarani): primeira descrição linguística. *LIAMES*, Campinas, v. 17, n. 2, p. 411–506, jul./dic. 2017.

REIS SILVA, Maria Amélia; SALANOVA, Andrés Pablo. A assessoria linguística nos projetos escolares indígenas: o caso da formação de professores mēbêngôkre. In: SILVA, Aracy Lopes da; FERREIRA, M. Leal (orgs.), *Antropologia, história e educação*, p. 331–59. São Paulo: Global, 2001.

RIBEIRO, Eduardo Rivail. Empréstimos Tupí-Guaraní em Karajá. *Revista do Museu Antropológico*, Goiânia, v. 5–6, n. 1, p. 75–100, jan./dez. 2001/2002.

RIBEIRO, Eduardo Rivail. *A grammar of Karajá*. 2012. 292 f. Tese (Doutorado em Linguística) – Department of Linguistics, University of Chicago. 2012.

RODRIGUES, Patrícia de Mendonça. *A caminhada de Tanyxiwè: uma teoria Javaé da História*. 2008. xix + 933 f. Tese (Doutorado em Antropologia) – Department of Anthropology, University of Chicago. 2008.

RUMSEY, Alan. *A comparative lexicon and glottochronology of some Ge languages*. Manuscrito, Universidade de Chicago, 1971.

SAARES RANTA, Tiina; DÍAZ MAQUERA, Rufino; HINOJOSA ROMÁN, Magaly. Educación indígena originaria campesina: perspectivas de la educación intracultural. La Paz: Fundación PIEB, 2011.

SALA, Rv. P^o. Antonio Maria Sala. Ensaio de grammatica Kaiapó. *Lingua dos Índios Kaiapós Brasil Central*. *Revista do Museu Paulista*, São Paulo, v. 12, p. 395–429, 1920.

SALANOVA, Andrés Pablo. *A nasalidade em Mebengokre e Apinayé: o limite do vozeamento soante*. 2001. viii + 93 f. Dissertação (Mestrado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas. 2001.

SALANOVA, Andrés Pablo. A flexão de terceira pessoa nas línguas Jê. *LIAMES*, Campinas, v. 11, p. 75–114, 2011.

SALANOVA, Andrés Pablo. A estrutura do sintagma nominal em Mebengokre. In: GOMES, Dionei Moreira; QUEIXALÓS, Francesc (orgs.). *O sintagma nominal em línguas amazônicas* (p. 203–39). Campinas: Pontes, 2017.

SALANOVA, Andrés Pablo; LEA, Vanessa. *A linguistic analysis of Mëbêngôkre triadic terms*. Comunicação a ser apresentada no *Amazônicas 8*. Universidade Federal de Goiás, em preparação.

SANTOS, Bárbara Heliodora Lemos de Pinheiro. *Empréstimos lexicais do português para a Língua Geral: século XVI ao XXI*. 2020. 109 f. Dissertação (Mestrado em Linguística) – Departamento de Linguística, Português e Línguas Clássicas, Instituto de Letras, Universidade de Brasília. 2020.

SÃO FÉLIX DO XINGU. *Lei n.º. 571/2019, de 13 de novembro de 2019*. Dispõe sobre a cooficialização da língua Mebêngôkre (Kayapó) no Município de São Félix do Xingu–PA e o incentivo da disciplina de estudo da língua no currículo escolar, nas escolas da rede municipal de ensino localizadas nas regiões em que predominam a população descendente no município. São Félix do Xingu: Câmara Municipal, 2019.

SILVA, Cristiane Oliveira da. *O processamento da dêixis em Karajá*. 2010. 122 f. Dissertação (Mestrado em Linguística) – Faculdade de Letras, Universidade Federal do Rio de Janeiro. 2010.

SOUSA, Suseile Andrade. *Contribuições para a história linguística do subgrupo Tupí-Guaraní Norte-Amazônico, com ênfase na língua Zo'ê*. 2013. Dissertação (Mestrado em Linguística) – Departamento de Linguística, Português e Línguas Clássicas, Instituto de Letras, Universidade de Brasília. 2013.

SOUSA, Suseile Andrade. *Elementos do léxico e da gramática Apiaká (subramo VI da família linguística tupí-guaraní)*. 2017. 204 f. Tese (Doutorado em Linguística) – Departamento de Linguística, Português e Línguas Clássicas, Instituto de Letras, Universidade de Brasília. 2017.

SOUZA, Shirley Dias Cardoso de. *The Morphology of Nouns in the Ugoronmo Language (Arara of Pará)*. 2010. xiii + 125 f. Dissertação (Mestrado em Linguística) – University of North Dakota. 2010.

STEINEN, Karl von den. *Durch Central-Brasilien: Expedition zur Erforschung des Schingú im Jahre 1884*. Leipzig: F. A. Brockhaus, 1886.

STOUT, Mickey; THOMSON, Ruth. Fonêmica Txukuhamêi (Kayapó). In: BRIDGEMAN, Loraine Irene (ed.). *Série Lingüística, no. 3*, p. 153–76. Brasília: Summer Institute of Linguistics, 1974.

STRADELLI, Ermano. Vocabulários da língua geral português-nheêngatú e nheêngatú-português, precedidos de um esboço de Grammatica nheênga-umbuê-sáua mirî e seguidos de contos em língua geral nheêngatú poranduu. *Revista do Instituto Historico e Geographico Brasileiro*, Rio de Janeiro, v. 104, n. 158, p. 9–768, 1929.

TAPIRAPÉ, Diego Manawari. Lòrilòri. In: LIMA, Nei Clara de; LEITÃO, Rosani Moreira (orgs.). *Iny Tkylysinamy Rybèna: arte Iny Karajá, patrimônio cultural do Brasil*, p. 128–32. Goiânia: IPHAN-GO, 2019.

TORAL, André Amaral de. *Cosmologia e sociedade Karajá*. 1992. 280 f. Dissertação (Mestrado em Antropologia Social) – Museu Nacional, Universidade Federal do Rio de Janeiro, 1992.

TURNER, Terence. Os Mebengokre Kayapó: história e mudança social. De comunidades autônomas para a coexistência interétnica. In: Carneiro da Cunha, Manuela (org.). *História dos índios no Brasil*. 2ª ed. (p. 311–38). São Paulo: Companhia das Letras, 1998.

URBAN, Greg. A história da cultura brasileira segundo as línguas nativas. In: Carneiro da Cunha, Manuela (org.). *História dos índios no Brasil*. 2ª ed. (p. 87–102). São Paulo: Companhia das Letras, 1998.

VASCONCELOS, Eduardo Alves. *Investigando a hipótese Cayapó do Sul–Pinará*. 2013. 291 f. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas. 2013.

VASILYEV, Mikhail; SAENKO, Mikhail. K voprosu o tochnosti glottoxronologii: datirovanie jazykovoj divergencii po dannym romanskix jazykov. [How accurate can

glottochronology be? Dating language divergence on the basis of Romance data.] *Journal of Language Relationship*, Moscou/Piscataway, v. 15, n. 2, p. 114–35, 2017.

VERSWIJVER, Gustaaf. Intertribal relations between the Juruna and the Kayapó Indians (1850–1920). *Jahrbuch des Museums für Völkerkunde zu Leipzig*, Berlim, v. 34, p. 305–15, 1982.

VERSWIJVER, Gustaaf. *The club-fighters of the Amazon: warfare among the Kaiapo Indians of Central Brazil*. Ghent: Rijksuniversiteit te Gent, 1992. (Werken uitgegeven door de Faculteit van de Letteren en Wijsbegeerte, v. 179.)

VIVEIROS DE CASTRO, Eduardo. *Araweté. Os deuses canibais*. Rio de Janeiro: Jorge Zahar, 1986.

VIDAL, Lux Boelitz. *Morte e vida de uma sociedade indígena brasileira. Os Kayapó-Xikrin do Rio Cateté*. São Paulo: Hucitec/EdUSP, 1977.

WILBERT, Johannes. *Material lingüístico ye*. Caracas: Sucre, 1964. (Monografías, n. 10.).

THE STORY TOLD BY THE MĒBÊNGÔKRE LEXICON

ABSTRACT

The goal of this paper is to identify non-native elements in the lexicon of MĒbêngôkre (an indigenous language of the Jê group, Macro-Jê family), spoken by the Kayapó and Xikrin nations in the state of Pará as well by the Kayapó Mĕtyktire (Txucarramãe) in northern Mato Grosso. New loan etymologies are proposed for a number of MĒbêngôkre words, including waxi ‘fishing line’, benorã ‘peacock bass’, awo ‘tree sp.; bark boat’, ixé ‘mirror’ (from Xambioá); môtôbi’y ‘peanut’, xãn ‘cat’, mokà ‘kind of basket’, xoko ‘tiger heron’ (from Língua Geral Amazônica), karaxu ‘spoon’ (from Yudjá), among others; in addition, several MĒbêngôkre loans in other languages are identified. We conclude by proposing a synthesis of the migration history of the Northern Jê peoples based both on our novel findings and on previous scholarship.

Keywords: MĒbêngôkre, language contact, lexical borrowings, Jê languages.

Recebido em 23/08/2020.

Aprovado em 21/09/2020.