



LUND
UNIVERSITY

Centro de Línguas e Literatura (SOL)

Divisão de Linguística

Neologismos na revitalização das línguas indígenas no Nordeste do Brasil

Tradução do título original em inglês: Neologisms in Revitalizing Indigenous Languages in Northeast Brazil

Göran Jarl

Orientadora: Karolin Obert

Tese de bacharelado em Linguística

Primavera de 2026

Índice

<i>Resumo, Abstract</i>	3
1. <i>Introdução</i>	4
2. <i>Antecedentes</i>	7
2.1 <i>Línguas indígenas no Brasil</i>	7
2.1.1 <i>Breve história</i>	7
2.1.2 <i>Documentação linguística</i>	8
2.3.3 <i>Famílias linguísticas indígenas</i>	9
2.1.4 <i>Escala de risco de extinção das línguas</i>	11
2.1 <i>Revitalização das línguas – O panorama global</i>	11
2.2. <i>Neologismos</i>	16
2.2.1. <i>Definição de neologismos</i>	16
2.2.2. <i>Classificação do processo de cunhagem</i>	17
2.2.3 <i>Alguns aspectos sociolinguísticos dos neologismos</i>	20
3. <i>Contexto para a revitalização no Nordeste brasileiro</i>	22
3.1 <i>Patxohã: povo e língua</i>	24
3.2 <i>Kariri-Xocó: povo e língua</i>	27
3.3 <i>Fulni-ô, Yathê: Povo e língua</i>	29
3.4 <i>Krenak, Borum: Pessoas e Língua</i>	31
3.5 <i>Xukuru: Povo e Língua</i>	34
3.6 <i>Sistemas numéricos na língua</i>	36
4. <i>Métodos</i>	38
5. <i>Resultados</i>	40
5.1 <i>Patxohã</i>	40
5,2 <i>Kariri-Xocó</i>	43
5.2 <i>Yathê</i>	46
5.3 <i>Borum</i>	48
5,5 <i>Xukuru</i>	49
6. <i>Discussão</i>	51
7. <i>Conclusões</i>	54
8. <i>Perspectivas</i>	56
9. <i>Referências</i>	58
10. <i>Apêndice</i>	
A. <i>Fontes para neologismos</i>	66
B. <i>Distribuição de neologismos por processo de cunhagem. Quantidades e porcentagens</i>	67
C. <i>Abreviações nas glosas</i>	68

Lista de figuras

Figura 1 O mapa do Brasil mostra a linha divisória a leste da qual a maioria das línguas indígenas está inativa. (Mirim, s.d.)	5
Figura 2 Família linguística Macro-Jê (Instituto Socioambiental. (s.d.)).....	10
Figura 3 Família linguística Tupi (Instituto Socioambiental. (s.d.))	10
Figura 4 Escala de grau de ameaça de línguas. Tradução de UNESCO (2003)	11
Figura 5 Modelo Chickasaw para revitalização da língua (Fitzgerald & Hinson 2013, p. 59).....	14
Figura 6 Região Nordeste do Brasil com as comunidades indígenas cujas línguas são estudadas neste artigo	22
Figura 7 Distribuição dos processos de cunhagem de neologismos	51

Lista de tabelas

Tabela 1 Lista de palavras Yathê (Branner 1887, P.330)	9
Tabela 2 Parâmetros sociolinguísticos para as línguas estudadas	23
Tabela 3 Sistema de determinantes definidos em patxohã (Bomfim & Lima, 2020).....	26
Tabela 4 Sufixos verbais para marcar o tempo verbal em Patxohã (Braz da Conceição, 2016, pp. 25-26).....	26
Tabela 5 Sufixos derivacionais em substantivos em Patxohã (Braz da Conceição, 2016, p. 27).....	26
Tabela 6 Lista de neologismos em Patxohã.....	40
Tabela 7 Numeros em Patxohã	42
Tabela 8 Lista de neologismos em Kariri-Xocó	43
Tabela 9 Numerais em Kariri-Xocó (Silva 2021)	44
Tabela 10 Lista de neologismos em Yathê	46
Tabela 11 Numerais em Yathê (Lapenda, 1968)	47
Tabela 12 Números ordinais em Yathê (Lapenda, 1968)	48
tabela 13 Lista de neologismos em Borum	48
Tabela 14 Lista de neologismos em Xukuru	49

Nota do autor: A tradução deste texto para o português contou com o apoio da ferramenta de IA DeepL, com revisão integral do autor.

Neologismos na revitalização das línguas indígenas no Nordeste do Brasil

Resumo:

Este artigo compara a formação de neologismos em cinco línguas indígenas do Nordeste do Brasil que estão atualmente em processo de revitalização ou recuperação pós-dormância. Com foco na criação de substantivos e numerais, o estudo categoriza os neologismos de acordo com processos morfológicos (por exemplo, composição, derivação, empréstimo) e semânticos (por exemplo, extensão) e avalia sua distribuição entre línguas com diferentes níveis de ameaça de extinção. A análise mostra que, embora a composição e o empréstimo sejam estratégias amplamente preferidas, o tipo de empréstimo e outras estratégias de criação variam substancialmente e se correlacionam com a história sociolinguística, em vez de restrições tipológicas. Línguas com transmissão ininterrupta apresentam perfis de neologismos mais próximos dos de línguas totalmente vivas, enquanto línguas com interrupção prolongada dependem mais de formas emprestadas ou criadas artificialmente. Essas descobertas destacam a importância do contexto sociolinguístico na formação da inovação lexical em línguas revitalizadas.

Palavras-chave: Revitalização; Neologismo; Línguas Indígenas; Nordeste Brasileiro; Cunjagem de Palavras; Substantivo; Numeral

1. Introdução

Em todo o mundo, a revitalização de línguas extintas ou quase extintas durante o século XX tem estado cada vez mais ligada a um movimento mais amplo de retomada de terras e cultura. Inicialmente, os esforços concentravam-se frequentemente na documentação do léxico e da gramática antes que os últimos falantes falecessem. Isso evoluiu para um movimento em que as comunidades recuperam ativamente a língua de seus ancestrais por meio de escolas, oficinas comunitárias e materiais didáticos modernos. Paralelamente, cada vez mais indivíduos se identificam como membros dessas comunidades indígenas. Sucessos bem conhecidos — como a revitalização do Hebraico como língua falada, bem como do Maori e do Havaiano — inspiraram iniciativas semelhantes em todo o mundo. Hoje, a tecnologia digital e as plataformas de comunicação, juntamente com a pesquisa linguística, desempenham papéis centrais em muitas dessas iniciativas, permitindo que as comunidades não apenas preservem o patrimônio linguístico, mas também o restabeleçam como parte viva da vida cotidiana.

Hoje, esse movimento também é uma realidade entre a população indígena do Brasil, que está recuperando as línguas de seus ancestrais. Juntamente com o aumento da autoidentificação dos indígenas, muitas das comunidades indígenas, que foram declaradas extintas, estão agora recuperando suas terras e sua cultura. O processo de recuperação de uma língua, muitas vezes adormecida ou extinta, tem vários efeitos positivos para a comunidade em questão (por exemplo, apoiar a educação comunitária), bem como para os indivíduos (por exemplo, reafirmar a identidade étnica).

Esses processos em andamento são o resultado da dizimação da maioria das línguas indígenas no Brasil, que se extinguiram ou ficaram adormecidas devido à brutalidade do processo de colonização (Rodrigues, 2002). Isso se aplica à maioria das línguas faladas a leste de uma linha entre São Luiz, no Maranhão, e Porto Alegre, no Rio Grande do Sul (ver Figura 1). Essa área corresponde a aproximadamente 25% do território brasileiro total. As comunidades indígenas dessa região foram as que mais sofreram, em alguns aspectos, pois foi ali que a colonização portuguesa se iniciou e a maioria das políticas, tanto da coroa portuguesa quanto da Igreja Católica, foram implementadas de forma mais severa e sistemática.



FIGURA 1 O MAPA DO BRASIL MOSTRA A LINHA DIVISÓRIA A LESTE DA QUAL A MAIORIA DAS LÍNGUAS INDÍGENAS ESTÁ INATIVA. (MIRIM, S.D.)

O processo de colonização fragmentou deliberadamente as comunidades indígenas e as reagrupou em missões e locais de trabalho. Diferentes grupos indígenas foram reunidos com africanos escravizados e pessoas de ascendência europeia, o que torna as noções anteriores de populações indígenas ou afrodescendentes “puras” menos relevantes no Brasil do que em algumas outras ex-colônias, onde as comunidades foram mantidas separadas e, assim, preservaram suas culturas distintas (Telles, 2019).

Esse processo também se reflete na língua, pois as tradições linguísticas foram fundidas e remodeladas durante períodos de supressão cultural e extinção. Declarações formais da extinção de comunidades indígenas eram frequentemente usadas como pretexto para distribuir terras a colonos brancos, e casamentos interétnicos eram apresentados como evidência de que essas comunidades haviam deixado de existir. Um exemplo dessa extinção administrativa é descrito em Silva (2007), onde as autoridades locais do município de Pesqueira (município em que se situava a comunidade Xukuru) informaram à província que não existiam mais índios Xukuru, apenas *caboclos*¹, ao mesmo tempo em que informavam sobre a necessidade de expansão do município e solicitavam a concessão da *Terra Indígena*. Em resposta a esses pedidos persistentes, em 1879, o Governo Imperial decretou oficialmente a extinção da Missão Cimbres em Pesqueira. Os inquilinos, em sua maioria ilegais, que viviam no Território Indígena foram favorecidos, muitos deles vereadores e proprietários de terras que haviam invadido as terras Xukuru e eram membros da elite local com influência considerável na política nacional e regional.

Uma parte importante da revitalização das línguas é a criação de novas palavras, ou seja, a cunhagem, para incluir as mudanças na sociedade que ocorreram desde a extinção ou desde a última documentação da língua. Isso vale para línguas dormentes e extintas, mas muitas vezes também para aquelas que estão sujeitas a grave risco de extinção e declínio. Neste trabalho, vou me concentrar

¹ *Caboclo* é uma categoria social brasileira que se refere principalmente a pessoas de ascendência mista indígena e europeia e, de forma mais ampla, às populações rurais moldadas por vários graus de influência cultural indígena em diferentes regiões do Brasil.

nessa parte específica do processo de revitalização: a criação de novas palavras em algumas das línguas dessa zona linguisticamente amplamente dormente no Nordeste brasileiro (a leste da linha azul na Figura 1). Explorarei as semelhanças e diferenças dos processos de cunhagem de palavras e sua relação com os diferentes contextos de cada comunidade falante. Especificamente, questiono quais estratégias são empregadas para criar novas palavras nas línguas revitalizadas do Nordeste brasileiro e como essas estratégias se refletem nas características estruturais e culturais de cada língua.

Para isso, classificarei os neologismos em cada língua, identificarei os processos de cunhagem utilizados e examinarei possíveis correlações com o contexto sociolinguístico. As línguas abordadas encontram-se em diferentes fases de revitalização e foram escolhidas para ter amostras de uma ampla gama de níveis de extinção e tipos de processos de revitalização. As línguas escolhidas são Xukuru, Kariri-Xocô, Yathé, Borum e Patxohã, todas pertencentes à família linguística Macro-Jê ou isoladas. Essa escolha é interessante, pois permite abordar o processo de cunhagem de neologismos especificamente em diferentes modelos e fases de revitalização e pode, possivelmente, agregar novos conhecimentos úteis para novos projetos de revitalização. A criação de novas palavras em línguas em grande parte dormentes é muitas vezes uma questão de sobrevivência, não apenas para a língua, mas, em muitos casos, para a comunidade como tal.

Este trabalho está estruturado da seguinte forma: após abordar o contexto geral das línguas indígenas no Brasil e os fundamentos teóricos da revitalização e dos neologismos, começarei por descrever o contexto específico da revitalização no Nordeste do Brasil no que diz respeito ao contexto sociolinguístico, bem como alguns parâmetros tipológicos utilizados no meu trabalho de classificação para cada uma das línguas estudadas no capítulo 3. Em seguida, descrevo os métodos utilizados para a descoberta, codificação e classificação e os no capítulo 4. O capítulo 5 apresenta os resultados das classificações dos neologismos substantivos, bem como a estrutura escolhida para a criação de numerais em cada uma das línguas selecionadas. A discussão dos resultados e as correlações com o contexto sociolinguístico e as conclusões seguem-se nos capítulos 6 e 7. Este trabalho é concluído com uma perspectiva sobre como a revitalização futura nesta região poderia ser apoiada utilizando um modelo e uma estrutura desenvolvidos com base nas experiências da revitalização do Chickasaw (uma língua indígena de Oklahoma, EUA).

2. Antecedentes

2.1 Línguas indígenas no Brasil

2.1.1 Breve história

A história linguística do Brasil é uma interação entre colonização, migração e resistência dos povos indígenas. Na época da colonização portuguesa, estimava-se que entre 1.000 e 1.300 línguas indígenas distintas eram faladas em todo o território que hoje é o Brasil (Instituto Socioambiental, s.d.). Hoje, restam apenas cerca de 274 línguas indígenas (G20 Brasil, 2024). Essa perda dramática não é a mesma em todo o país. Especialmente no nordeste do Brasil, quase todas as línguas indígenas desapareceram, enquanto outras regiões — especialmente a Amazônia e o centro do Brasil — ainda apresentam uma diversidade linguística muito maior. Os povos indígenas das zonas costeiras — como os Tupinambás e os Potiguarás — foram os primeiros a sofrer deslocamentos em massa, escravidão e extensa mistura com europeus e africanos. Os primeiros missionários jesuítas impuseram uma *Língua Geral* (uma língua franca baseada no Tupi)² durante os séculos XVI e XVII. Em meados do século XVIII, muitas línguas indígenas foram proibidas, incluindo a *Língua Geral*, e o Português foi imposto como única língua oficial (Durazzo & Costa, 2020). Isso foi imposto pelos colonizadores e pela Igreja como um só. Como o Nordeste estava mais integrado à economia colonial, a homogeneização linguística teve um efeito mais forte e duradouro nessa região do que em regiões mais remotas. Paradoxalmente, grande parte do material que permanece escrito dessas línguas proibidas foi registrado e estruturado pela mesma organização, a Igreja Católica. Eles também traduziram partes da Bíblia, orações, hinos etc. para muitas dessas línguas e, em muitos casos, documentaram dicionários, gramáticas parciais etc. de várias línguas indígenas, que mais tarde se extinguíram ou ficaram inativas (mais sobre os diferentes status posteriormente). Essa dualidade, tanto participando do genocídio e deslocamento, etc. dos povos indígenas e, ao mesmo tempo, às vezes protegendo as culturas e línguas indígenas (embora de uma posição paternalista e colonialista, tendo a evangelização como objetivo final), pode explicar de certa forma as opiniões complexas e divididas que prevalecem hoje sobre, por exemplo, o padre José de Anchieta. Por um lado, ele fazia parte do sistema missionário jesuíta que visava converter os povos indígenas ao cristianismo, suprimindo rituais indígenas e estruturas sociais tradicionais; por outro lado, ele foi um dos poucos que aprendeu Tupi, protegeu vários grupos indígenas dos portugueses e, em alguns aspectos, é celebrado como um “protetor” dos indígenas. Muito provavelmente, Nunes Duarte Silveira (2018) se aproxima da realidade ao estudar a obra teatral de Anchieta quando afirma que sua motivação doutrinária residia na crença de que esses povos não tinham salvação e conhecimento de Cristo e, portanto, era necessário cristianizá-los, mesmo que, para isso, fosse necessário alterar de alguma forma sua realidade cultural e sua memória.

A bacia Amazônica — embora geograficamente vasta e desafiadora para o assentamento sustentável dos colonizadores — tornou-se linguisticamente interconectada a partir do século XVII. Principalmente as missões Jesuítas e as redes coloniais promoveram a disseminação do Nheengatu (Língua Geral Amazônica), uma língua franca baseada no Tupi que facilitou a comunicação entre diversos grupos indígenas, missionários, colonos e, mais tarde, também populações afrodescendentes. Em vez de permanecerem isoladas, muitas comunidades e as do Rio Negro e da região amazônica em geral

² Simplificado: Três línguas francas diferentes foram usadas em diferentes regiões e períodos: Século XVII: Tupinamba, mais tarde chamada Tupi (para a qual Padre Anchieta criou uma gramática (Anchieta, 1595); Século XVIII: Língua Geral Paulista, às vezes chamada de Tupi Austral (vulg.: a língua dos bandeirantes; pioneiros e aventureiros que exploravam Minas Gerais, Goiás, Mato Grosso e o Sul); Século XVIII: Língua Geral Amazônica, baseada no Nheengatu. As três línguas francas interagiram e se desenvolveram com base nas “necessidades” mercantis e missionárias (Rodrigues, 1986).

interagem por meio dessa língua comum, mesmo mantendo suas próprias identidades culturais e linguísticas distintas.

Em outras partes do Brasil — particularmente nas terras altas centrais e nas regiões sul — grupos como os Xavante, Kayapó e Guaraní mantiveram maior autonomia territorial e independência sociopolítica. Isso lhes permitiu manter suas próprias línguas na interação diária, na prática ritual e na transmissão intergeracional, mesmo com o aumento das pressões coloniais.

A região Nordeste é normalmente considerada como incluindo os seguintes estados: Alagoas, Bahia, Ceará, Maranhão, Paraíba, Pernambuco, Piauí, Rio Grande do Norte e Sergipe, e tem hoje uma população de cerca de 60 milhões, quase um terço do país. As condições socioeconômicas e institucionais também desempenharam um papel importante na extinção das línguas indígenas. No Nordeste, a urbanização, a assimilação e o enfraquecimento das comunidades indígenas autônomas levaram o Português a se tornar a língua da mobilidade social e da sobrevivência. Nas regiões onde os grupos indígenas mantiveram autonomia política ou territorial, suas línguas continuaram a servir como marcadores de identidade e como ferramentas práticas de comunicação diária (Storto, 2019). A perda da identidade cultural e do patrimônio está entre os efeitos mais duradouros da extinção dessas línguas e contribui para os esforços de revitalização das línguas.

Na década de 1970, movimentos indígenas foram iniciados, começando a valorizar as culturas e línguas indígenas. Para grande parte da população indígena, porém, a língua havia sido perdida e falar a língua passou a ser o marcador da indigenidade. A revitalização da língua e da cultura tornou-se, para vários desses grupos, parte da luta para recuperar a identidade como nação e indivíduo, juntamente com a recuperação da terra e da cultura. A definição de exatamente qual língua “revitalizar” cabe, no final das contas, a cada comunidade.

Devido à história colonial brutal do Brasil — durante a qual muitos indígenas se esconderam ou evitaram reconhecer suas origens (por exemplo, identificando-se como *caboclos* em vez de “índios” para escapar de mais discriminação ou pior; ver nota de rodapé, p. 4) — e devido aos extensos casamentos mistos entre diferentes grupos indígenas e a população não indígena, não é possível nem necessário estabelecer uma definição fixa da língua na qual os esforços de revitalização devem se basear. Como veremos mais adiante, tanto a “base” linguística quanto os próprios grupos étnicos são frequentemente compostos.

2.1.2 Documentação linguística

A maioria dos dados escritos disponíveis sobre as línguas que ficaram inativas durante o início da colonização da América do Sul foi coletada por missionários jesuítas, como o padre José de Anchieta, autor de Anchieta (1595). O objetivo de seus esforços era certamente facilitar a evangelização, e esses esforços continuaram mais tarde também pelos franciscanos e carmelitas.

No século XIX, vários exploradores-naturalistas europeus deram continuidade a esse empreendimento documental, normalmente combinando notas linguísticas com a redação de etnografias e a documentação da flora e da fauna. Ao contrário de seus antecessores missionários — cujo objetivo principal era a evangelização e a instrução catequética nas línguas locais —, essas figuras perseguiram objetivos científicos: registrar e classificar povos, línguas e ambientes naturais como parte de disciplinas comparativas emergentes. Entre os mais conhecidos estão Carl Friedrich Philipp von Martius (1794-1868), o príncipe Maximiliano zu Wied-Neuwied (1782-1867), Auguste de Saint-Hilaire (1779-1853) e Henry Walter Bates (1825-1892). Voltaremos ao príncipe Maximiliano mais adiante, pois ele documentou, entre outros, os Botocudos/Aimorés, os Puris, os Maxakalis e os Pataxós.

No início do século XX, Curt Nimuendajú fez um trabalho de documentação sobre as línguas Xukuru e elaborou uma lista inicial de 150 palavras, que foi complementada em 1944, quando Cícero Cavalcanti visitou os descendentes Xukuru dessa tribo. Nessa época, porém, os Xukuru haviam perdido tanto de sua língua que não conseguiam mais construir frases ou sentenças com as poucas palavras da língua nativa que lhes restavam (Hohenthal, 1954). A falta de documentação da língua Xukuru também coloca em dúvida muitas das palavras, se são Xukuru ou possivelmente Paratío, outra comunidade indígena frequentemente mencionada, mas não temos evidências reais de que os Paratío tivessem uma língua diferente.

O padre Mamiani, em 1699, escreveu uma gramática básica da língua Kariri, Mamiani (1877/1699), José Barbosa de Sá fez um trabalho mais descritivo durante o século XVIII, que foi posteriormente seguido por Friedrich Müller (1834–1898) e Karl von der Steinen (1855–1929).

Finalmente, no que diz respeito à língua *Yathé*, a referência mais antiga que identifiquei aparece em Branner (1887), que descreve um grupo conhecido pelos brasileiros como *Carnijós*, mas que se autodenomina *Förniö*. Branner (1887) relata uma breve conversa noturna com um membro idoso dessa comunidade e relata os detalhes linguísticos que conseguiu coletar, incluindo a seguinte lista inicial de palavras:

ever it is possible.	
{ Itó = my eyes.	{ Tehó = the hand.
{ Ató = another person's eyes.	{ Tehó tchá = the lower arm.
{ Tocáno = two eyes.	{ Tehá fuá = the whole arm.
{ Deretá = my nose.	{ Tehó túlll = all the fingers.
{ Aretá = the nose of another.	{ Tehó kũ = the little finger.
{ Dutchi = my mouth.	{ Tehó sã = the thumb.
{ Aotchi = another's mouth.	D'hó hó = the breast, bosom.
{ Jaxi = my tooth.	Iehitá = the belly.
{ Axí = another's tooth.	(Ieh in this word has the sound of the
{ Huatór = one.	German <i>feh</i> .)
{ Jaxí huatór = all my teeth, my	Sçi ku = food.
whole set of teeth.	Ôú tchí = meat.
{ Taláw = horse.	Kí tchí ùh = <i>farinha</i> .
{ Tskùh = cow.	Tökùh = salt.
{ Refeitlùh = cattle.	Ôyè = water.
Tóch = fire.	Ôyè téhn i tichina = give me some
Táthá = firewood.	water.

TABELA 1 LISTA DE PALAVRAS YATHÊ (BRANNER 1887, P.330)

2.1.3 Famílias linguísticas indígenas

A maioria das línguas indígenas em solo brasileiro pertence a uma das duas famílias linguísticas: Tupi e Macro-Jê, como pode ser visto nas Figuras 2 e 3 abaixo. Todas as línguas deste estudo pertencem à família Macro-Jê, exceto o Yathê, cuja classificação ainda é questionada. O Patxohã, por ser uma língua “nova”, não está classificado neste tipo de figura, mas deve ser colocado junto com o Maxakali e o Pataxó.

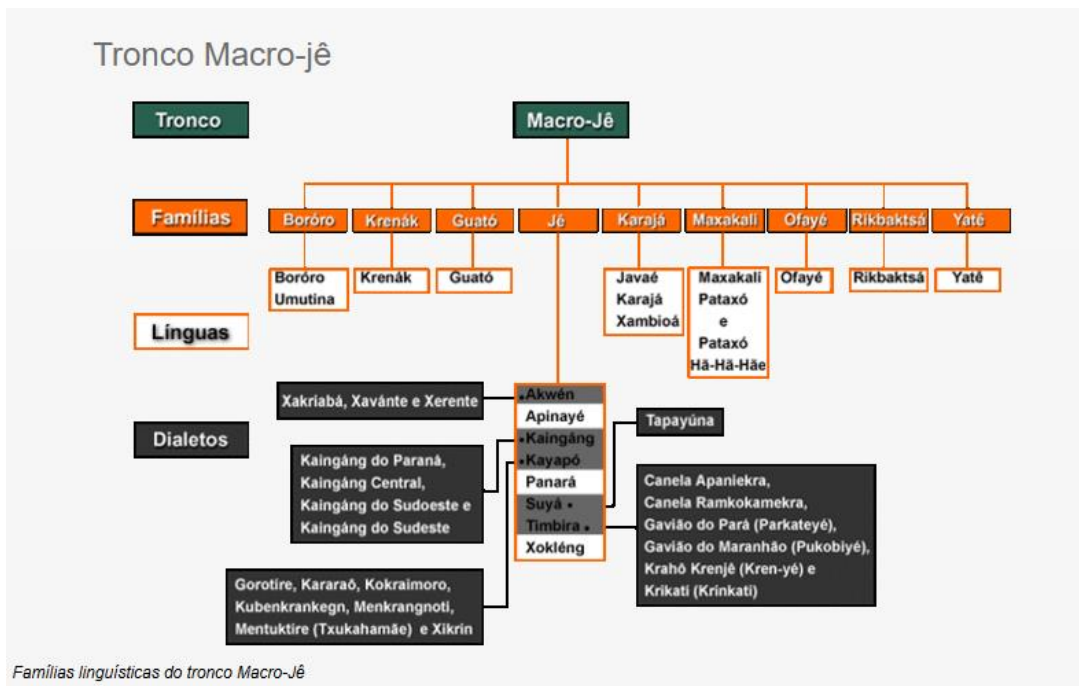


FIGURA 2 FAMÍLIA LINGUÍSTICA MACRO-JÊ (INSTITUTO SOCIOAMBIENTAL. (S.D.))

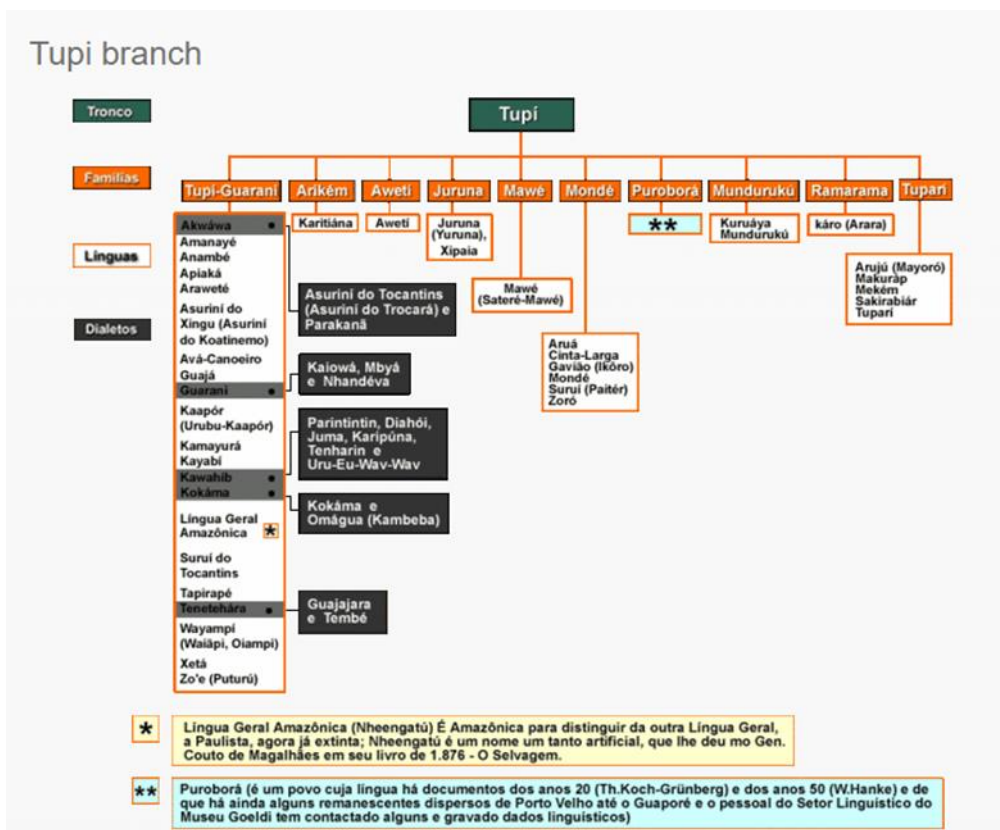


FIGURA 3 FAMÍLIA LINGUÍSTICA TUPI (INSTITUTO SOCIOAMBIENTAL. (S.D.))

Línguas em cerca de 20 outras famílias linguísticas, bem como oito isoladas, ainda são faladas no Brasil (Storto, 2019). As maiores dessas famílias (além do Tupi e do Macro-Jê) são Aruak, Karib, Nanambikwara e Pano. A maioria das línguas dessas famílias é falada principalmente no norte e noroeste do Brasil.

O Censo Populacional de 2022 mostra que existem 391 etnias, povos e grupos indígenas vivendo no Brasil (Brasil, 2018). As três etnias mais populosas são os Tikúna (74.061), os Kokama (64.327) e os Makuxí (53.446). Em 2022, havia 295 línguas indígenas faladas por povos indígenas com 2 anos de idade ou mais. As três línguas com mais falantes são: Tikúna (51.978), Guarani Kaiowá (38.658) e Guajajara (29.212). Em 2010, foram identificadas 274 línguas entre pessoas com 5 anos ou mais. Entre 2010 e 2022, o número de falantes de línguas indígenas entre indígenas com cinco anos ou mais aumentou de 293.853 para 433.980. No entanto, a porcentagem caiu de 37,35%, em 2010, para 28,51%, em 2022 (Instituto Brasileiro de Geografia e Estatística (IBGE), 2025).

2.1.4 Escala de risco de extinção das línguas

Muitas línguas indígenas, também no Brasil, estão em risco de extinção em algum grau. Uma categorização amplamente utilizada do status da língua é a Escala de grau de ameaça de línguas da UNESCO (2003), com seis níveis que vão de seguro até extinto/inativo, conforme mostrado abaixo na Figura 4 in translation.

Grau	Nível	Descrição
5	Segura	A língua é usada por todos os grupos etários, inclusive pelas crianças.
4	Insegura	A língua é usada por algumas crianças em todos os domínios; é usada por todas as crianças em alguns domínios.
3	Ameaçada	A língua é utilizada principalmente pelas gerações dos pais e gerações mais velhas.
2	Severamente ameaçada	A língua é usada principalmente pela geração dos avós e gerações mais velhas.
1	Criticamente ameaçada	A língua é conhecida por muito poucas pessoas, principalmente da geração dos bisavós.
0	Extinta	Não há falantes.

FIGURA 4 ESCALA DE GRAU DE AMEAÇA DE LINGUAS. TRADUÇÃO DE UNESCO (2003)

O Ethnologue usa uma escala diferente, a EGIDS (Escala Expandida de Interrupção Intergeracional Graduada), com 10 níveis diferentes. Ambas as escalas usam como fatores principais a transmissão intergeracional da língua e o número de falantes. A diferenciação entre línguas dormentes, extintas e esquecidas não é levada em consideração em nenhuma dessas escalas. Outro termo frequentemente usado neste contexto é moribundo, que significa uma língua que não é mais aprendida pelas crianças e é falada apenas pelas gerações mais velhas.

A seguir, utilizarei a Escala de Risco de grau de ameaça de línguas da UNESCO, em parte devido à agenda subjacente não transparente do Ethnologue, mas utilizarei o termo dormente em vez de extinto (Harrison, 2007), pois o termo é mais apropriado para as línguas deste estudo. O paralelo semântico entre a extinção de espécies e de línguas também é muito próximo, o que torna essa escolha de termo ainda mais necessária.

2.2 Revitalização das línguas – O panorama global

Em muitos casos, as práticas linguísticas coloniais tiveram um efeito destrutivo sobre a vitalidade das línguas locais em todo o mundo. “Muitos povos indígenas citam a cura como a principal razão para a revitalização. Eles muitas vezes sentem (com razão) que suas línguas lhes foram tiradas à força, juntamente com os direitos à autodeterminação e à decisão sobre o próprio destino” (Olko & Sallabank, 2020, p. 14). Olko e Sallabank (2020) também apontam como a revitalização da língua pode

fazer parte de um processo mais amplo de renascimento cultural e reivindicação do direito de determinar o próprio destino.

A importância da língua também para a saúde geral nas comunidades indígenas também foi demonstrada em estudos: Calma (2009, p. 61), por exemplo, relata que dados de pesquisas longitudinais demonstram uma correlação entre uma língua e cultura fortes nas comunidades indígenas e resultados positivos para a saúde.

Olko e Sallabank (2020) identificam um total de seis grandes grupos de razões para a revitalização, abrangendo categorias sociais, psicológicas e físicas:

(1) conexão com os ancestrais, o passado e o patrimônio cultural

(2) cura

(3) construção da comunidade

(4) conhecimento e cultura

(5) bem-estar

(6) benefícios cognitivos

Um projeto de revitalização bem conhecido é o do Hebraico, que passou de uma língua inativa, usada apenas para rituais religiosos, a uma língua nacional em Israel, bem como uma língua usada internacionalmente por algumas comunidades judaicas em todo o mundo (Shavit, 2017), embora o Inglês continue sendo a língua principal para a maioria delas (Spolsky, 2014).

A maioria dos projetos de revitalização é realizada com línguas ainda vivas, mas ameaçadas de extinção, como o Maori na Nova Zelândia. O processo de revitalização de línguas extintas ou inativas, sem falantes vivos, levanta algumas questões adicionais, mas muitas vezes pode se basear nas experiências do processo de revitalização de línguas ameaçadas de extinção.

Papua Nova Guiné é o país com as línguas mais ameaçadas, com mais de 840 línguas faladas (Ethnologue, 2025), das quais a grande maioria está ameaçada, severamente ameaçada ou criticamente ameaçada, de acordo com a classificação da UNESCO (3, 2 ou 1, ver figura 4). Na Papua-Nova Guiné, há mais esforços para documentar as línguas vivas, mas ameaçadas, do que para tentar revitalizar as já extintas (Kik et al., 2021). Até o momento, não há dados confiáveis sobre quantas línguas foram extintas no país.

O maior número de projetos de revitalização está hoje nas Américas, tanto na América do Norte quanto na América do Sul (Pérez Báez et al., 2019). Contudo, quando considerada separadamente, a Europa ocupa o segundo lugar, atrás da América do Norte. Na América do Norte, por exemplo, o renascimento do Havaiano e do Cherokee teve efeitos mensuráveis para as comunidades. A língua Wampanoag, em Massachusetts, é um exemplo de revitalização de uma língua extinta que esteve adormecida por quase 200 anos e agora é amplamente ensinada em escolas comunitárias (Woodward, 2018).

No entanto, a revitalização não é um tema exclusivo das comunidades indígenas. De acordo com a UNESCO, existem 128 línguas ameaçadas de extinção na União Europeia, entre as quais se destacam o Livoniano e o Córnico (Parlamento Europeu, 2013). Os conceitos desenvolvidos pelos pesquisadores sobre como revitalizar uma língua são igualmente válidos para esses grupos, embora tais estruturas sejam frequentemente discutidas principalmente em relação às línguas indígenas e suas histórias coloniais específicas. Os conceitos provaram ser aplicáveis à revitalização de línguas como a de

Guernsey (onde o Guernésiais continua gravemente ameaçado) e a de Wilamowice, na Polônia (onde o Wilamowiciano tinha menos de 20 falantes em 2020), onde todos os falantes são idosos e as línguas continuam a perder falantes em ritmo acelerado (Olko & Sallabank, 2020). O número de projetos de revitalização na Europa é bastante elevado (Pérez Báez et al., 2019) e, atualmente, na maioria dos países Europeus, conta com um forte apoio institucional (em comparação com muitas outras regiões do mundo), tanto a nível nacional³ como através da União Europeia. Alguns bons exemplos disso são o Sami no norte da Escandinávia, o Córnico no Reino Unido, o Bretão na França, o Galego na Espanha e o Livônio na Letônia.

Os esforços de revitalização linguística na Ásia cresceram significativamente nas últimas décadas, à medida que as comunidades respondem às rápidas mudanças sociolinguísticas, à urbanização e à pressão da língua nacional (Pérez Báez et al., 2019). Os exemplos incluem o renascimento do Ainu no Japão por meio de ninhos linguísticos administrados pela comunidade, centros de educação cultural e documentação apoiada pelo Estado, e a revitalização do Manchu na China por meio de programas universitários, plataformas digitais e cursos de imersão liderados pela comunidade. Em toda a Ásia, centenas de línguas são consideradas ameaçadas de extinção — por exemplo, só a Indonésia tem mais de 400 línguas em risco (UNESCO, 2010) —, mas o número exato de projetos de revitalização é difícil de determinar, pois muitas iniciativas são de pequena escala, informais ou não documentadas. A maioria dos projetos de revitalização enfrenta restrições estruturais, como domínios de uso limitados, falta de falantes fluentes e compromisso político de longo prazo insuficiente (Bradley, 2019). Os casos mencionados acima mostram que, embora a revitalização na região seja possível, o sucesso depende do envolvimento duradouro da comunidade, de estratégias educacionais direcionadas e do apoio governamental sustentado.

Teoricamente, uma vez que uma língua fica dormente (no sentido de que o último falante vivo falece), ela deve parar de se desenvolver e chegar a um ponto fixo. Não deve haver acréscimos ou alterações, nem nada deve ser removido. No entanto, isso permanece uma teoria e nunca é completamente realizado, pois qualquer pessoa que comece a pesquisar a língua tirará suas próprias conclusões a partir da documentação ou das memórias existentes sobre a língua. Além disso, durante o período de inatividade, as mudanças ao longo do tempo, à medida que os locais onde a língua permanece (ou seja, rituais religiosos, canções, sonhos, etc.) estão na memória dos seres humanos vivos. Essa memória não é uma entidade fixa. Isso levanta muitas questões, como qual língua está realmente sendo revitalizada, como lidar com diferentes dialetos na documentação e durante o processo de revitalização, quais novas palavras devem ser incluídas, etc. Todas essas são questões que precisam ser abordadas em projetos de revitalização.

Vários guias para a revitalização da língua foram desenvolvidos ao longo do tempo; Olko e Sallabank (2020) é um guia muito prático para uma comunidade, Grenoble e Whaley (2005) apresentam uma abordagem mais acadêmica do processo e um manual com um escopo mais global pode ser encontrado em Hinton et al. (2018).

Uma estrutura de alto nível para um processo contínuo de revitalização orientado pelo feedback foi desenvolvida por Fitzgerald e Hinson (2013), com base no trabalho com a língua Chickasaw (*Chikashshanompa*; Muskogean), anteriormente falada em partes dos atuais Mississippi e Alabama.

³ O Wilamowiciano infelizmente está em um país onde esse não é o caso, a Polônia, e um projeto de lei para reconhecê-lo como língua minoritária foi recentemente vetado pelo presidente polonês em outubro de 2025) (Gabinete do Presidente da República da Polônia, 2025)

A nação Chickasaw foi removida à força no século XIX e agora está localizada principalmente em Oklahoma. Essa estrutura é comumente chamada de modelo Chickasaw.

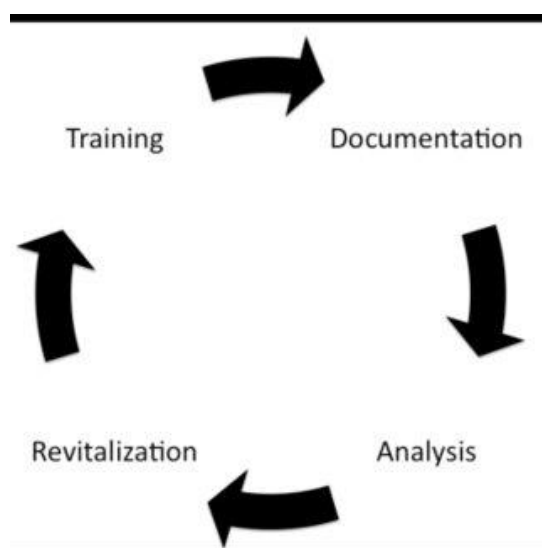


FIGURA 5 MODELO CHICKASAW PARA REVITALIZAÇÃO DA LÍNGUA (FITZGERALD & HINSON 2013, P. 59)

Cada uma das partes, conforme mostrado na Figura 5, neste processo precisa ser detalhada e desenvolvida para garantir que elas alimentem não apenas a comunidade, mas também a etapa seguinte nesta estrutura. O modelo Chickasaw de revitalização linguística opera com base na premissa de que, quando a transmissão intergeracional é interrompida e há poucos falantes fluentes, a revitalização só pode ocorrer por meio de novos falantes (em vez de depender da transmissão familiar natural). O modelo combina imersão intensiva de adultos, colaboração estreita com os falantes mais velhos remanescentes e forte apoio institucional da Nação Chickasaw para treinar falantes proficientes da segunda língua que possam usá-la na interação cotidiana e, eventualmente, ensiná-la a outras pessoas. Uma vez treinados, eles devem apoiar a documentação da língua, mas, em vez de priorizar a documentação isoladamente, o modelo trata a aquisição da língua como um gargalo central, com o objetivo de longo prazo de reconstruir o uso da língua dentro da comunidade e criar condições nas quais a transmissão intergeracional possa ser restabelecida posteriormente. A documentação será utilizada posteriormente na análise para desenvolver ainda mais a língua, etc.

Campos (2023) apresenta uma abordagem ligeiramente diferente, na qual define as diferentes fases da revitalização como documentação, criação de gramática e lexicografia, implementação do uso e, finalmente, planejamento. Referirei-me principalmente a essa estrutura a seguir, já que o processo de revitalização no contexto das línguas dormentes no Nordeste do Brasil é discutido principalmente nesses termos na literatura. Campos (2023) também aborda a diferença entre glotogênese e revitalização. Ao contrário da glotogênese — a criação de línguas inteiramente novas —, a revitalização se concentra na recuperação de línguas ancestrais que foram marginalizadas ou suprimidas por meio de políticas de colonização e assimilação (Coronel-Molina & McCarty, 2016). No entanto, tanto a glotogênese quanto os esforços de recuperação lexical pura são frequentemente incluídos no termo revitalização, como será mostrado abaixo em alguns dos casos estudados.

Harrison (2007) aborda a questão do que acontece quando as línguas se tornam inativas ou extintas, em termos da informação que se perde. Sua resposta inicial a essa questão é bastante solene: dizemos que não sabemos. De uma perspectiva linguística, a ciência entra em um modo de “coleccionar selos” quando uma língua está em perigo — uma alusão ao relato de Bernal (1939, p. 25) de que Ernest

Rutherford (1871–1937), frequentemente considerado o pai da física nuclear moderna, costumava classificar a ciência como física ou “coleccionar selos”. Nesse caso, isso significa que todo o foco está em salvar ou, pelo menos, documentar o que é possível, e pouco ou nada no que não é. Algumas respostas à questão, do ponto de vista linguístico ou científico, sobre o que se perde quando as línguas morrem, são apresentadas por Harrison (2007) quando ele argumenta que grande parte da base de conhecimento humano (por exemplo, os nomes das plantas medicinais muitas vezes dão pistas sobre seu uso), nossa herança cultural humana e nossa compreensão científica da mente humana estão incorporadas na língua. Isso também significa, de acordo com Harrison (2007), que a diversidade das línguas é um importante contributo para os alicerces e a própria arquitetura do pensamento humano. Todos estes são argumentos convincentes para promover a revitalização das línguas, mas não abordam o processo de cura que a comunidade falante muitas vezes considera prioritário. No entanto, na área das línguas ameaçadas de extinção, Harrison (2007) assume a posição de que “embora a ciência também possa servir às necessidades da comunidade falante, esse não é o objetivo principal dos cientistas”, com a simples motivação de que apenas os falantes podem “salvar” ou “resgatar” as línguas ou reverter a tendência. Não existe uma língua humana viva sem falantes. A “coleção de selos” muitas vezes não aborda as principais razões para a revitalização, como descrito anteriormente como cura ou qualquer outra das razões listadas por Olko e Sallabank (2020), conforme mostrado acima. Esses dois objetivos diferentes devem ser mantidos em mente para garantir que não se tornem opostos um ao outro. Outro exemplo de perda de informação com a perda de línguas pode ser encontrado em Cámara-Leret e Bascompte (2021), que mostram que na América do Norte, noroeste da Amazônia e Nova Guiné, o conhecimento indígena sobre plantas medicinais apresenta uma baixa redundância entre as línguas, o que é típico de sistemas com alto conteúdo informativo. Isso significa que não é importante apenas revitalizar línguas ameaçadas ou dormentes, mas fazê-lo com o maior número possível delas.

A revitalização das línguas indígenas é um processo dinâmico e multifacetado que visa reverter a mudança linguística e restaurar a vitalidade linguística nas comunidades. Experiências do Havaí, com o havaiano, e do Alasca, com o Alutiiq, demonstram que a revitalização bem-sucedida geralmente combina iniciativas populares com apoio institucional, enfatizando a educação, as práticas culturais e a transmissão intergeracional. Kimura e Counciller (2009) explicam isso, concentrando-se no processo de cunhagem de novas palavras nessas duas regiões.

Artigos acadêmicos de Guerrettaz e Engman (2023) e Campos (2023), por exemplo, destacam que esses processos não são meramente linguísticos, mas profundamente ligados à identidade, já que muitas comunidades reafirmam identidades étnicas anteriormente ocultas para sobreviver. Essa especificidade na revitalização das línguas indígenas está frequentemente ligada ao fato de pelo menos uma geração do grupo étnico já ter perdido a oportunidade de aprender a língua como língua materna. Modelos como a abordagem Chickasaw ilustram que a revitalização normalmente se desenrola em etapas: documentação, desenvolvimento pedagógico e integração na vida pública. Uma característica notável na maioria dos esforços de revitalização é a criação de neologismos, que servem tanto para fins práticos quanto simbólicos — preenchendo lacunas lexicais e, ao mesmo tempo, afirmando a criatividade e a resiliência linguísticas.

2.3. Neologismos

2.3.1. Definição de neologismos

O primeiro a afirmar a existência de neologismos parece ter sido Humboldt (1999/1836), quando defendeu a ideia de que a linguagem muda com o tempo e estimula as mentes. Isso causou uma verdadeira revolução na época e resultou em inúmeros estudos sobre o desenvolvimento da linguagem. Humboldt (1999/1836, pp. 51-51) também escreveu: “Nunca pode haver um momento de verdadeira estagnação na linguagem, assim como não há no pensamento incessante dos homens. Por natureza, é um processo contínuo de desenvolvimento.” Nem todos viram esse desenvolvimento como algo positivo, como pode ser visto diariamente nos comentários e discussões sobre a “decadência” da língua devido aos empréstimos, os esforços contínuos na França com a *Académie Française* e a *Commission d’enrichissement de la langue française (CELF)* para definir o que é e o que não é bom Francês, etc. Os neologismos são, por sua própria natureza, a extensão da língua que usamos hoje, uma vez que muitas das palavras atuais foram neologismos em algum momento. Kuldashova e Safarbaeva (2025) discutem os vários fatores que contribuem para a criação de neologismos, desde os avanços científicos até a poesia em uzbeque, mas sempre com o objetivo de que os neologismos reflitam e se adaptem às mudanças no ambiente da língua. A maioria das línguas exibe um certo conservadorismo linguístico, juntamente com mudanças contínuas (Trudgill, 2020). Termos como “Inglês correto” ou gíria⁴ em Português refletem isso. Como todas as línguas vivas estão em um processo contínuo de mudança, o conservadorismo linguístico necessariamente levanta a questão de qual momento histórico deve definir a forma “correta” de uma língua — ou seja, em que momento ela era falada, escrita, lida ou ouvida *corretamente*. Qualquer tentativa de fixar tal momento determina implicitamente o que conta como uma inovação legítima e o que é descartado como um neologismo.

Bauer (1983) forneceu uma classificação sistemática e baseada em teoria sobre os processos de formação de palavras em geral, que desde então tem sido a espinha dorsal dos estudos sobre como os neologismos são classificados. Ela fornece uma arquitetura descritiva duradoura para a formação de palavras, rejeitando modelos rígidos baseados em regras. Essa arquitetura foi desenvolvida em conjunto com a emergente linguística descritiva, que ganhou impulso a partir do final da década de 1980. Com sua abordagem baseada na realidade, ela também se conectou bem com o surgimento de grandes corpora na década de 1990, que muitas vezes exigem capacidade para lidar com dados linguísticos confusos, desiguais e do mundo real.

Pesquisadores de diferentes formações acadêmicas apresentam diferentes definições de neologismo. De acordo com Greenough e Kittredge (1901, p. 78), um neologismo pode ser descrito como uma palavra ou frase nova que ainda não foi totalmente reconhecida na variante padrão. Eles continuam especificando que, se um neologismo parece, para a maioria dos falantes, suprir uma lacuna na língua ou ser particularmente adequado para a expressão de alguma ideia especial, ele certamente se manterá contra os protestos da guilda literária e acadêmica (Algeo, 1991, p. 2). tem uma definição um pouco diferente quando argumenta que uma palavra nova é uma forma linguística ou o uso de uma forma linguística não registrada em dicionários gerais, podendo ser uma palavra isolada, um composto ou até mesmo uma frase. Alves (1990) argumenta, ao discutir neologismos em Português, que uma palavra nova pode ser formada por empréstimo ou por processos e mecanismos internos à língua. Ela também propõe uma classificação dos diferentes processos de formação neológica como: neologia fonológica, neologia sintática (derivação, composição, formação por acrônimos, composição sintagmática), neologia semântica, neologia por empréstimo, conversão (derivação imprópria), entre

⁴ As línguas indígenas brasileiras também são, por vezes, consideradas gíria num sentido pejorativo.

outros processos. Alves (1990) também argumenta que a simples criação de um neologismo não é suficiente para que ele se torne um membro integrante do estoque lexical de uma língua. Na verdade, segundo Alves (1990), é a comunidade linguística, por meio do uso do neologismo ou de sua falta de disseminação, que decide sobre a integração dessa nova formação e na língua. Oreški (2021) sugere que a crescente popularidade dos neologismos na comunidade linguística acadêmica poderia ser usada para trabalhar em direção a uma definição unificada, facilitando assim um desenvolvimento mais rápido nos estudos de neologia.

As discussões acadêmicas continuam sobre vários tópicos relativos à definição e ao escopo dos neologismos, tais como: se os ocasionalismos devem ser separados dos neologismos (Schippan, 2002) e, em caso afirmativo, com base em quais parâmetros; se os acrônimos são neologismos ou apenas variantes de lexemas existentes (Elsen, 2004); se as mudanças de valência como extensão semântica ou ressignificação devem ser consideradas neologismos (Xuefu, 1989), etc. Kinne (1998) chega a propor como uma possível definição que apenas inovações lexicais devem ser consideradas neologismos. Optei por usar uma definição que é uma ligeira modificação da de Greenough e Kittredge (1901, p. 78) para abordar melhor os neologismos nas línguas indígenas. Para este artigo, usarei a seguinte definição de neologismo:

“um neologismo é um elemento linguístico recém-introduzido em uma língua”.

Essa definição não define “recém” (há quanto tempo o neologismo existe), o que considero apropriado se aceitarmos o fato de que uma língua está em um processo contínuo de desenvolvimento. Ela também não responde à questão do que significa “introduzido” para que o neologismo seja aceito. Para a discussão sobre neologismos em línguas que estão sendo revitalizadas, faz sentido não abordar essas questões impondo limitações, pois isso poderia excluir certos processos de cunhagem, como veremos mais adiante.

A maior parte dos trabalhos sobre a descrição de neologismos concentrou-se no que, neste artigo, chamarei de línguas majoritárias (Inglês, Alemão, Português, Francês, etc.), como por exemplo (Kinne, 1998), (Ryskina et al., 2020) e (Shi, 2025). Existem alguns artigos interessantes sobre a criação de neologismos em línguas indígenas da América do Norte, como (Pentland, 1982) sobre o Cree, (Kimura & Counciller, 2009) comparando os processos de cunhagem de neologismos no Havaiano e no Alutiq, no Alasca. (Gruda, Haimovich & Sullivan, 2023) descreve um estudo empírico sobre neologismos nas variedades modernas do Nahuatl (México) e aponta como a formação de palavras está sendo aplicada a novos conceitos. Outro artigo recente, (Camacho, 2024), analisa como o processo de cunhagem para o Quíchua (ou, pelo menos, o processo de propor neologismos) pode ser automatizado usando redes neurais.

2.3.2. Classificação do processo de cunhagem

O processo de criar neologismos é frequentemente chamado de cunhagem. Uma classificação comum dos diferentes processos para neologismos (morfosintáticos) é: *composição, derivação, palavra-valise, conversão, truncamento e acrônimos*. Essas classificações são baseadas principalmente em Bauer (1983). Neologismos semânticos são frequentemente classificados em metáfora, metonímia, extensão e ressignificação. Existem também, pelo menos teoricamente, neologismos fonéticos, como, por exemplo, reduplicação e onomatopeia.

Usarei uma combinação dessas classificações neste artigo. Quando se trata especificamente de empréstimos, abrirei a categoria de empréstimos e diferenciarei dois: empréstimos da língua majoritária e empréstimos de línguas indígenas vizinhas. Essa distinção é útil no presente estudo, pois

permite uma comparação entre a influência do Português e a das línguas indígenas vizinhas nos processos de revitalização e reconstrução.

No que diz respeito aos neologismos semânticos, utilizarei duas categorias: extensão semântica e ressignificação semântica e, caso sejam encontrados, integrarei metáforas na extensão semântica e metonímias na ressignificação semântica. A extensão semântica ocorre quando o significado de uma palavra é ampliado dentro da área semântica ou para uma área diferente; portanto, trato as metáforas como um subtipo de extensão. A ressignificação semântica, por outro lado, captura processos de ressignificação e, como a mudança metonímica também envolve a reatribuição de um termo a novos objetos, optei por integrar a metonímia a essa categoria. Não abordarei de forma alguma a neologia fonológica, pois o material utilizado não permite isso (a maioria dos dados que utilizei é proveniente de pequenos materiais do tipo dicionário para uma área semântica específica ou de artigos acadêmicos e não possui glosas). Um estudo detalhado do processo de neologia fonológica no contexto da revitalização de línguas dormentes poderia certamente acrescentar informações importantes à base de conhecimento, não apenas sobre línguas indígenas, mas sobre os processos de neologia em geral.

Optei por limitar o estudo a substantivos e numerais principalmente devido à melhor disponibilidade de dados. No entanto, verbos e outras classes gramaticais merecem um estudo separado, pois isso poderia aprofundar a compreensão desses processos de cunhagem de neologismos. Estudar numerais e o sistema numérico também tem a vantagem de ter um grupo de palavras que todas as línguas abordaram em seu processo de revitalização.

Shi (2025) propõe uma classificação comparável da formação de neologismos, combinando análises qualitativas e quantitativas com base no Oxford Advanced Learner's Dictionary e distinguindo categorias como composição, derivação, truncamento (iniciais, acrônimos e truncamentos), conversão, palavra-valise, epônimo, variação, onomatopeia e outras.

O presente estudo exclui epônimos, variações e onomatopeias, pois os dados disponíveis não permitem sua identificação sistemática nas línguas analisadas. O foco é, portanto, limitado aos processos morfossintáticos e semânticos de criação de palavras.

Segue uma breve descrição de cada uma das dez categorias que utilizarei: ⁵

Empréstimos. Trata-se da adoção de palavras de outra língua, muitas vezes com adaptação da ortografia e da fonologia. Exemplos em Inglês são *sofa* (Arabe), *piano* (Italiano) e *tattoo* (Taitiano). Um tipo específico de empréstimos são os calques, que correspondem a uma tradução direta dos elementos de uma palavra para a língua de destino. Por exemplo, a palavra Inglesa “skyscraper” foi calcada em várias línguas: *gratte-ciel* (Francês), *Wolkenkratzer* (Alemão), *skyskrapa* (Sueco).

Composto. É a união de duas ou mais palavras separadas para produzir uma única forma. Alguns exemplos são “wallpaper” (papel de parede), “fast-food” (comida rápida) em Inglês, *portefeuille* (*porte* (carregar) + *feuille* (folha, folha de papel)), que significa pasta, *porte-monnaie* (*porte* (carregar) + *monnaie* (dinheiro, moeda)), que significa carteira em Francês, e *guarda-chuva* (*guarda* (proteger) + *chuva* (chuva)), que significa guarda-chuva, e *segunda-feira* (*segunda* (segundo) + *feira* (dia litúrgico em Português antigo)), que significa segunda-feira em Português. O Alemão é uma das línguas mais produtivas quando se trata de composição, como por exemplo

⁵ Baseei as categorias em (Yule, 2023), mas acrescentei as explicações e a maioria dos exemplos. Nos exemplos, não fui rigoroso quanto ao momento em que as palavras foram consideradas neologismos, para me alinhar à definição de neologismo usada neste artigo.

Kühlschranktür (*Kühlschrank* (geladeira) + *Tür* (porta)), que significa porta da geladeira, onde *Kühlschrank*, por si só, também é uma composição de *Kühl* (frio) e *Schrank* (armário).

Derivação. Isso é muito comum e geralmente é feito adicionando diferentes afixos a um morfema central. Os afixos podem ser prefixos, sufixos ou, em alguns casos, infixos. Alguns exemplos antigos em Inglês seriam “foolishness” (tolice), em que o núcleo é “fool” (tolo) e dois sufixos são adicionados *-ish* e *-ness*, e “untighten” (afrouxar), com o núcleo “tighten” (apertar) e o prefixo *un-*. Em Francês, exemplos antigos seriam *impossibilité* (impossibilidade), com o núcleo *possible*, um prefixo *im-* e um sufixo *-ité*; em Espanhol, alguns exemplos antigos são *felicidad* (felicidade), do núcleo *feliz* e do sufixo *-idad*; e em Holandês, *kennisgeving* (notificação), do núcleo *kennis* (conhecimento) e do sufixo *-geving* (semelhante ao sufixo Inglês *-giving*). Todas essas palavras já foram neologismos em algum momento. Para dar alguns exemplos mais modernos em Inglês: “unfriend” e “influencer”, e, em Português, *apagão* (falha de energia elétrica), do verbo *apagar* com o sufixo aumentativo *-ão*, e *flamenguista* (torcedor/fã do clube de futebol Flamengo), criado a partir da forma abreviada *Flamengo* deste importante time de futebol do Rio de Janeiro, *Clube de Regatas do Flamengo*, e do sufixo *-(u)ista*, que produz torcedor/fã ou profissões.

Palavra-valise. De certa forma semelhante à composição, a palavra-valise combina duas formas separadas para produzir um único termo novo, na maioria das vezes pegando o início de uma palavra e juntando o final da outra, como “brunch” (breakfast + lunch) em Inglês, *Macronautes* (*Macron* + *astronautes*) para fãs de Macron em Francês ou a palavra-valise Portuguesa *portanhol* (*Português* + *espanhol*), que significa uma língua que mistura Português e Espanhol.

Conversão Significa uma mudança na classe gramatical, ou seja, um verbo que se torna um substantivo ou vice-versa. Conversões comumente usadas em Inglês seriam “to glue something” (colar algo), “having doubts” (ter dúvidas) e “handouts” (folhetos), e em Alemão: *das Essen* (a refeição, do verbo *essen*), *parken* (estacionar, do substantivo *der Park*) e *die Schöne* (a bonita, a bela, do adjetivo *schön*). Alguns neologismos mais modernos que utilizam o processo de conversão seriam “to ghost someone” (ignorar alguém) em Inglês e *être bof* (ser indiferente) da interjeição *bof* (meh, não me importo) em Francês.

Truncamento (recorte) Esse processo mostra ainda mais o elemento de redução na mistura, pois não há adição de outra palavra. Isso também acontece frequentemente com nomes. Alguns exemplos: “gas” (gasolina), “fax” (fac-símile), “condo” (condomínio), “Ron” (Ronald) em Inglês, *ado* (*adolescente*), *bac* (*bacharelado*), *Alex* (*Alexandre*), *prof* (*professeur/professeure*) em Francês, e bio (cinema de biograf), *matte* (matemática de *matematik*), *tel* (número de telefone de *telefonnummer*) em Sueco.

Acrônimo (sigla) constituem uma classe de neologismos formados a partir das letras iniciais de expressões com várias palavras. Elas podem ser pronunciadas articulando cada letra individualmente ou, mais comumente, lendo a sequência resultante como uma única unidade lexical. Exemplos do primeiro tipo incluem o Inglês “CD” (*compact disc*) e “ATM” (*automatic teller machine*), bem como o Francês TGV (*train à grande vitesse*) e RER (*réseau express régional*). Os acrônimos do segundo tipo — aquelas

pronunciadas como palavras completas — são bem atestadas em Inglês, como “laser” (*Light Amplification by Stimulated Emission of Radiation*) e “scuba” (*Self-Contained Underwater Breathing Apparatus*). O discurso público contemporâneo ocasionalmente produz formações semelhantes, por exemplo, o slogan político “TACO” (*Trump Always Chickens Out, ou Trump sempre se acovarda*). A formação de acrônimos também desempenhou um papel significativo em outras línguas. Em Russo, por exemplo, ВУЗ (*VUZ, Высшее учебное заведение*, ou instituição de ensino superior/universidade) tornou-se um substantivo comum. Outro caso notável é ГУЛАГ (*GULAG, Главное управление лагерей*, “Administração Principal dos Campos”), que não só se tornou um substantivo em russo, mas também entrou em várias outras línguas. Padrões semelhantes podem ser observados em acrônimos russas reconhecidas internacionalmente, como KGB (*Комитет государственной безопасности, Komitet Gosudarstvennoy Bezopasnosti*, “Comitê de Segurança do Estado”) e АК-47 (*Автомат Калашникова 1947, Avtomat Kalashnikova 1947*, “Fuzil de assalto Kalashnikov, modelo 1947”). Esses exemplos ilustram como os acrônimos podem se transformar em itens totalmente lexicalizados que circulam muito além de seus contextos originais.

Extensão semântica É quando o significado de uma palavra é ampliado dentro da área semântica ou para uma área diferente. Vários exemplos disso podem ser vistos, em muitas línguas, quando se trata de tecnologia da informação. Exemplos disso são “mouse” (Inglês e, por empréstimo, Português) / *mus* (Sueco) / *souris* (Francês) para o pequeno dispositivo portátil usado para controlar um ponteiro na tela do computador e “net” (Inglês) / *rede* (Português), *réseau* (Francês) / *nät* (Sueco) para o sistema de computadores e servidores conectados na internet.

A ressignificação semântica ocorre quando uma palavra mantém sua forma, mas seu significado muda significativamente. Alguns exemplos disso em Inglês são “gay” (de feliz para homossexual), “cool” (de frio para socialmente atraente) e “hacker” (de alguém que hackeia ou corta grosseiramente para alguém que é habilidoso em programação de computadores ou segurança digital). Alguns exemplos em Francês seriam *clé* (de chave para memória USB) e *portable* (de algo que pode ser transportado para um celular).

Combinação destes. A palavra sueca *paraply* (guarda-chuva) é um tipo de combinação de composição e empréstimo, uma vez que *parapluie* é uma composição francesa (*parer + pluie*) que foi emprestada ao Sueco com apenas uma alteração ortográfica mínima. A palavra sueca *vabba* também é uma combinação interessante, em que *VAB* é o acrônimo para *Vård Av (sjukt) Barn* (cuidar de criança (doente)) e isso mudou, por meio de conversão, para o verbo *vabba*, sendo principalmente o ato de cuidar de uma criança do ponto de vista administrativo/remuneratório.

2.3.3 Alguns aspectos sociolinguísticos dos neologismos

De uma perspectiva sociolinguística, os neologismos são frequentemente vistos como uma expressão da relação — muitas vezes de resistência — com a língua majoritária (Campos (2023); Loncón Antileo e Castillo Sánchez (2019)), especialmente quando se trata de línguas indígenas. Isso, é claro, não deve ser confundido com purismo linguístico, como no caso do Islandês, que usa a fonologia e a morfologia existentes para criar neologismos semânticos como *hitabelti* (“trópicos”) a partir de *hiti* (calor) e *belti*

(cinto) ou *veðurfræði* (meteorologia) a partir de *veður* (clima) e *fræði* (ciência). A Islândia mantém uma política linguística prescritiva tradicional e a ideologia atual ainda é de continuidade linguística ou “manter a linha”, coordenada pelo Conselho da Língua Islandesa (criado em 1964) (Leonard & Árnason, 2011). Calvet (2005) descreve como as palavras carregam significado político, uma vez que aqueles que nomeiam o mundo também têm o poder de defini-lo, e ele continua explicando como isso aconteceu durante o colonialismo e a glotofagia resultante⁶. Essa perspectiva é igualmente relevante quando se discute a revitalização de línguas dormentes ou anteriormente faladas, e constitui um elemento importante do processo mais amplo de recuperação da língua pelas comunidades indígenas.

Loncón Antileo e Castillo Sánchez (2019) descrevem, nesse contexto, como os neologismos espontâneos em Mapuzugun (língua isolada de acordo com Adelaar e Pache (2022), entre falantes que vivem em Santiago do Chile, estão sendo criados predominantemente com bases de processos de formação de substantivos inerentes à língua, em oposição a empréstimos do Espanhol ou do Inglês com adaptações morfológicas e fonológicas. Os autores relacionam isso à forte conexão ideológica dos participantes com a cultura Mapuzugun.

Em De Paula Camargo (2020), o autor discute isso com o exemplo da nação Wakalitesu (Nambikwara) da terra indígena Tirecatinga, no nordeste do Mato Grosso. Os neologismos nessa língua são frequentemente criados usando a forma, o uso ou outros parâmetros dos objetos, em vez de pegar diretamente a palavra em Português e adaptá-la à fonologia e morfologia, como mostrado nos exemplos (1) e (2), enquanto os empréstimos do Português também podem formar a base, como mostrado no exemplo (3).

(1) [ˈwaĩsakanũsu]
 waĩsa- ka- nũ- su
 ADJ.forte- ?- COMP.pulverulento- S.REF
 “café em pó”
 Lit.: “algo forte e em pó” (De Paula Camargo, 2020. p. 67)

(2) [naũˈdakisu]
 naũ- da- ki- su
 ADJ.doce- ? COMP.redondo- S.REF
 ‘doce’
 Lit.: “algo redondo e doce” (De Paula Camargo, 2020, p. 66)

(3) [kanetaˈkatsu]
 kaneta- kat- su
 pen COMP.long S.REF
 ‘caneta’
 Lit.: “caneta longa” (De Paula Camargo, 2020, p. 60)
 (usado por falantes mais jovens do Wakalitesu)

⁶ Glottophagie. Termo cunhado por CALVET, L.J. para descrever o processo em que uma língua “engole” outra, particularmente no contexto colonial. (Calvet, 2005. p.11)

3. Contexto para a revitalização no Nordeste brasileiro.

A extinção linguística quase total no nordeste do Brasil é, em grande parte, resultado de ter sido a primeira região a ser colonizada e de que isso ocorreu de forma muito mais intensa do que em outras regiões. Os portugueses estabeleceram seus primeiros assentamentos permanentes ao longo da costa nordeste, começando na Bahia e em Pernambuco, que se tornaram os primeiros centros de produção de açúcar e colonização europeia.

A leste da linha marcada na Figura 1 (página 5), a maioria das línguas indígenas desapareceu (Rodrigues, 1986 e 2002). No nordeste, as únicas exceções são o Yathé, falado pelo povo Fulni-ô em Pernambuco, o Maxakali e o Borum (língua falada pelo povo Krenak no nordeste de Minas Gerais, Bahia e São Paulo). Quanto ao Borum, a língua é hoje considerada moribunda, com apenas idosos sendo bilíngues em Português/Borum (Instituto Socioambiental, s.d.; Silva, 2025). Na Figura 1 (página 5), também é possível ver vários exemplos de Tupi-Guarani em vários estados, Rio Grande do Sul, Santa Catarina, São Paulo, Rio de Janeiro, Paraná e Espírito Santo. Todos eles são descendentes do oeste, principalmente do Vale do Rio Paraná.

As línguas selecionadas para este estudo são o Patxohã, o Yathé, o Kariri-Xocó, o Xukuru e o Borum. Essa seleção teve como objetivo permitir uma comparação dos processos de cunhagem em línguas em diferentes estágios de revitalização e que empregam diferentes estratégias de revitalização. O território dos Krenak (falantes do Borum) está hoje fora do que é chamado de Nordeste, mas eu os incluí por duas razões: em primeiro lugar, porque eles também têm uma longa história no Nordeste e muito contato, especialmente com os Pataxó. Em segundo lugar, o Borum foi uma das poucas línguas que resistiram à colonização e nunca perderam completamente seus falantes ativos. A Figura 6 mostra a situação geográfica, até hoje, das principais comunidades com essas línguas.



FIGURA 6 REGIÃO NORDESTE DO BRASIL COM AS COMUNIDADES INDÍGENAS CUJAS LÍNGUAS SÃO ESTUDADAS NESTE ARTIGO

O Krenak/Borum é classificado como *severamente ameaçado* (nível 2) em vários bancos de dados de línguas ameaçadas, enquanto o Xukuru, o Pataxó e o Kariri-Xocó são listados como *dormentes ou extintos* (nível 0). O Yaté e o Maxakali, por outro lado, são classificados como *inseguros ou ameaçados* (nível 4). O Patxohã, sendo uma língua contemporânea distinta da variedade ancestral Pataxó, não está incluído em nenhum banco de dados importante sobre línguas ameaçadas. Essa variedade de status de ameaça é vantajosa para o presente estudo, pois oferece uma perspectiva comparativa mais ampla sobre os processos de revitalização, especialmente considerando que essas línguas também se encontram em estágios marcadamente diferentes de revitalização.

Alguns parâmetros sociolinguísticos para as diferentes línguas estudadas seguem na Tabela 2:

Língua	Família linguística	Povo	Tamanho aproximado da comunidade	Grau de ameaça (UNESCO 2013)	Número aproximado de falantes
Patxohã	(Macro-Jê)	Pataxó	20 000	(Não classificado)	-
Kariri-Xocó	Macro-Jê	Kariri-Xocó	2000 - 4000	0	-
Yathé	Isolado	Fulni-ô	4000	4	500 - 1000
Borum	Macro-Jê	Krenak	800	2	<10
Xukuru	Possivelmente Macro-Jê	Xukuru	12.000	0	0

TABELA 2 PARÂMETROS SOCIOLINGUÍSTICOS PARA AS LÍNGUAS ESTUDADAS

Quando se trata dos fatores sociolinguísticos mais comumente citados na pesquisa geral sobre neologismos, eles são: vitalidade da língua, atitudes da comunidade, grau de bilinguismo, relações históricas e políticas, domínio de uso, presença de planejamento linguístico e distribuição demográfica dos falantes, conforme discutido, por exemplo, em Kirilenki (2024). Como apenas em um caso, o Yathé, o bilinguismo esteve presente durante todo o processo de colonização, dei muito pouca atenção a esse fator. Todos os outros fatores acima serão considerados, mas em maior ou menor grau, dependendo do contexto e da história de cada língua.

Estudei os parâmetros tipológicos das diferentes línguas, mas principalmente como uma ferramenta para identificar neologismos. A tipologia e a proximidade tipológica com línguas vizinhas também podem se refletir nos neologismos e na escolha dos processos de cunhagem, mas isso está fora do escopo desta tese. As informações tipológicas são, portanto, utilizadas metodologicamente neste trabalho, e não como variável explicativa.

Nas seções seguintes deste capítulo, descreverei o contexto sociolinguístico, a língua e a fase do processo de revitalização, bem como alguns parâmetros tipológicos para cada uma das línguas estudadas. Este contexto será posteriormente correlacionado com os processos de cunhagem encontrados para cada língua.

Também incluí uma subseção 3.6 descrevendo as estruturas numéricas básicas utilizadas nas línguas para poder ligar os processos de cunhagem de palavras numéricas escolhidos a isso, de forma semelhante ao que foi feito com os neologismos substantivos.

3.1 Pataxó, Patxohã: Língua e povo

História do povo Pataxó

A língua Patxohã é usada pelo povo Pataxó, com uma população de cerca de 20.000 pessoas em um total de 44 aldeias nos estados da Bahia e Minas Gerais. A língua histórica dos Pataxó era chamada de Pataxó, do tronco Matakali da família Macro-Jê (Rodrigues, 1986).

O povo Pataxó continua lutando por seus direitos à terra em três frentes: administrativamente, para obter terras registradas em algumas áreas e ampliar as áreas já registradas; e fisicamente, em terras registradas, contra invasões, cultivo, etc. por não indígenas em suas terras (Bomfim, 2017).

Como a maioria da população indígena do nordeste do Brasil, os Pataxó foram forçados, como parte das políticas de assimilação colonial, a usar apenas o Português para se comunicar, e a língua foi lentamente silenciada. Oficialmente, o povo Pataxó foi considerado extinto (assim como a língua) até 1970 (apesar de ser, e ter sido em 1970, a maior população indígena do estado da Bahia), e suas terras só começaram a ser registradas em 1992. Durante o regime de Getúlio Vargas (primeiro ditador militar do “Estado Novo” 1937-1945) e, mais tarde, como presidente eleito (1951-1954), a população indígena recebeu, no papel, algum reconhecimento. Ele instituiu o *Dia do Índio* em 19 de abril de 1943 e cooptou a figura do “índio” como símbolo das raízes autênticas do Brasil. Esse uso da população indígena como símbolo, aliado a políticas expansionistas como a *Marcha para o Oeste*, ilustra como sua política era de assimilação, deslocamento e transferência de territórios para o setor agrícola (tudo por meio da agência SPI (*Serviço de Proteção ao Índio*)), em vez de conceder-lhes autonomia genuína. Em 1943, na região Pataxó, ao sul de Porto Seguro, no estado da Bahia, foi criado o Parque Nacional do Monte Pascoal. Isso fazia parte de uma política geral de criação da identidade nacional, e o parque começou a ser protegido contra o uso para outros fins, como a pesca e as plantações dos Pataxó, mas também contra o uso por não indígenas. A proteção contra estes últimos, como costuma acontecer, não foi muito eficiente e o conflito se intensificou, levando ao terrível massacre da população Pataxó chamado *O fogo de 51 (Joôpek 51)* na aldeia de Barra Velha (Garfield 2000). Essa aldeia é o que se tornou mais tarde a *Aldeia Mãe* (em Patxohã: *pataxi Imamakã*, “Aldeia Mãe”), uma espécie de ponto central dos Pataxó. A Aldeia Mãe tem sido um ponto central de comércio entre as diferentes nações e, portanto, parte do “caldeirão” do grupo linguístico Maxakali. Como o processo de revitalização utilizou o Maxakali como base principal, compararei frequentemente o Patxohã ao Maxakali a seguir.

História da língua Patxohã

No final da década de 1990 e início da década de 2000, um grupo de professores de Barra Velha — juntamente com educadores de outras aldeias Pataxó no sul da Bahia, como Coroa Vermelha — começou a coletar fontes históricas (por exemplo, Wied-Neuwied, 1989) e a registrar as memórias antes silenciadas de seus anciãos. Quando começaram esse trabalho, eles conseguiram recuperar apenas cerca de 200 a 250 palavras da língua (Bomfim, 2017; Oliveira Gonçalves, 2025).

Os Pataxó tinham historicamente um contato muito próximo com a nação Maxakali e frequentemente entravam em conflitos com os Botocudo (também chamados de Krenak) (Bomfim, 2017). Ao iniciar o projeto, as influências e discussões com ambas as nações foram naturalmente parte do processo. Como o Maxakali é uma língua muito próxima do Pataxó, ela era frequentemente usada como base na criação de novas palavras, mas palavras como *jokana* (“mulher”) e *miãga* (“água”) eram as mesmas usadas pelos Krenak. Não se sabe se isso é um empréstimo do Pataxó ao Krenak ou se os Krenak o emprestaram anteriormente, como costuma acontecer entre línguas em contato. Barra Velha, agora conhecida como *Aldeia Mãe*, tem sido um ponto central de comércio entre as diferentes nações e,

portanto, parte do “caldeirão” do tronco linguístico Maxakali. Como o processo de revitalização e usou o Maxakali como base principal, compararei frequentemente o Patxohã ao Maxakali a seguir.

O processo de revitalização é melhor descrito em Franchetto (2021, p.30):

“Patxohã não ressuscitou uma língua antiga, mas foi gradualmente construída por meio de uma espécie de bricolagem criativa e com notável rigor, explorando documentos históricos, as memórias dos anciãos e a intensa troca com os Maxakali, que ainda hoje têm falantes nativos fluentes de uma língua muito próxima do antigo Pataxó”.

Desde o início, o processo de revitalização foi deliberadamente definido como um projeto “conduzido pelos Pataxó para os Pataxó”, sem dependência de especialistas externos. Esse princípio tem sido mantido de forma consistente e fortaleceu significativamente a aceitação entre os alunos da comunidade. Ao mesmo tempo, a iniciativa inclui linguistas e estudiosos indígenas — principalmente Anari Braz Bomfim — que descreve em Bomfim (2017) como os Pataxó estão reconstruindo o Patxohã com base no conhecimento dos anciãos, nos arquivos organizados pela comunidade e nas tradições da memória oral. Seu relato demonstra que o renascimento linguístico é inseparável de um movimento mais amplo de reafirmação cultural e política e que, no caso dos Pataxó, muitas das decisões, incluindo aquelas sobre os processos de cunhagem de palavras, são reflexos diretos da trajetória histórica do povo Pataxó.

Lançado no início dos anos 2000 e ainda em andamento, o projeto tomou decisões iniciais sobre ortografia, estruturas gramaticais básicas e os critérios para selecionar e criar novas palavras. Como os Pataxó haviam mudado para o Português como sua L1 nessa época, certas escolhas alinharam intencionalmente o Patxohã com os padrões estruturais do Português — por exemplo, adotando a ordem das palavras SVO como uma alternativa disponível ao padrão SOV encontrado em Maxakali (Bomfim, 2017, p. 78).

Relacionado ao povo Pataxó, os Pataxó Hãhãhãe, no interior da Bahia, estão seguindo um caminho paralelo, mas muito diferente, para revitalizar sua língua ancestral. O povo Pataxó Hãhãhãe é muito mais resultado da mistura racial com diferentes povos: Baenã, Kamakã, Tupinambá, Kariri-Sapuyá e Guerén. Sua língua ancestral também é resultado da combinação histórica dessas línguas. O projeto de revitalização é mais voltado para reafirmar a identidade do que para a fluência na língua. Nelson (2018) descreve a criação do léxico a ser dividido em uma seção Baenã, uma seção Kamakã, uma seção Kariri-Sapuyá e uma seção Pataxó Hãhãhãe. O plano é que uma seção Tupinambá seja incluída posteriormente. A comunidade considera importante a preservação das informações sobre a *raiz*. Isso dá mais um sinal de como o contexto sociolinguístico se reflete diretamente nas decisões linguísticas, incluindo os processos de cunhagem de novos termos durante a revitalização. No entanto, não entrei em mais detalhes sobre as especificidades do processo de revitalização do Pataxó Hãhãhãe.

Alguns parâmetros tipológicos

A língua linguisticamente mais próxima do Pataxó é o Maxakali, que tem uma ordem de palavras SOV, mas o Português brasileiro tem SVO. Ao definir o Patxohã, tanto SOV quanto SVO são variantes aceitas e parece que a ordem SVO é, em geral, preferida pelos falantes (Bomfim, 2017). Quando se trata da ordem do substantivo e do adjetivo, tanto Substantivo-Adjetivo quanto Adjetivo-Substantivo são variantes aceitas.

É uma língua que permite argumentos nus (como mostrado no exemplo (4)), mas contém determinantes (dois gêneros (masculino e feminino, singular/plural) de acordo com a Tabela 3:

Número	Feminino	Masculino
Singular	iě	iõ
Plural	iěp	iõp

TABELA 3 SISTEMA DE DETERMINANTES DEFINIDOS EM PATXOHÃ (BOMFIM & LIMA, 2020)

(4) João dxê'ã jokana

João viu mulher

“João viu uma/algumas mulheres”

(Bomfim & Lima, 2020, p. 328)

Os verbos são marcados por meio de categorias verbais, como o tempo verbal, com um conjunto de sufixos, muito semelhante à lógica do Português e completamente diferente do Maxakali (Bomfim & Lima, 2020). Um exemplo com o verbo caçar *kôhay* é mostrado na Tabela 4.

TABELA 4 SUFIOS VERBAIS PARA MARCAR O TEMPO VERBAL EM PATXOHÃ (BRAZ DA CONCEIÇÃO, 2016, PP. 25-26)

sufixo	forma verbal/verbal finita	forma verbal	tradução para o Português
-rê	infinitivo	<i>kôhayrê</i>	caçar
-irá	gerúndio	<i>kôhay'irá</i>	caçando
-txê	particípio	<i>kôhaytxê</i>	(foi caçado)
-ã	passado finito	<i>kôhay'ã</i>	caçado (ontem)

Os substantivos também carregam sufixos derivacionais para determinadas funções, o que também difere do Maxakali, como pode ser visto na tabela 5.

sufixo	Função	forma base	forma derivada	Significado em Inglês
-txê	Coletivo	<i>mukusuy</i>	<i>mukusuytxê</i>	um peixe -> cardume de peixes
-ará	Profissão	<i>hamiá</i>	<i>hamiará</i>	dançar (verbo) -> dançarino
-aô	Objeto	<i>goyá</i>	<i>goyáô</i>	bebida -> bebedouro

TABELA 5 SUFIOS DERIVACIONAIS EM SUBSTANTIVOS EM PATXOHÃ (BRAZ DA CONCEIÇÃO, 2016, P. 27)

O Patxohã apresenta alinhamento nominativo-acusativo, como o Português, enquanto o Maxakali é ergativo. Essa escolha, assim como muitas outras feitas pela⁷ Atxohã ao definir a língua Patxohã,

⁷ Atxohã é o grupo comunitário que coordena a revitalização da língua Patxohã

conforme descrito por Bomfim (2017), provavelmente faz parte da tendência geral de adotar parâmetros sintáticos do Português para promover a aceitação da língua Patxohã pelos alunos Pataxó.

3.2 Kariri-Xocó: povo e língua

História do povo Kariri-Xocó

A língua Kariri-Xocó (também conhecida como dzubukuá) é utilizada pelo povo Kariri-Xocó, com uma população entre 2.000 e 4.000 pessoas, no município de Porto Real do Colégio, no estado de Alagoas, ao longo do rio São Francisco. Faz parte do tronco Kariri da família Macro-Jê (Rodrigues, 1986). A denominação Kariri-Xocó é consequência da fusão, em 1873, entre os Kariris que viviam na Missão de Porto Real de Colégio e parte dos Xocós da ilha fluvial de São Pedro⁸, em Sergipe, no rio *Opará* (que hoje se chama São Francisco. *Opará* significa rio ou rio-mar em algumas línguas Tupi-Guarani, conforme afirmado em Florêncio (2022)). Quando as aldeias indígenas foram extintas por decreto imperial (declarando extintos os povos, línguas e culturas indígenas do Nordeste do Brasil), os Xocó tiveram suas terras invadidas e foram se refugiar com os Kariri do outro lado do rio. Vários outros grupos étnicos também se juntaram à Missão Jesuíta em Porto Real Do Colégio, como os Fulni-ô, Pankaru, Aconã, etc. A fusão foi certamente um longo processo, família por família e casal por casal, e a identidade resultante dessa fusão é o que hoje é chamado de nação Kariri-Xocó.

História da língua Kariri-Xocó

O Kariri-Xocó é uma língua da família linguística Kariri, frequentemente chamada de Kipeá ou Dzubukuá-Kipeá. As outras línguas da família Kariri são o Dzubukuá, o Sabuyá e o Kamurú.

Houve desde cedo um forte interesse em descrever as línguas Kariri e, já em 1699, o padre Jesuíta Mamiani escreveu a primeira gramática da língua Kariri (Mamiani, 1877/1699).

Não há uma distinção muito rígida entre as diferentes línguas da família Kariri no projeto de revitalização. Nhenety Kariri-Xocó escreve em seu blog (Kariri-Xocó, blog de 25 de dezembro de 2018) sobre a gramática básica que a língua, moldada como o Kariri moderno, é o resultado de uma compilação dos dialetos Kariri (Dzubukuá, Kipeá, Sapuya ou Sabujá, Kamuru ou Pedra Branca e Natu). O processo de revitalização começou a partir da base linguística do dialeto Dzubukuá, e as lacunas foram preenchidas com palavras de outros dialetos.

Os Kariri-Xocó empreenderam a recuperação de sua língua como uma atividade popular, sem contribuições significativas de pesquisadores com formação acadêmica (Kariri-Xocó, s.d.; Nelson & Pitman, 2020). Nhenety Kariri-Xocó tem sido uma figura-chave no projeto de revitalização. Ele começou a coletar palavras de memória, entrevistando os mais velhos e a partir de material de arquivo da coleção da FUNAI no Museu do Índio, no Rio de Janeiro (Nelson et al., 2023). A forte adoção de novas tecnologias de comunicação, como a criação de blogs comunitários e grupos no WhatsApp, tem sido fundamental para a retomada da língua. Em 2019, a sobrinha de Nhenety, Idiane Kariri-Xocó, discursou em uma conferência Viva Língua Viva na UFRJ (Universidade Federal do Rio de Janeiro) e iniciou seu discurso em Kariri-Xocó, uma língua que havia sido classificada como extinta pelo governo brasileiro 150 anos antes!

Como há esforços de revitalização em vários desses dialetos, minha primeira hipótese foi limitar meu estudo à língua Dzubukuá (baseado em Nhenety, blog de 25 de dezembro de 2018), porém, segundo

⁸ A ilha foi-lhes dada como agradecimento por terem ajudado os portugueses a expulsar os holandeses no início do período de colonização (Dantas & Dallar, 1980).

- (7) dó *addo*
 IMP 2SG.PRES.comer
 “Coma!” (Correia de Queiroz, 2012, p. 143)
- (8) dó *addoá*
 IMP 2PL.PRES.comer
 “Comam!” (Correia de Queiroz, 2012, p. 143)

3.3 Fulni-ô, Yathé: povo e língua

História do povo Fulni-ô

Yathé (também grafado Ia-Tê e Yaathê) é a língua falada pelo povo Fulni-ô, cuja população atual é de aproximadamente 4.000 indivíduos que vivem principalmente nos arredores da cidade de Águas Belas, no estado de Pernambuco (Pinto & Monteiro, 2024). Os Fulni-ô conseguiram manter sua língua, embora ela tenha permanecido gravemente ameaçada durante os períodos colonial e imperial da história brasileira.

Os Fulni-ô são um dos poucos grupos indígenas do Nordeste brasileiro que preservaram tanto sua língua original quanto suas práticas culturais. Historicamente conhecidos como Carnijó, eles resistiram à assimilação durante a época colonial e mantiveram um certo grau de autonomia, apesar da intensa pressão dos colonizadores portugueses e, mais tarde, da sociedade brasileira. Em determinado momento, as autoridades governamentais de Pernambuco chegaram a declará-los extintos, a fim de redistribuir suas terras aos colonos brancos. Durante a Guerra do Paraguai, os Fulni-ô — juntamente com outros povos indígenas do Nordeste brasileiro, como os Xukuru e os Wassú — foram arrastados para o conflito em circunstâncias coercivas. Relatos históricos indicam que muitos homens Fulni-ô foram recrutados à força ou mesmo sequestrados para servir no Exército Imperial Brasileiro, muitas vezes sem consentimento e em condições adversas (Silva, 2007; Silva, 2015). As histórias orais indígenas preservadas entre os Fulni-ô lembram esse período como um de sofrimento e perda, o que contrasta fortemente com as narrativas oficiais que retratam a guerra como um triunfo nacional. Essas memórias exemplificam as perspectivas contrastantes sobre os eventos históricos mantidas pela sociedade majoritária brasileira e pelas comunidades indígenas. Após a guerra, a maioria dos soldados Fulni-ô não retornou, mas uma parte do território original dos Fulni-ô foi concedida a algumas centenas de famílias como compensação (Porantim, 1989).

Fundamental para a identidade Fulni-ô é o ritual Ouricuri, uma cerimônia sagrada realizada anualmente em uma área florestal isolada e com duração de cerca de três meses. O ritual reforça a continuidade espiritual e cultural, enfatizando o sigilo e a resiliência contra influências externas (Fundação Nacional dos Povos Indígenas, 2022). Ao contrário de muitos outros grupos indígenas do Nordeste que perderam suas línguas, os Fulni-ô preservaram o Yathé como um marcador de identidade étnica, transmitindo-o juntamente com o Português (Campani & Urani, 2025). A profunda conexão com o ritual Ouricuri — e sua longa duração — tem sido crucial para permitir essa continuidade.

História da língua Yathé

O termo *Yathé* deriva de *ya:thê*, que significa “nossa língua”. Ainda não foi estabelecido de forma conclusiva se o Yathé pertence à família linguística Macro-Jê. Greenberg (1987) e Kaufman (1990)

classificaram-no dentro da família Macro-Jê, embora sem comentários adicionais. Ribeiro e van der Voort (2010) questionaram essa afiliação com base em uma análise fonêmica inicial; Jolkesky (2016) também expressou dúvidas, mas não chegou a nenhuma conclusão. Usando reconstrução histórico-comparativa, Nikulin (2020) incluiu vários troncos centrais da Macro-Jê, mas excluiu explicitamente o Yathé — junto com o Borôro, Purí, Guató, Karirí e Otí — devido à ausência de características fonológicas e morfológicas compartilhadas. Seguindo as descobertas de Nikulin (2020), o Yathé é atualmente considerado principalmente como uma língua isolada.

O Yathé nunca se extinguiu totalmente. Em 2023, o número de falantes foi estimado em cerca de 900 (Eberhard et al., 2023), aproximadamente 25% da população total Fulni-ô. Isso representa um declínio significativo em relação aos 2.000-3.000 falantes estimados algumas décadas antes.

A comunidade Fulni-ô implementou várias estratégias culturalmente enraizadas para manter e revitalizar o Yathé. Muitos esforços estão centrados no ritual Ouricuri (Instituto Socioambiental, s.d.), que reforça a transmissão intergeracional e a identidade cultural. Além disso, as escolas locais em Águas Belas incorporam o ensino do Yathé juntamente com o Português, introduzindo vocabulário básico e tradições orais aos alunos mais jovens. Linguistas colaboraram com a comunidade para documentar narrativas orais e vocabulário, produzindo materiais usados em oficinas educacionais (Jolkesky, 2016; Nikulin, 2020). Programas de rádio ocasionais e eventos culturais também promovem o Yathé em espaços públicos, fomentando o orgulho e a visibilidade (Eberhard et al., 2023). Embora essas iniciativas sejam principalmente impulsionadas pela comunidade, e não financiadas pelo Estado, elas desempenham um papel importante na manutenção e revitalização do Yathé.

O Yathé está bastante bem documentado em termos de fonologia e gramática. Na década de 1960, Lapenda (1962) produziu o primeiro estudo sistemático desses domínios, que mais tarde serviu de base para Da Costa (1999, 2015) e Silva (2011), cujo trabalho explora a morfologia da língua com maior profundidade.

Alguns parâmetros tipológicos

O Yathé é geralmente descrito como tendo uma ordem básica das palavras SOV e posposições, porém não de forma rígida. A ordem das palavras parece sinalizar funções pragmáticas para se alinhar com os objetivos comunicativos dos falantes ou ordens fixas para certos verbos, como mostrado nos exemplos (9) e (10).

(9) *ohya-te i nétkaka*
 água-com eu quero-LIM
 “Eu quero água” (Meland & Meland, 2009, p.72)

(10) *néma ta ik koka hle-de*
 e-eu ele para dar-LIM agora-então
 “e então ele me deu” (Meland & Meland, 2009, p.72)

É altamente flexional e morfológicamente rica, como mostrado nos exemplos (11) e (12)

(11) *efo'dwãne*
 [e fo 'dwã ne]
 3SG marido PRIV FEM
 Lit. A mulher sem marido
 “mulher solteira” (Meland & Meland, 2009)

- (12) *Se kfeInsesato jaded^wasato diwnekase*
 [se kfeInese -sato jaded^wa- sato diw -ne -ka -se]
 REF.INDEF jogo -PL menino -PL cansar -CAUS -IND -PASS

“os jogos cansaram os meninos”

Lit: os jogos fizeram os meninos se cansarem

(Da Costa, 1999, p.297)

Os afixos nominais são, por exemplo, usados para marcar parentesco (*e-*) como em *efo'd^wãne* acima para mostrar a relação com o sujeito (aquela sem marido), bem como derivação nominal (*-ke*), clítico possessivo (*ya-*, *wa-*, *tha-* para a 1ª, 2ª e 3ª pessoa do singular), gênero (*-ne* para feminino), privatividade para indicar ausência (*'d^wã*), negação (*-doá*), etc. O Yathé é nominativo-acusativo em sua marcação de argumento central e usa várias partículas discursivas (*ta* para afirmação, *néma* para marcar sequências, *yasa* para marcar causalidade, etc.) (Da Costa, 1999; Silva, 2011; Meland & Meland, 2009).

3.4 Krenak, Borum: Povo e Língua

História do povo Krenak

A história do povo Krenak (autônimo Borum) é marcada por uma resiliência duradoura em meio ao deslocamento colonial e à sobrevivência cultural. Como descendentes dos Botocudos Orientais, os Krenak habitavam o vale superior do Rio Doce (no nordeste de Minas Gerais) antes do contato com os europeus e viviam um estilo de vida semi-nômade de caçadores-coletores centrado no rio Watu (Rio Doce), seu eixo espiritual (Falco Genovez & Santana, 2025). A comunidade principal vive hoje entre os municípios de *Resplendor* e *Conselheiro Pena*, dentro de um Território Indígena reconhecido federalmente com aproximadamente 4.000 hectares. No total, a comunidade Krenak é composta por 600 indivíduos, embora parte deles esteja em comunidades menores em outros estados, como São Paulo e Mato Grosso.

As pressões coloniais e pós-coloniais — incluindo as “Guerras Justas” (nas quais os botocudos foram designados como inimigos da civilização) declaradas pela coroa portuguesa, a construção de ferrovias, a apropriação de terras agrárias, as remoções forçadas e o confinamento em reservas — dizimaram sua população e limitaram seu território. A campanha conhecida como “Guerras Justas” contra os botocudos — dos quais os Krenak (também chamados de Borum) eram um grupo oriental — foi formalmente iniciada pela coroa portuguesa no início do século XIX. Em 1808, logo após a chegada da corte real portuguesa ao Brasil, o príncipe regente Dom João emitiu uma *Carta Régia* declarando guerra aos povos botocudos, citando como justificativa sua suposta resistência à civilização e alegadas práticas antropofágicas (Steward, 1946-1959). Esse conflito se intensificou em todo o Espírito Santo e Minas Gerais a partir desse momento até aproximadamente 1824, culminando com o cerco fracassado de Vitória pelos guerreiros botocudos. Embora a *Carta Régia* tenha lançado uma ofensiva, as hostilidades formais foram oficialmente rescindidas em 1831, com o fim do período da Regência no Brasil (Falco Genovez & Santana, 2025; Steward, 1946–1959).

Parte de seu território foi restaurada em 1997. O rompimento da barragem de rejeitos de Fundão em 2015 liberou resíduos tóxicos de mineração que causaram uma catástrofe no ecossistema do Rio Doce, deslocando ainda mais os Krenak, deteriorando suas estratégias de sobrevivência e estimulando ações judiciais transnacionais, ainda em andamento (Falco Genovez & Santana, 2025).

História da língua Borum

A língua Borum, também conhecida como Krenak, é a última sobrevivente do tronco Botocudo (Borum) da família linguística Macro-Jê, que anteriormente incluía variedades agora extintas, como o Guerén e o Nakrehé (Seki, 1987). O termo Botocudo tem origem na palavra portuguesa *botoque*, que significa tampão, para descrever os adornos de madeira que eles usavam nos lábios e nas orelhas (Seki, 1987). Apesar de sua documentação histórica remontar a encontros no início do século XIX, dados linguísticos abrangentes só surgiram por meio de listas de vocabulário fragmentadas e trabalhos de campo no final do século XX. O trabalho de Lucy Seki no início dos anos 80 contou com quatro falantes nativos para documentar características fonológicas e gramaticais, ressaltando o extremo risco de extinção da língua — com menos de 15 falantes parciais, principalmente idosos, restantes (Seki, 1987). Esforços descritivos anteriores, incluindo o dicionário bilíngue de Rudolph (Rudolph, 1909), surgiram em iniciativas de revitalização mais recentes — como workshops baseados em materiais antigos — para transmitir o conhecimento da língua entre os membros mais jovens da comunidade (Seki, 1987; Pessoa, 2008).

A língua Borum é classificada como moribunda e criticamente ameaçada, sobrevivendo principalmente entre mulheres idosas, e toda a transmissão intergeracional praticamente cessou. As estimativas de falantes fluentes ou semi-fluentes variam consistentemente entre 10 e menos de 20 indivíduos. Dados da UNESCO indicam “uma dúzia de falantes, principalmente mulheres idosas”. O Instituto Max Planck de Antropologia Evolutiva (s.d.) observa aproximadamente 10 falantes em Minas Gerais e no Espírito Santo, com usuários fluentes frequentemente estimados em um número ainda menor (UNESCO, 2010).

A revitalização contemporânea do Borum está focada na recuperação lexical e em recursos orientados para professores, projetados para uso imediato em sala de aula. Um marco é a tese de bacharelado de Frassetto (2018), que reconstrói e padroniza listas de palavras históricas em um dicionário moderno Português-Krenak/Krenak-Português, alinhando as entradas com a IPA e a ortografia atual da comunidade para apoiar a implantação pedagógica nas escolas Krenak. Esse esforço acadêmico, juntamente com uma compilação endossada pela comunidade, o *Vocabulário Unificado Português-Krenak (Botocudo)/Krenak-Português* (2019), produzida com parceiros como a ONG KAMURI, a FUNAI e o grupo InDIOMAS da UNICAMP, e disponibilizada através do Museu Índia Vanuíre (2019) para fortalecer o planejamento de aulas e atividades de alfabetização. Juntos, esses recursos abordam diretamente a escassez crônica de materiais lexicais utilizáveis por professores e alunos, ancorando a revitalização na prática em sala de aula e no controle comunitário das normas ortográficas (Frassetto, 2018; Museu Índia Vanuíre, 2019).

Um segundo pilar é a pedagogia digital orientada para a comunidade. A plataforma web *Kjeme Itchok: “Casa da Língua Krenak”* (Itchok, s.d.) organiza tutoriais sobre fonologia, morfossintaxe e formação de frases, e compila dissertações, dicionários e gramáticas com base em consultas à comunidade que priorizaram o amplo acesso ao conhecimento existente em detrimento da obtenção de novos dados. A documentação pública destaca o plano da plataforma para um dicionário interativo e treinamentos de professores, transformando um repositório vivo em um ambiente de aprendizagem participativo para famílias Krenak dispersas (Silva, 2024; Itchok, s.d.). Ao posicionar os educadores e detentores de conhecimento Krenak no centro das decisões de conteúdo, a plataforma incorpora os princípios das melhores práticas de revitalização — propriedade local, acesso aberto e co-design iterativo — ao mesmo tempo em que reduz as barreiras geográficas à participação (Silva, 2024; Itchok, s.d.).

O trabalho de revitalização também é apoiado por redes regionais e programação cultural que integram o ensino de idiomas com práticas de identidade e colaboração intercomunitária. O dossiê

“Levante Linguístico Indígena” documenta como as redes indígenas no Nordeste, Espírito Santo e Minas Gerais se mobilizam para o ressurgimento linguístico, vinculando iniciativas em sala de aula a rituais comunitários e estruturas de transmissão intergeracional (Bonfim et al., 2021). Atividades paralelas baseadas em museus — exposições, acesso a coleções de referência e apoio ao desenvolvimento de professores indígenas — ampliam os contextos em que o Borum é aprendido e usado, aumentando assim as oportunidades de ativação da língua além da escolaridade formal (Museu Índia Vanuíre, 2019; Bonfim et al., 2021).

Por fim, as mensagens de revitalização aproveitam a defesa pública que enquadra a língua como uma tecnologia central para a sobrevivência cultural e a gestão ecológica. A cobertura das iniciativas lideradas por intelectuais Krenak enfatiza práticas — canções, performances, eventos comunitários — que reintegram a língua na vida cultural cotidiana e motivam a participação de jovens e adultos. Essa programação está alinhada com a Década Internacional das Línguas Indígenas da UNESCO (2022-2032) e promove a solidariedade entre as comunidades em torno dos objetivos de revitalização, complementando as estratégias digitais e em sala de aula ao sustentar os domínios simbólicos e experienciais do uso da língua (Bonfim et al., 2021).

Os primeiros materiais linguísticos relacionados ao Borum datam do início do século XIX. Considera-se que a primeira documentação do Krenak foi realizada pelo príncipe Wied-Neuwied, que veio da Alemanha para viajar pelo Brasil (precedido apenas por um vocabulário muito pequeno anotado por Langsdorff, o e a russo cônsul no Rio de Janeiro (Feest, 2022) e, certamente, pelo vocabulário de Marlière (1825); no entanto, este último é inédito). O conjunto de materiais existentes sobre o Krenak foi produzido por uma variedade de pessoas de diferentes nacionalidades (alemães, russos, franceses, brasileiros, americanos...), com várias profissões (biólogos, farmacêuticos, geólogos, antropólogos, naturalistas, engenheiros, clérigos, etnólogos, militares, linguistas...) e por várias razões e interesses. Devido a essa diversidade de autores, os documentos estão bastante dispersos e a qualidade de muitos deles é questionável, uma vez que vários autores procuraram encontrar no Krenak as mesmas características das línguas indo-europeias (Frassetto, 2018).

A atenção acadêmica ao Borum (Krenak) tem sido muito limitada, sendo uma das línguas indígenas menos estudadas do Brasil. Parece haver vários fatores convergentes que resultaram nessa falta de interesse acadêmico pela língua. Primeiro, o Borum chegou a um estado de grave risco de extinção no final do século XX, com apenas um punhado de falantes idosos espalhados por Minas Gerais e São Paulo, o que tornou o trabalho de campo de longo prazo e a descrição abrangente excepcionalmente difíceis (Seki, 2004). Segundo, a violência estatal, as remoções forçadas e a repressão cultural, incluindo a proibição do uso de línguas indígenas durante a ditadura militar brasileira, prejudicaram a transmissão intergeracional e desestimularam o envolvimento com pesquisadores externos, limitando assim as oportunidades para a documentação (Bonfim et al., 2021). Em terceiro lugar, os materiais iniciais eram léxicos fragmentados, em vez de gramáticas — as fontes do século XIX priorizavam listas de palavras para áreas específicas com transcrição inconsistente —, deixando os linguistas modernos a reconstruir padrões gramaticais a partir de dados escassos e sob pressão de tempo (Seki, 2004). Quarto, barreiras geográficas e institucionais (infraestrutura limitada, base pequena de falantes e financiamento historicamente escasso para a documentação de microcomunidades) reduziram os incentivos e a viabilidade de projetos sustentáveis, mesmo quando características tipologicamente significativas — por exemplo, ordem SOV, pós-posições extensas e efeitos de contato com o Português — foram observadas em breves esboços (Seki, 2004). Por fim, embora os recentes esforços de revitalização e lexicografia tenham melhorado o acesso aos materiais, eles chegaram tarde na trajetória de extinção, de modo que os registros acadêmicos ainda se inclinam para a morfossintaxe e a fonologia, em vez de gramáticas descritivas completas (Frassetto, 2018; Bonfim et al., 2021).

Alguns parâmetros tipológicos

A ordem padrão das palavras em Borum é consistentemente SOV, como mostrado no exemplo (13).

- (13) [nãŋtõndɔn kuparak pip]
 criança onça-pintada ver.PST
 “a criança viu a onça” (Seki, 2004, p.134)

As frases equativas são tratadas usando duas frases nominais não marcadas, uma das quais tem uma função predicativa e a outra serve como sujeito da predicação, como mostrado no exemplo (14).

- (14) [ŋgi-tʃukɪ n kwaha ŋɪ m]
 1SG-marido homem este/aqui
 “meu marido é este homem” (Seki, 2004, p.134)

A posse pode ser expressa de diferentes maneiras, os exemplos (15) e (16) mostram o uso do morfema [wãŋ], que é sempre colocado após o possuidor. A ordem possuidor-posse também pode ser vista no exemplo (14) com a posse do marido.

- (15) [ŋg-wãŋ ŋɔŋ]
 1SG-POSS cão
 “Eu tenho um cachorro / o cachorro é meu” (Seki, 2004, p.136)

- (16) [ŋɔŋ ŋɪ m nãŋtõndɔn wãŋ]
 cão esta criança POS
 “este cachorro é da criança” (Seki, 2004, p.136)

A negação é sempre expressa com o morfema *-nuk*, como mostrado no exemplo (17).

- (17) [ti ŋgri -nuk nɛʔ]
 eu canto NEG FUT
 “Não vou cantar” (Seki, 2004, p. 133)

Como os estudos sobre o Borum começaram muito tarde, quando já havia poucos falantes nativos vivos, ainda hoje há muitas incertezas a serem esclarecidas, especialmente no que diz respeito ao sistema de codificação de pessoas. Sabe-se que, tipologicamente, o sistema pronominal nas línguas Macro-Jê é complexo e, além disso, como o contato com os últimos falantes nativos de Borum foi breve e havia grande variação na pronúncia de certos itens, ainda não foi possível obter textos coesos e, frequentemente, os falantes tinham dificuldade em dar o equivalente da afirmação solicitada em Português. Seki (2004) explica essas dificuldades e dá o exemplo de que, ao pedir o equivalente a “você me viu”, recebemos o equivalente a “eu te vi” ou até mesmo uma tradução da pergunta.

3.5 Xukuru: povo e língua

História do povo Xukuru

O povo Xukuru vive hoje principalmente na Serra de Ororubá, em Pesqueira, no estado de Pernambuco. Há outro grupo étnico que vive em Palmeira dos Índios, no estado de Alagoas, chamado Xukuru-Kariri, porém sua língua é da família Kariri. A origem desse grupo é muito provavelmente resultado do deslocamento forçado de ambos os grupos, Xukuru e Kariri, para uma comunidade comum. A comunidade Xukuru de interesse neste artigo é a da Serra de Ororubá. Atualmente, essa

comunidade é composta por cerca de 12.000 indivíduos em 24 *aldeias* diferentes (das quais cerca de 200 indivíduos se mudaram para a cidade de Pesqueira), o que a torna um dos maiores grupos indígenas do Nordeste (Silva E.H., 2007).

Em 1879, o Governo Imperial decretou oficialmente a extinção da aldeia missionária de Cimbres. Esse ato administrativo foi usado localmente para afirmar que “não havia mais índios, apenas caboclos”, e abriu caminho para reivindicações de terras privadas sobre o território Xukuru na Serra do Ororubá. Fugindo da perseguição, as famílias Xukuru se dispersaram pela região ou foram morar nas terras de outros antigos assentamentos e na periferia das cidades. Outras famílias, que eram discriminadas e chamadas de “*caboclos de Orubá*”, resistiram em pequenos lotes de terra, fazendas em locais de difícil acesso, ou continuaram trabalhando em suas próprias terras, exploradas como mão de obra pelos proprietários invasores. Seus rituais religiosos tradicionais continuaram em segredo na mesma região (Silva, 2007). Durante as primeiras décadas do século^{xx}, os Xukuru começaram a reivindicar a terra com mais força e, em 1954, um Posto do SPI (*Serviço de Proteção aos Índios*) foi estabelecido na Serra de Ororubá. Uma vitória para os Xukuru, mas que não resolveu os conflitos com os fazendeiros que haviam invadido suas terras. Isso porque a política oficial da agência indigenista para o Nordeste utilizava a falta de provas documentais/legais dos territórios indígenas para resolver o conflito, seja reduzindo as áreas originais, seja adquirindo pequenos lotes de terra para as comunidades indígenas. No final da década de 1980 e início da década de 1990, surgiu um líder na luta pela terra, o cacique Xicão, que conseguiu consolidar a luta da comunidade, em grande parte vinculando-a ao movimento de reafirmação da identidade Xukuru, que havia sido suprimida por mais de um século sob a narrativa da “extinção”. Ele organizou assembleias comunitárias e revitalizou práticas tradicionais, tornando a identidade cultural a base da luta pela terra. Xicão liderou a campanha jurídica e política pelo reconhecimento da *Terra Indígena Xukuru*. Ele trabalhou em estreita colaboração com a FUNAI, organizações indígenas e aliados para promover estudos e o reconhecimento oficial. Sua liderança transformou a causa Xukuru em um símbolo nacional da resistência indígena ao desafiar abertamente os poderosos proprietários de terras que ocupavam as terras Xukuru, o que o tornou alvo de ameaças e violência. Apesar dos riscos, ele manteve uma estratégia de mobilização pacífica combinada com ações legais, mas foi tragicamente assassinado em maio de 1998, durante o auge do conflito pela terra. Sua morte atraiu a atenção nacional e internacional, acelerando o processo de demarcação. Finalmente, em 2001, a TI Xukuru foi oficialmente homologada — uma vitória amplamente atribuída à sua liderança e sacrifício (Silva, 2007; Hohenthal, 1954; Santana & Magalhães, 2022).

História da língua Xukuru

A língua original do povo Xukuru é considerada extinta e pouco documentada. A afiliação genética do Xukuru permanece provisória devido à extinção da língua e à escassa documentação, mas a maioria dos estudiosos a coloca provisoriamente dentro da esfera Macro-Jê com base em evidências lexicais limitadas e padrões de distribuição regional. As primeiras compilações (por exemplo, Pompeu Sobrinho, 1958) e o único esboço descritivo, Lapenda (1962), fornecem listas curtas de palavras e notas fonológicas que serviram de base para classificações posteriores; no entanto, esses materiais são fragmentários e, às vezes, misturados com itens de repertórios vizinhos (Kariri, Pankararu, Yatê e Língua Geral), o que complica a comparação. Com base nessas listas, Loukotka (1968) propôs uma microfamília “Xukuruan” ligando Xukuru e Paratió (e provisoriamente “Garanhun”), um arranjo mantido (com cautela) em catálogos modernos como o Glottolog, que tratam a filiação e o status de ameaça de extinção do Xukuru como não resolvidos, mas provavelmente pertencentes ao Macro-Jê (Loukotka, 1968). Dada a ausência de textos contínuos, morfologia robusta e dados sintáticos, a afiliação não pode ser demonstrada além das correspondências lexicais; a visão predominante é que o Xukuru está relacionado ao Macro-Jê — possivelmente mais próximo do Kariri —, mas ainda não

pode ser classificado com certeza, dependendo da descoberta de fontes primárias adicionais (Lapenda, 1962; Pompeu Sobrinho, 1958).

As políticas coloniais indígenas e a assimilação forçada levaram à rápida perda da língua Xukuru no final do século XIX, deixando apenas listas fragmentadas de palavras registradas por missionários e viajantes. Os últimos falantes nativos faleceram em meados do século XX. O que foi preservado da língua ancestral estava principalmente nos rituais religiosos secretos, como é o caso de muitas comunidades indígenas do Nordeste do Brasil. Hoje, os Xukuru falam Português como sua L1, e sua identidade cultural é mantida por meio de tradições orais e práticas rituais, em vez de uma língua indígena viva. Os esforços contemporâneos para revitalizar a língua acabaram se concentrando no renascimento simbólico, em vez da reconstrução linguística. A perda da língua ancestral para a comunicação diária continua sendo uma fonte de angústia entre os Xukuru (Carvalho, 2019). Os esforços que estão sendo realizados envolvem, entre outros, a reintrodução de palavras, nomes e expressões ancestrais nas práticas culturais, rituais e materiais educacionais como marcadores de identidade e herança. Essa abordagem fortalece a memória coletiva e afirma a continuidade étnica, mesmo quando a língua não pode ser restaurada como meio de comunicação cotidiana (Silva, 2007).

Alguns parâmetros tipológicos

Lapenda (1962) é um esforço para documentar a língua, mas ele descobriu, por exemplo, que não há vestígios de pronomes pessoais nem de como a posse ou os demonstrativos são realizados. O mesmo foi encontrado para conjunções e interjeições. Um possível advérbio de negação “biá” foi observado e um “i” nasalizado (do Português “em”) parece ter sido incluído como um advérbio temporal, como nos exemplos (18) e (19)

(18) *nõyẽ* *biá*
falar não Lapenda (1962, p. 16)

(19) *i(n)* *bemen*
em a manhã Lapenda (1962, p.16)

Alguns sufixos foram encontrados, porém com função incerta, como *go* em muitos verbos, tais como *arágo* (matar), *montógo* (correr) e *kringó* (comer), e também em alguns substantivos verbais, tais como *xukégo* (ladrão) e *jubégo* (feiticeiro).

3.6 Estrutura numeral nas línguas

As estruturas numéricas variam significativamente entre as culturas, incluindo até onde se estendem os sistemas de contagem e qual base numérica é usada para formar os números (geralmente cinco ou dez). Em muitas línguas, a mão e os dedos servem como base para a contagem; no entanto, a estrutura ainda pode diferir muito dependendo de fatores como se os polegares são incluídos, se os dedos são adicionados ou subtraídos. Outra diferença importante é se cada dedo conta como um ou três, contando os três falanges⁹. Isso, juntamente com o uso do polegar para contar, é a base dos sistemas numéricos duodecimais, presentes principalmente na Ásia e na África (Huylebrouck, 2010)).

⁹ O sistema duodecimal (base 12) ainda tem alguns vestígios históricos em muitas culturas indo-europeias, como na compra de ovos por dúzia e em palavras como *grosa* (12x12=144).

Muitas línguas — como o Mundurukú da Bacia Amazônica e o Krenak, uma das línguas em foco neste estudo — têm palavras apenas para números até três ou quatro (Instituto Max Planck de Antropologia Evolutiva, s.d.). Além desses, são usadas expressões como “*um punhado*” ou “*muitos*” (Overmann, 2021). Tanto o Krenak quanto o Mundurukú agora usam os números em Português a partir do quatro, pois a participação na sociedade moderna criou a necessidade de expressar quantidades maiores (Calude, 2021). No entanto, adotar a língua majoritária não é a única estratégia disponível quando surgem tais demandas externas. Isso torna particularmente interessante examinar quais estratégias de criação de palavras diferentes línguas em processo de revitalização adotaram para os numerais em seus respectivos contextos socioculturais.

4. Métodos

Para examinar como os neologismos são formados nas cinco línguas em revitalização do Nordeste do Brasil, adotei uma metodologia em várias etapas, combinando diferentes fontes de dados e análise estrutural dos neologismos identificados nessas línguas. Quanto às fontes de dados, tive que recorrer a fontes diferentes para cada língua, com base em sua disponibilidade e diferentes níveis de acesso. Para cada língua, selecionei termos tanto de materiais de dicionários (contemporâneos, quando disponíveis, e históricos, quando não disponíveis) quanto de textos escritos recentemente. Em muitos casos, alguns neologismos também foram publicados por pesquisadores e/ou linguistas indígenas ao descrever a língua. Nesses casos, as publicações incluem informações sobre como novas palavras são criadas pelos grupos ou indivíduos envolvidos nesse trabalho, e incluí suas descobertas como classificação.

Também incluí palavras relacionadas a cuidados de saúde, educação e tecnologia, quando disponíveis, uma vez que essas são áreas em que a necessidade de novo vocabulário surgiu claramente desde o início da colonização. Meu objetivo era ter mais de dez neologismos com processos de cunhagem identificados para cada idioma.

Os numerais foram incluídos no estudo, pois fornecem algumas informações adicionais sobre o tipo de processos criativos que estão em vigor para cada língua. Os sistemas numéricos são quase sempre abordados em projetos de revitalização e são especiais por pertencerem, ao mesmo tempo, a uma área semântica natural e a uma ciência criada pelo homem (matemática, historicamente muitas vezes baseada em uma lógica decimal, etc.). Uma razão pela qual os numerais são quase sempre abordados em projetos de revitalização é, na maioria das vezes, para valorizar melhor a língua. Campos (2023, p. 109) elabora sobre isso:

Do ponto de vista prático, não haveria problema em uma língua usar os numerais portugueses, mas, da perspectiva da valorização da língua, ter um vocabulário próprio para discutir qualquer assunto, no campo da matemática, por exemplo, parece fazer toda a diferença, especialmente para os jovens. Se imaginarmos uma situação em que uma pessoa não indígena pergunta a um falante de Maxakali como se diz “dois reais” em sua língua, é possível que seja constrangedor para o falante de Maxakali ter que explicar que a expressão em Maxakali é praticamente a mesma que em Português, o que poderia criar a falsa impressão de que o Maxakali é menos eficiente ou menos importante que o Português. Da mesma forma, poderia ser constrangedor dizer que os números na língua são os mesmos que os usados em Português.

Uma lista detalhada das fontes encontra-se no Apêndice A.

Codificação de neologismos

Meu processo de codificação dos neologismos de acordo com os processos de cunhagem seguiu este protocolo:

Primeiro, baseei-me em classificações já estabelecidas na literatura acadêmica ou em projetos de revitalização liderados pela comunidade. Nos casos em que tais classificações não estavam disponíveis, usei as informações sobre formação de substantivos que havia reunido sobre cada língua para examinar sistematicamente os dicionários existentes, como o dicionário Maxakali existente (Popovich & Popovich, 2005), que usei como fonte primária para encontrar os processos nos neologismos Patxohã. Isso me permitiu rastrear as origens dos morfemas relevantes e, assim, identificar os processos pelos quais os neologismos foram criados. Nesse processo, é claro que houve vários

neologismos em que não consegui encontrar um processo de cunhagem razoável. Todos esses neologismos foram descartados no processo.

Um exemplo desse processo seria o neologismo Yathê *txhatkhweaiekdonkya*, que significa “avião”. Sabendo que o Yathê é uma língua altamente flexional, procurei palavras semelhantes para tentar identificar flexões. Por exemplo, a parte final *ekdonkya* é a mesma em *oyathaiekdonkya* (canoa). Em várias outras palavras, o morfema feminino *nkya* é usado com um verbo. Descobri que *Ek* é a forma 3SG de *ekkya* (correr) e que *doa* é um morfema usado para nominalização em outras palavras compostas. A parte inicial da palavra é encontrada nos dicionários Yathê com o significado de “céu”. Portanto, os processos de cunhagem utilizados são principalmente o encurtamento e a derivação. *Nãdudya*, também Yathê, que significa inovação, foi formado de uma maneira diferente. *Dudyaa* é um adjetivo que significa “estranho” e o morfema inicial *nã* é a raiz do verbo *nãkya*, que significa “introduzir/mostrar”. Portanto, esse neologismo também foi criado usando o processo de truncamento, mas também está mudando a função de verbo/adjetivo para substantivo, por isso é combinado com o processo de conversão. Outro exemplo são as palavras *tapitébm'aô* (geladeira), *patxéhiáb* (TV) e *patxékahab* (caderno), onde tive que iniciar a análise comparando uma página das versões em Patxohã e Português da Cartilha Pataxó (Ministério de Minas e Energia (s.d.)) para encontrar as palavras em Patxohã. Com as palavras identificadas, pude então continuar com o mesmo processo dos exemplos em Yathê acima.

Classificação dos neologismos

Classifiquei o processo de cunhagem de cada neologismo nas seguintes categorias (exceto para os numerais, onde optei por descrever o processo mais detalhadamente para cada língua). As classificações podem ser vistas no capítulo 4, em cada língua.

1. Emprestado da língua majoritária
2. Emprestado da língua indígena vizinha
3. Composto
4. Derivação
5. Palavra-valise
6. Conversão
7. Truncamento
8. Acrônimos
9. Resignificação semântica
10. Extensão semântica

Qualquer substantivo cujo processo de criação seja uma combinação de dois ou mais processos é contado tantas vezes quantos os processos combinados.

Este modelo de agrupamento de palavras baseia-se em (Bauer, 1983) e (Alves 1990) no que diz respeito a neologismos sintáticos, e eu acrescento a divisão entre 1 e 2 e adiciono 9 e 10 de acordo com a minha motivação descrita anteriormente.

5. Resultados

5.1 Patxohã

O processo mais utilizado em Patxohã é, de longe, o empréstimo do vizinho Maxakali. A amostra não dá uma imagem fiel disso, pois eu, propositalmente, tentei identificar quaisquer outros processos utilizados também. A lista completa pode ser encontrada na Tabela 5. Alguns desses outros processos de cunhagem refletem o processo de formação de palavras em Maxakali, como a derivação utilizada na criação de *guyãô* (fonte de água) abaixo.

Palavra	Português	Princípios de formação	Descrição da formação
ipakâié	Professor	Palavra-valise Empréstimo de língua vizinha (Maxikali) Composto	<i>Ipamakã</i> (pai) + <i>akâiéko</i> (líder) = <i>ipakâié</i> (professor)
kijëtxawê	Escola	Palavra-valise Empréstimo da língua vizinha (Maxikali) Composto	<i>Kijemi</i> (casa) + <i>etxawê</i> (ensinar) = <i>kijëtxawê</i> (escola)
Patxohã	A língua Patxohã, língua do guerreiro	Palavra-valise	<i>Patxó</i> (o povo patoxa) + <i>hã</i> (falar) (às vezes, a palavra <i>xohã</i> (guerreiro) é incluída na explicação)
miãga	água	Empréstimo da língua vizinha (Maxikali)	Havia várias palavras disponíveis para designar água, mas <i>miãga</i> foi a escolhida. Outras, como <i>txonãg</i> , tiveram seu significado alterado/ampliado para água da chuva
maion	luz	Empréstimo de línguas vizinhas (sol em Maxakali: <i>mãÿõn</i>)	<i>mãÿõn</i> = sol em Maxakali
tapitébm'aô	geladeira	Empréstimo da língua vizinha (maxikali) Derivação	<i>tapitébm</i> = clima frio/gelo (clima frio <i>xaptep-tex</i> em Maxakali). - <i>aô</i> é um sufixo derivacional (agente), <i>tapitébm'aô</i> - geladeira (agora <i>tapitébm</i> = gelo)
patxéhiáb	televisão	Composto Extensão semântica	<i>patxé</i> = ferramenta/janela, <i>hiáb</i> = terra/concepção
patxékahab	caderno	Composto Extensão semântica	<i>patxé</i> = ferramenta/janela, <i>kahab</i> = calcular/contar/viver (<i>kahãbá</i> = dinheiro)
nõte ùpú enuay	agente de saúde	Empréstimo (calque) da língua majoritária	<i>assistente</i> = <i>nõte</i> (ajudar em Maxakali: <i>nõ tehe</i> , <i>ùpú</i> = <i>de</i> , <i>enuãÿ</i> =saúde)

TABELA 6 LISTA DE NEOLOGISMOS EM PATXOHÃ

Palavra	Português	Princípios de formação	Descrição da formação
guyãô	bebedouro	Derivação	<i>goyá</i> = beber. -aô é um sufixo derivacional (agente), <i>goyáô</i> - bebedouro
jokana korihé	enfermeira	Empréstimo (calque) da língua majoritária	<i>jokana</i> = mulher, <i>korihé</i> = <i>cuidar</i> (cuidar de)
ĩdhixi	não indígena	Derivação	ĩd - negação + <i>thixi</i> - pessoa indígena
Atxôhã imamakã	língua materna	Empréstimo (calque) da língua majoritária	<i>Atxôhã</i> - língua, <i>imamakã</i> - mãe

TABELA 6 (CONTINUAÇÃO)

Como a língua havia sido silenciada e o projeto de revitalização começou com uma base de aproximadamente 250 palavras, a maioria das palavras em Patxohã são, por definição, neologismos. Fica claro, a partir de Bomfim (2017), que o vocabulário Maxakali influenciou amplamente o trabalho. Usei o dicionário Maxakali existente (Popovich & Popovich, 2005) como base para encontrar as possíveis fontes dos neologismos. Isso, juntamente com a estrutura básica dos sufixos e comparações com a forma como outros compostos foram criados, me deu a maioria das respostas para os processos de neologismo. O processo para duas das palavras (escola e professor) foi divulgado pelo grupo Atxohã (ver nota de rodapé, p. 26).

Embora os Paxotó vivam em um ambiente de língua majoritariamente portuguesa, há muito poucos exemplos de empréstimos da língua portuguesa (excluindo a estrutura da língua, gramática etc.), mas parece que “Matemática” foi integrada como uma palavra Patxohã, e as organizações administrativas oficiais mantiveram seus nomes portugueses também em Patxohã. O que é comum, porém, é o empréstimo por calque da construção genitiva “de” do Português. Uma construção típica do nome em Português com o genitivo está sendo calcada como no nome do prato tradicional de farinha de milho com caranguejo: *angu de caranguejo*. No livro de receitas bilíngue “*Fappet Mägute Pataxó*” (livreto de comida Pataxó desenvolvido por mulheres Pataxó na aldeia de Cumuruxatiba em 2010-2011), foi dado o nome *de Angu ũpú Kãbará* (de Tassio Borges da Silva, 2019). *Angu* é uma palavra de origem africana (provavelmente do Fon, uma língua Gbe da família linguística Níger-Congo ainda falada principalmente no Benim) (Pereira Serra & de Maria de Araujo Ramos, 2021), “*ũpú*” é a palavra Patxohã para “de” e *Kãbará* significa “caranguejo” em Patxohã.

Empréstimos de uma língua indígena vizinha, palavra-valises, derivações e composições são usados para criar neologismos em Patxohã. Não encontrei nenhum exemplo de acrônimos, truncamentos ou conversões.

No entanto, os processos de neologia semântica são frequentemente misturados com os acima mencionados, tanto como extensão semântica do significado como, num caso, resignificação semântica. Um exemplo é o processo de escolha entre várias palavras existentes para “água”: a comunidade decidiu usar *miãga* e *txiäng* foi resignificado para significar apenas água da chuva. Não incluí este neologismo, pois a escolha deliberada está, de qualquer forma, um pouco fora da definição normal deste processo.

Um exemplo de extensão é a palavra composta para televisão *patxéhiáb*, uma combinação das palavras para janela e terra, e outro exemplo é quando, no processo de descrição da língua, o grupo

Atxohã (ver nota de rodapé p. 26) precisava de um termo para descrever o alfabeto e decidiu estender o significado da palavra *Atxūhū* (*linguajar*, dialeto).

Numerais

Apenas o número “um” foi encontrado no material original (Wied-Neuwied, 1989). Os outros números, de “dois” a “dez”, foram criados pela comunidade (no entanto, sem usar numerais de nenhuma língua intimamente relacionada, e a fonte utilizada não foi revelada pelo grupo Atxohã¹⁰). A palavra *apêtxienã* (um) é a palavra original adaptada à ortografia atual em Patxohã. Os primeiros dez números são apresentados na tabela 6.

Nº	Palavra	N	Palavra
0	Txuĩba	5	Nigrê
1	Apêtxienã	6	Patxiá
2	Kroxxí	7	Dxãdxê
3	Mitxê	8	Nitxê
4	Rãtxê	9	Rawata
		10	Rótsa

TABELA 7 NÚMEROS EM PATXOHÃ

O sistema decimal foi adaptado para outros números, 11 – 19 é *Ruê* + o número correspondente 1-9, portanto, onze é *ruê Apêtxienã*. Os números acima de 20 são formados por um prefixo (*krã-*) e o múltiplo de dez e, finalmente, o numeral final (Bomfim & Lima, 2020). Isso significa que a estrutura dos números de 11 a 19, bem como acima de 20, apresenta todas as características da calque da formação portuguesa dos numerais.

Por exemplo, cinquenta e nove:

Krãnigrê rawata, onde *nigrê* é cinco e *rawata* é nove, de acordo com a tabela acima.

Os numerais ordinais entre 1º e 10º têm origens completamente diferentes dos numerais (por exemplo, o primeiro é: *Pukuixê*, mas a partir do 11º eles seguem a mesma lógica decimal dos numerais cardinais acima com prefixos. Além disso, eles seguem uma estrutura de calque baseada no Português.

¹⁰ São necessários mais estudos para determinar a estratégia utilizada pela comunidade Pataxo no processo de criação de palavras numéricas além do um. Esses numerais não são paralelos aos numerais na língua mais próxima” (Bomfim & Lima, 2020)

5.2 Kariri-Xocó

A composição é o processo de criação de palavras mais utilizado em Kariri-Xocó, seguido pela palavra-valise, como pode ser visto na Tabela 7. A língua utilizou muito poucos empréstimos e neologismos semânticos.

Palavra	Português	Princípios de formação	Descrição da formação
Nhinhodeá	Mulheres Indígenas	Derivação	<i>Nhinho</i> (indígena) + <i>de</i> (sufixo plural) + <i>á</i> (sufixo feminino)
Erayõnrakua	Biblioteca	Composto Palavra-valise	<i>Era</i> (casa) + <i>aerãkuara</i> (livro), (<i>Aerãkuara</i> também é uma composição: <i>aerã</i> (folha) + <i>kuará</i> (listra))
Dzutsáko	Café		<i>Dzu</i> (água) + <i>tsá</i> (fogo) + <i>ko</i> (<i>dço</i> - preto)
Ipoacró	Óculos	Composto	<i>Ipoá</i> (óculos) + <i>cró</i> (pedra)
Dzunhicró	Geladeira	Composto Palavra-valise	<i>Dzu</i> (água) + <i>nhi</i> (derivado de <i>cunhi</i> (frio)) + <i>cró</i> (pedra)
Memará	Telefone	Composto Extensão semântica	<i>Mé</i> (conversa) + <i>mará</i> (ouvido)
Waruamé	Televisão	Composto Extensão semântica	<i>Waruá</i> (espelho) + <i>mé</i> (conversa)
Subatekerá	Escola	Palavra-valise	<i>Subatekié</i> (conhecimento) + <i>era</i> (casa)
Erawodzo	Hospital	Composição Truncamento	<i>Era</i> (casa) + <i>wodzodzó</i> (<i>cura</i>)
Wyracrotã	Drone	Composição Derivação	<i>Uyrá</i> (pássaro) + <i>cró</i> (pedra) + <i>-tã</i> (resistente, duro)
Tonadzó	Sanduíche	Palavra-valise Composto	<i>Vtona</i> (pão) + <i>cradzó</i> (carne)
Carneiro	Carneiro	Empréstimo da língua majoritária	Empréstimo direto do Português “carneiro”
Ibirakuá	Caneta	Palavra-valise	<i>ibirá</i> (árvore, galho) + <i>kuará</i> (listra)

TABELA 8 LISTA DE NEOLOGISMOS EM KARIRI-XOCÓ

A base principal para a criação de neologismos em Kariri-Xocó, bem como alguns dos dados, está disponível online, principalmente no blog mantido por Nhenety Kariri-Xocó. O processo de criação de neologismos se alinha muito ao processo tradicional de formação de substantivos em Dzubukuá, usando afixos existentes, e também reflete como a composição é produzida na língua base. Um exemplo disso é (embora eu não tenha nenhuma glosa, o uso de acentos graves e agudos dá algumas indicações) a palavra *Ibirakuá* (“lápiz”).

[*Ibirá* + *kuará*] -> [*Ibirá* + *rákua*] -> [*Ibirakuá*].

Onde a inversão *kuará* -> *rákua* e a eliminação do acento médio e a repetição da mesma coda (*rá*) seguem a formação de substantivos existente (como onde [*subatekié + erá*] -> [*subatekerá*], “casa do conhecimento” (mais tarde também significando escola)

Encontrei alguns empréstimos diretos do Português, aparentemente da parte inicial da colonização, como alguns animais importados (“carneiro” (Português *carneiro*, Kariri-Xocó: *carneiru*), “cabra” (Português *cabra*, Kariri-Xocó: *cabara*).

Os exemplos de extensão semântica que encontrei parecem concentrar-se apenas em novas tecnologias, além disso, a mistura de composição, fusão e truncamento parece ser o processo dominante de criação de palavras em Kariri-Xocó.

A sintaxe e a gramática do Kariri-Xocó foram construídas com base em uma gramática bastante estável do dzbukúá e mantiveram a maioria de suas especificidades. No entanto, para ter uma visão melhor de quanto da influência do Português aparece nos neologismos produzidos em Kariri-xocó, seria necessário um estudo no local. Isso também poderia dar uma visão melhor das influências fonológicas, tanto do Português quanto das línguas indígenas vizinhas, especialmente aquelas do mesmo tronco linguístico, como o Dzbukúá revitalizado.

Numerais

O Kariri-Xocó encontrou as palavras do dzbukúá para os cinco primeiros números e construiu uma estrutura quinairial a partir disso até 10, de acordo com o seguinte, com um sistema decimal começando em 11, conforme mostrado na tabela 8.

Nº	Palavra	Nº	Palavra	Nº	Palavra
1	<i>hæ</i> (< <i>bihe</i>)	10	<i>bœ</i> (< <i>bœ todos</i>)	20	<i>wibœ</i>
2	<i>wi</i> (< <i>witane</i>)	11	<i>bœhœ</i>	30	<i>kebœ</i>
3	<i>ke</i> (< <i>witanedike</i>)	12	<i>bœwi</i>	40	<i>subœ</i>
4	<i>hu</i> (< <i>humarã anrobœ</i> , imitando o equivalente Kipeá <i>sumarã orobae</i>)	13	<i>bœke</i>	50	<i>muihbœ</i>
5	<i>mæ</i> (< <i>mædha</i> mão)	14	<i>bœdhù</i>	60	<i>muihbœ</i>
	Começando com o número seis, são adições à “mão completa” em <i>mæ</i>	15	<i>bœmuidhù</i>	70	<i>muihbœ</i>
6	<i>muihbœ</i> (< <i>mui bihe moedha hã</i> carregando um e uma mão para ele)	16	<i>bœmuidhù</i>	80	<i>muihbœ</i>
7	<i>muiwi</i> (< <i>mui witane moedha hã</i> carregando dois e uma mão para ele)	17	<i>bœmuidhù</i>	90	<i>muidhubœ</i>
8	<i>muike</i> (< <i>mui witanedike moedha hã</i> carregando três e uma mão para ele)	18	<i>bœmuidhù</i>	100	<i>ye</i> (lit.: grande)
9	<i>muidhù</i> (< <i>mui hamarã anrobœ moedha hã</i> carregando quatro e uma mão para ele)	19	<i>bœmuidhù</i>	1000	<i>hodè</i> (lit: muito ser)

TABELA 9 NUMERAIS EM KARIRI-XOCÓ (SILVA 2021)

O sistema de 1 a 10 é bastante semelhante ao da comunidade Desana da Amazônia brasileira, cujos termos para “cinco” e “dez”, *yuhuru mōhōtō* (“uma mão”) e *pemōhōtō* (“duas mãos”), são

acompanhados por termos correspondentes para os dedos restantes; o termo para “seis”, por exemplo, é *yuhuru mōhōtō yuhuru niãã*, “uma mão, uma da outra mão” (Overmann, 2021; Silva, 2012). O sistema numérico Desana vai além em sua lógica quina e inclui também os dedos dos pés, continuando um sistema de contagem com partes do corpo até 20. Aqui temos outro exemplo de uso da língua de origem e adaptação ao novo contexto linguístico apenas pela expansão do sistema numérico existente.

5.3 Yathé

O empréstimo do Português e o truncamento são os dois processos de cunhagem mais utilizados em Yathê, muitas vezes em combinação com outros processos. A lista completa dos neologismos pode ser encontrada na tabela 9.

Palavra	Português	Princípios de formação	Descrição da formação
,tape'tiwa	Tapetinho	Empréstimo de língua majoritária (calque) Derivação	,tape'ti (tapete) + wa (diminutivo)
banko	banco	Empréstimo da maioria língua	
doãthoyane	televisão	Composto Extensão semântica	dofa (espelho, imagem) + thoya + ne (FEM, afetivo): possivelmente nosso espelho íntimo (Xavier de Oliveira, 2018)
sprita	hospital	Empréstimo da maioria língua	
mestle	professor	Empréstimo da maioria língua	de <i>Mestle</i> (Pt. <i>mestre</i> - professor)
Fowaxia-neho	Geladeira	Truncamento Composição	Fowa (pedra) + exia (frio) + neho (INSTR)
Towenese	Fogão	Derivação Composto	Towe (fogo) + nekal (ne (afetivo) + kal (LOC)) + se (REF) Possivelmente nosso criador de fogo próximo. Também pode usar <i>fugãw</i> (empréstimo do Português <i>fogão</i>)
Txhatkhweaiekdonkya	Avião	Truncamento Derivação	Txhatkhwea (espaço/céu) + i (LINK) + ek (de <i>ekky</i> (correr)) + do (de <i>doa</i>) + nkya (FEM) = Aquela que corre no céu
Khotkya-neho	manicuro	Truncamento Derivação	Khoho (mão) + tkya (INSTR) + neho (INSTR)
Feyla	feira	Empréstimo da maioria língua	
Oyathaiekdonkya	Canoa	Derivação Truncamento	Oyakofo (ostra) + thaiky (carregar) + ek (de <i>ekky</i> (correr)) + do (de <i>doa</i>) + nkya (FEM) = a ostra que corre carregando
Td'ahesane	Caminhão	Extensão semântica Extensão semântica	Td'a (estrada) + hesane (grande)
Nãdudya	Inovação	Truncamento Conversão	nãkya (introduzir) + dudya (estranho)

TABELA 10 LISTA DE NEOLOGISMOS EM YATHÊ

Embora esteja em perigo de extinção, o Yathê sempre manteve um pequeno número de falantes nativos ativos. Há um grande número de empréstimos do Português, com pequenas alterações ortográficas e fonológicas para melhor se adequar à língua. Na amostra, há exemplos da maioria dos processos de cunhagem, exceto palavra-valise, acrônimos e ressignificação semântica. Todos os neologismos estudados seguem as regras normais de formação de substantivos da língua, tal como seria de esperar numa língua viva. Uma vez que o yathê é uma língua altamente flexiva, isso também se expressa na forma como os neologismos são criados, como seria de esperar numa língua que se manteve viva ao longo da colonização.

Numerais

Na primeira documentação disponível sobre o Yathé, Brenner (1886) observa que “os números (*em Yathé*) vão apenas até dez: tudo além disso é muitos”.

Lapenda (1968) tem uma descrição mais detalhada do sistema numérico, que muito provavelmente é composto em parte por neologismos:

Para os numerais cardinais, a lógica quina é usada em Yathé, já que cinco foi o número mais alto por muito tempo (Lapenda, 1968). Hoje, com a influência da civilização ocidental, existem números até 50 usando a lógica quina. A palavra para cinco é mão (*khoyá*) ou uma mão (*khoyáfthoa*) e a palavra para um é *fathoá*. A partir do número quatro, um empréstimo do Português com o sufixo “no” é uma alternativa aceitável. Abaixo, mostrarei isso a partir do seis. O Yathé adaptou-se, como se pode ver por isso, ao novo contexto linguístico colonial de uma forma muito flexível, aceitando dois sistemas em paralelo, mas colocando seu próprio “sabor” com o sufixo “-no” também onde se usa o empréstimo do Português.

Os primeiros dez números são mostrados na tabela 10.

Não	Palavra	Não	Palavra
1	<i>fathoá</i> (fem. <i>fathoane</i>)	6	<i>khoyáfthoa -thak-fathoá, sezeno</i>
2	<i>tkano</i>	7	<i>khoyáfthoa-thakê-tkáno, seteno</i>
3	<i>lixino</i>	8	<i>khoyáfthoa-thak-lixino, oiteno</i>
4	<i>sa'totkáno</i>	9	<i>khoyáfthoa-thakê-sa'totkáno, nono</i>
5	<i>khoyá, khoyáfthoa</i>	10	<i>khoyatkano</i> (duas mãos), <i>dezeno</i>
<i>A lógica continua com</i>			
15	<i>khoyalixino</i> (três mãos)	20	<i>khoyasa'totkáno</i> (quatro mãos), etc.

TABELA 11 NUMERAIS EM YATHÊ (LAPENDA, 1968)

Os numerais ordinais têm uma lógica diferente, em que, com exceção de “primeiro”, o numeral cardinal do número anterior é usado com o sufixo *-mtëa* e adicionando o sufixo feminino *-ne* quando aplicável, conforme o seguinte para os cinco primeiros números:

	masculino	feminino
1º	klehê	klehêne
2º	fathoá'mtëa	fathoá'mteane
3º	tkano'mtëa	tkano'mtëa
4º	lixino'mtëa	lixino'mtëa
5º	sa'totkáno'mtëa	sa'totkáno'mtëa

TABELA 12 NÚMEROS ORDINAIS EM YATHÊ (LAPENDA, 1968)

5.4 Borum

O empréstimo do Português é o processo de criação de palavras preferido em Borum, como mostra a tabela 12. Isso reflete, em grande medida, a dizimação brutal da população e expressa a necessidade de preencher rapidamente as lacunas lexicais.

Palavra	Português	Princípios de formação	Descrição da formação
<i>txúca</i>	açúcar	Empréstimo da língua majoritária	Adaptação à fonética do Borum
<i>reróge</i>	<i>relógio</i>	Empréstimo da língua majoritária	Adaptação à fonética Borum
<i>rápis</i>	lápis	Empréstimo da língua majoritária	Adaptação à fonética Borum
<i>joncát-nac</i>	trem de ferro	Empréstimo da língua majoritária (calque)	<i>tchon cate</i> (canoa) + <i>krak</i> (faca)
<i>papé</i>	papel	Empréstimo da língua majoritária	Adaptação à fonética Borum
<i>krenkat</i>	chapéu	Palavra-valise	<i>kren</i> (cabeça) + <i>kat</i> (pele)
<i>nakam</i>	cama	Empréstimo da língua majoritária Conversão Composto	<i>na</i> (em) + <i>kama</i> (cama)
<i>numat</i>	mato	Empréstimo da língua majoritária Conversão Composto	<i>no</i> (em) + <i>mato</i> (floresta)
<i>bikot</i>	biscoito	Empréstimo da língua majoritária	
<i>bentong</i>	botão	Empréstimo da língua majoritária	
<i>kauvan</i>	cavalo	Empréstimo da língua majoritária	
<i>sauon</i>	sabão	Empréstimo da língua majoritária	
<i>kudhé</i>	colher	Empréstimo da língua majoritária	

TABELA 13 LISTA DE NEOLOGISMOS EM BORUM

Como a maioria dos esforços de revitalização do Borum se concentra na recuperação lexical, na documentação e na criação de recursos orientados para professores para uso imediato em sala de aula, não há um desenvolvimento real da língua. Isso também se aplica à criação de novas palavras

para acompanhar o desenvolvimento tecnológico e social. O uso direto de palavras em Português é a forma principal de neologismos, com alguma adaptação da fonologia e ortografia. Alguns exemplos disso são: açúcar (Português *açucar*, Borum: *txúca*) e relógio (Português *relógio*, Borum: *reróge*) (Abreu, 1929, p. 594-600). Como algumas mudanças sintáticas ocorreram em Borum, incluindo estruturas portuguesas também nesse nível, alguns neologismos refletem essas mudanças com um tipo de conversão de uma locução preposicional que se torna um substantivo. A preposição portuguesa para em: “na” ou “no” é composta com os substantivos portugueses, como em *nakam* (Português *na cama*, na cama) e *nomat* (Português *no mato*, na floresta), tornando-se novos substantivos para cama e floresta.

Numerais

Nas primeiras listas de palavras do Botocudo (ou Krenak), apenas os três primeiros números são listados: 1 – *putxik*, 2 – *ɲarimbo* e 3 – *krutuip*. O seguinte, *inhait*, significa quatro, mas também muitos ou muito. O quinto número é raramente usado entre o povo Krenák, sendo relatado por apenas um dos membros da comunidade no estudo feito por Silva (2025). A palavra usada para ele, *kuan*, também significa “cheio”.

Na prática, quando os falantes precisam enumerar além dessa faixa nativa — particularmente em contextos escolares, administrativos ou urbanos — eles tomam emprestados os numerais portugueses (por exemplo, *quatro*, *cinco*, *seis*, etc.), um padrão consistente com a mudança linguística em comunidades criticamente ameaçadas e documentado nas histórias orais e narrativas de revitalização do Borum que descrevem a mistura cotidiana do Borum e do Português, inclusive para contar (Silva, 2025). Em conjunto, essas fontes sugerem que nenhum neologismo Borum amplamente adotado para números mais altos substituiu as formas portuguesas ou, melhor dizendo, os numerais emprestados são integrados quando as demandas comunicativas excedem o sistema tradicional, enquanto as iniciativas comunitárias se concentram em preservar e ensinar as formas nativas atestadas e sua semântica cultural (Silva, 2025).

5.5 Xukuru

Os empréstimos, seja do Português ou de línguas indígenas vizinhas, eram o processo de criação de palavras preferido em Xukuru quando a língua ainda estava viva. Uma lista de neologismos desse período é apresentada na tabela 13.

Palavra	Português	Princípios de formação	Descrição da formação
<i>ximinéw</i>	<i>fumaça</i>	Empréstimo da língua majoritária Extensão semântica	Extensão semântica de <i>cheminé</i> “chaminé”
<i>xakr</i>	<i>águia, rio</i>	Empréstimo da língua majoritária Extensão semântica	Extensão semântica de <i>chakro</i> “poça”
<i>xako</i>	<i>casa</i>	Empréstimo de língua vizinha	Tupi: <i>s-oka</i>
<i>takwar</i>	<i>faca</i>	Empréstimo da língua vizinha	Tupi: <i>sakwaren</i>
<i>tuyá</i>	<i>barriga</i>	Empréstimo de língua vizinha	Tupi: <i>t-yê</i>
<i>xiá</i>	<i>frio</i>	Empréstimo da língua vizinha	Yathé: <i>xiá</i>
<i>akó</i>	<i>filho</i>	Empréstimo da língua vizinha	Yathé: <i>e-ka</i>
<i>xuá</i>	<i>vento forte</i>	Empréstimo da língua vizinha	Yathé: <i>xmwá</i>

TABELA 13 LISTA DE NEOLOGISMOS EM XUKURU

Não há evidências de neologismos criados desde que o último falante nativo faleceu, em meados do século XX. No entanto, até então, há evidências de empréstimos, conforme mostrado na tabela 14, com extensão semântica ou não, tanto do Português quanto de línguas vizinhas, como o Yathé e o Tupi.

Numerais

A língua Xukuru não possui um sistema numérico documentado. Registros históricos e estudos linguísticos não fornecem evidências de um conjunto padronizado de numerais ou termos de contagem usados pelo povo Xukuru. Essa ausência sugere que o sistema nunca foi formalmente codificado ou que se perdeu ao longo do tempo devido a mudanças culturais e linguísticas. Pesquisas adicionais sobre tradições orais ou análises comparativas com línguas indígenas vizinhas podem oferecer pistas sobre como os Xukuru expressavam conceitos numéricos na vida cotidiana. A hipótese mais provável é que o sistema numérico Português tenha sido incluído no vocabulário Xukuru no início da era da colonização.

5. Discussão

A distribuição geral dos processos de cunhagem de palavras identificados neste estudo é mostrada na Figura 7 (com detalhes da distribuição fornecidos no Apêndice B). O empréstimo surge como o processo de criação de palavras mais frequente na amostra, seguido pela composição. Os demais processos são representados em proporções relativamente semelhantes, exceto os acrônimos e a resignificação semântica, que não aparecem nos dados. Essa distribuição relativamente uniforme entre a maioria das categorias reforça a representatividade da amostra, pois indica que nenhum processo menor domina de forma esmagadora o conjunto de dados.

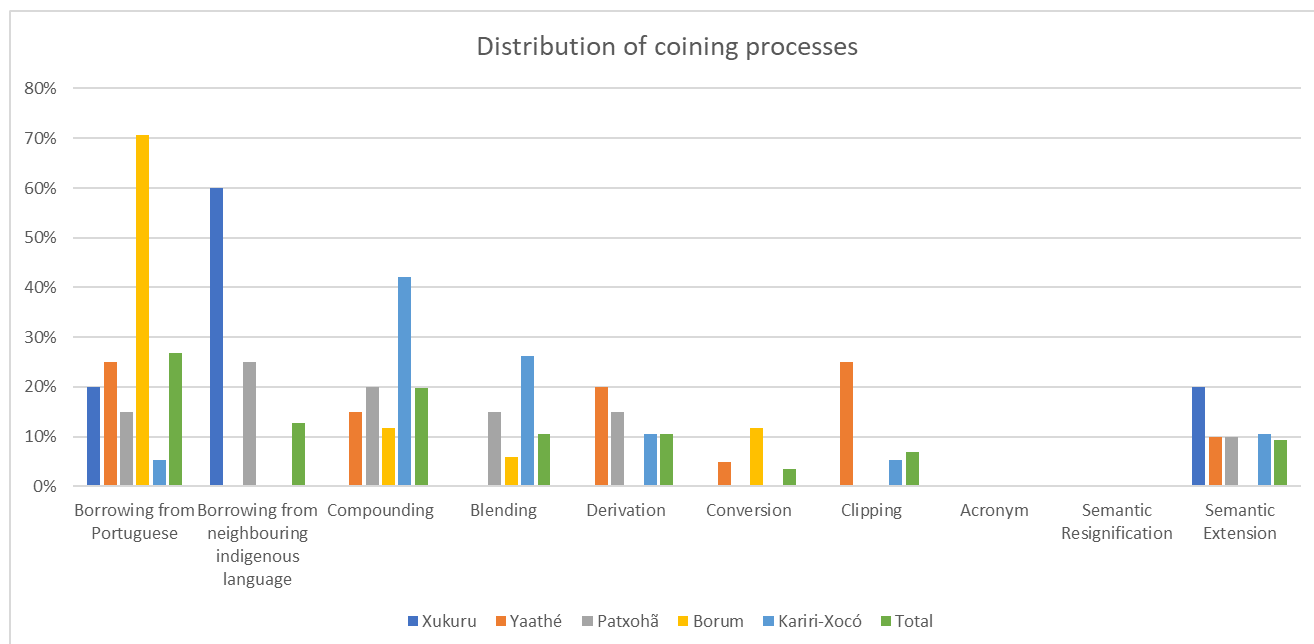


FIGURA 7 DISTRIBUIÇÃO DOS PROCESSOS DE CUNHAGEM DE NEOLOGISMOS

Padrões de empréstimo

No que diz respeito ao empréstimo lexical, o Português surge como a fonte dominante em geral, embora sua influência varie consideravelmente entre as línguas. O Patxohã apresenta um caso particularmente instrutivo. Apesar do contato extensivo com o Português, sendo que todos os falantes têm o Português como sua L1, a maioria dos seus empréstimos deriva, em vez disso, das línguas indígenas vizinhas — principalmente o Maxakalí — e não da língua colonial. A escolha deliberada de confiar no Maxakalí como principal língua de origem deve, portanto, ser entendida como uma estratégia política e cultural: um esforço explícito para distanciar as práticas de revitalização do legado da dominação colonial por meio do alinhamento linguístico com uma língua indígena viva intimamente relacionada ao Pataxó ancestral.

O Xukuru compartilha um perfil de empréstimos semelhante, com um número substancial de itens lexicais originários de línguas indígenas vizinhas. No entanto, neste caso, todos esses empréstimos são anteriores ao século XX e refletem mudanças regulares induzidas pelo contato em uma língua anteriormente viva, em vez de escolhas conscientes da era da revitalização. Ao contrário do Patxohã, onde o empréstimo é uma postura ideológica contemporânea, os dados do Xukuru capturam um estágio histórico de interação linguística anterior à dormência.

Em contraste, o Borum exibe a influência geral mais forte do Português. Esse padrão pode estar relacionado ao contato excepcionalmente precoce e intenso entre as comunidades Krenak e os

colonizadores portugueses, resultando em um bilinguismo assimétrico de longo prazo que favoreceu o empréstimo da língua dominante. O colapso populacional e a documentação deficiente amplificaram ainda mais essa tendência, forçando os falantes remanescentes a “remendar” lacunas lexicais usando formas portuguesas. Nesse caso, o empréstimo extensivo de palavras parece menos uma escolha ideológica e mais uma consequência da pressão sociolinguística externa.

Processos morfossintáticos de criação

Ao examinar os processos de neologismo que combinam material lexical existente — nomeadamente composição, palavra-valise e truncamento —, a composição surge como a estratégia preferida em todas as línguas. A palavra-valise e a truncamento ocorrem apenas marginalmente, com o Yathê como uma exceção significativa. O Yathê exibe uma proporção comparativamente alta de truncamentos, um padrão que pode ser explicado por ser único com sua transmissão intergeracional ininterrupta. A truncamento, que muitas vezes produz formas mais curtas e menos complexas morfológicamente, é normalmente usada para sinalizar familiaridade e uso informal contínuo (Hilpert et al., 2023), sugerindo que o perfil de neologismos do yathê se alinha mais estreitamente com o de línguas não ameaçadas de extinção. A rica morfologia do yathê também é evidente quando se observa como os neologismos são criados, já que muitos dos morfemas formadores de substantivos já presentes na língua também são usados no processo de cunhagem.

O Yathê também se destaca pelo uso relativamente frequente da derivação como estratégia de criação, enquanto o Patxohã não apresenta derivações na amostra analisada. Embora esse contraste seja marcante, seriam necessárias mais pesquisas para determinar por que isso ocorre: se reflete diferenças tipológicas, efeitos de amostragem ou se é resultado do processo de decisão muito ideológico e estruturado por trás dos processos de cunhagem do Patxohã.

O Borum apresenta outro desvio digno de nota dos padrões gerais, utilizando a conversão como parte de estratégias mistas de neologismo. Vários neologismos são formados através da composição de uma preposição locativa e um substantivo portugueses, seguidos da conversão num item lexical, como em *nakam* “na cama” (< Português *na cama*). A coexistência de empréstimos, composição e conversão dentro de um único processo de formação se destaca em uma língua sob forte pressão externa de uma língua majoritária historicamente hostil. Esse padrão pode estar novamente ligado à urgência de preencher lacunas lexicais em uma população de falantes em rápido declínio, mas também pode ser resultado da dificuldade (Seki, 2004) que os falantes tinham em produzir o equivalente à afirmação solicitada em Português.

A formação de acrônimos está totalmente ausente como processo de neologismo indígena na amostra. Os únicos acrônimos observados são aqueles emprestados integralmente do Português, principalmente em contextos administrativos, por exemplo, *TI* (*Terra Indígena*) e *CPF* (*Cadastro de Pessoas Físicas*), indicando que a acronimização não foi internalizada em nenhuma das línguas estudadas como uma estratégia morfológica produtiva. Esses casos não são considerados na amostra.

Processos de cunhagem semântica

Os processos de neologismo semântico revelam uma clara assimetria. Nenhum caso de resignificação semântica aparece nos dados. A extensão semântica, no entanto, é atestada em todas as línguas, exceto no Borum, e é particularmente proeminente na criação de termos para novas tecnologias. Uma possível hipótese para a ausência de neologismos cunhados usando resignificação semântica seria: *quando uma língua atinge um estado de grave perigo de extinção, a resignificação semântica em nível comunitário cessa*. Os dados deste estudo apoiam essa afirmação, mas não são suficientes como base para essa hipótese, pois a amostra é simplesmente muito pequena. Essa seria uma área interessante

para estudos futuros, pois aponta para as condições necessárias para os processos de cunhagem semântica, bem como para o que a transmissão intergeracional da língua proporciona em termos de desenvolvimento linguístico. Se for demonstrado que a ressignificação semântica reaparece em línguas revitalizadas com sucesso, ela pode até se tornar um marcador do sucesso da revitalização.

Processos de cunhagem de numerais

As estratégias de neologismo relacionadas com numerais e sistemas numéricos revelam um quadro um pouco diferente. Todas as línguas adotaram uma estrutura decimal ou quinal, ligada à contagem com os dedos. O Patxohã não empresta numerais do Português, mas adotou um sistema decimal tanto para numerais ordinais como cardinais. Como a documentação da língua ancestral atesta apenas um único numeral (“um”), o grupo Atxohã criou um sistema numérico totalmente novo, sem usar línguas vizinhas como modelos estruturais ou lexicais — uma exceção às suas práticas gerais de empréstimo. Não há pistas sobre o motivo disso, o que ainda precisa ser investigado/revelado pelos linguistas Patxohã. O Kariri-Xocó, por outro lado, emprega uma lógica quinal para os primeiros dez numerais e muda para um sistema decimal a partir daí, enquanto o Yathé possuía originalmente um sistema indígena até dez (conceituado como “duas mãos”), que foi expandido para vinte (“quatro mãos”) antes de integrar uma lógica decimal em valores mais altos. Os numerais portugueses coexistem em paralelo a partir dos valores baixos, marcados com sufixos nativos. Isso reflete muito bem a integração na infraestrutura colonial que os Fulni-ô adotaram. O Borum mostra a integração mais extensa dos numerais portugueses, com termos nativos limitados a um, dois e três, extensões de palavras em massa para quatro e cinco e dependência completa das formas portuguesas além disso. O Xukuru, sem nenhum numeral indígena documentado, parece ter integrado totalmente o sistema Português antes do período de dormência. Em todas as línguas, os sistemas de contagem começam com lógicas baseadas nos dedos e nas mãos e divergem na forma como se expandem além de cinco, dez ou vinte, refletindo diretamente diferentes respostas ao contato colonial intensificado.

Potenciais questões metodológicas

Uma fraqueza deste estudo é a desigualdade, confiabilidade e, às vezes, falta de disponibilidade dos dados nas cinco línguas. Especialmente o Xukuru e o Borum, dormentes ou gravemente ameaçados, têm o problema de que o material lexical existente é fragmentário e um tanto inconsistente. Isso afeta a representatividade da classificação dos neologismos, assim como a decisão tomada não apenas de descrever os neologismos em Patxohã com base em empréstimos do Maxikali, embora esta seja a esmagadora maioria. Além disso, a amostra de neologismos por idioma foi necessariamente pequena, o que significa que os padrões identificados podem não incluir todos os processos neológicos existentes em cada comunidade e a distribuição proporcional dos processos utilizados certamente não é muito exata. Por fim, o estudo se concentra exclusivamente em substantivos e numerais; incluir verbos ou outras classes de palavras, bem como neologia fonológica, poderia ter proporcionado uma compreensão mais completa da inovação lexical impulsionada pela revitalização. Pesquisas futuras com conjuntos de dados mais ricos, entrevistas e domínios lexicais mais amplos ajudariam a validar e refinar as conclusões aqui apresentadas.

6. Conclusões

A questão central da pesquisa deste artigo foi identificar as estratégias utilizadas para cunhar novos itens lexicais em línguas indígenas revitalizadas da região Nordeste do Brasil e examinar como essas estratégias refletem tanto as características estruturais de cada língua quanto as condições socioculturais em que a revitalização ocorre. A análise comparativa mostra que as estratégias de neologismo estão longe de ser uniformes; ao contrário, elas são moldadas pela trajetória histórica de cada língua, pelo grau de interrupção e pela relação tanto com o Português quanto com as línguas indígenas vizinhas. A análise confirma que, neste contexto específico, o que Campos (2023) afirmou sobre os processos de revitalização estarem fortemente ligados à reafirmação da identidade também é válido para a escolha dos processos de cunhagem.

Também é importante destacar que, embora o contexto sociolinguístico se reflita na preferência por determinados processos de cunhagem, as características tipológicas determinam a disponibilidade dos processos. Um exemplo que mostra isso claramente é o Patxohã, onde a maioria dos neologismos são empréstimos do Maxakali, embora o processo normal de formação de palavras para substantivos seja a derivação com sufixos. Se as características tipológicas da língua também se refletem na escolha do processo de criação é uma questão interessante, mas fora do escopo deste trabalho.

A partir deste estudo, fica claro que a análise de neologismos na revitalização da língua não pode ser adequadamente capturada através de uma lente puramente linguística. Em vez disso, requer uma estrutura sociolinguística e ideológica atenta às questões de identidade, bem como à autoridade e à história da língua. Essas dimensões moldam fundamentalmente não apenas *se* (como visto no exemplo Xukuru) os neologismos são criados, mas também *como* as práticas de criação são usadas, estruturadas, avaliadas e legitimadas dentro dos movimentos de revitalização. Nesse sentido, a inovação lexical é inseparável do ecossistema político e cultural mais amplo.

Os neologismos em línguas revitalizadas não são meramente soluções para lacunas lexicais; eles também funcionam como atos de resistência cultural e criação de identidade. Como observa Batista Silva (2025) em sua discussão sobre a revitalização do Krenak, um paradoxo central surge da dependência de arquivos coloniais, registros missionários e documentação etnográfica para reviver línguas que foram ativamente silenciadas pela expansão colonial. Essa dependência introduz uma tensão persistente entre a autenticidade linguística e as estratégias pragmáticas de revitalização. As comunidades devem navegar pela coexistência incômoda entre a ruptura histórica e a continuidade contemporânea, uma negociação que influencia profundamente as decisões sobre a expansão lexical e a legitimidade das formas recém-criadas.

Os processos de cunhagem também variam significativamente, dependendo do estágio do esforço de revitalização. Nas fases iniciais — particularmente em contextos onde a transmissão intergeracional foi interrompida — é comum o empréstimo da língua dominante (Português, no caso brasileiro) ou a adaptação de material lexical ancestral a domínios semânticos modernos. O Patxohã oferece um exemplo representativo: conceitos tecnológicos e contemporâneos têm sido expressos por meio da extensão semântica de morfemas tradicionais, ligando referentes ancestrais aos atuais. Um exemplo disso, da tabela 7 (neologismos Patxohã): [*tapitébm'aô*], que significa geladeira, combina um morfema emprestado do Maxakali *tapitébm* (clima frio/gelo) com *-aô*, um sufixo derivacional tradicional. Este exemplo e a experiência Patxohã em geral refletem, em muitos aspectos, o que Fitzgerald (2017, 2021) descreve no modelo Chickasaw, mas especificamente para neologismos no contexto do Nordeste brasileiro.

A neologia organizada em línguas revitalizadas raramente é impulsionada por ideologias puristas ou prescritivistas, como visto nas línguas deste estudo. Em vez disso, ela faz parte da luta pelo

reconhecimento cultural, identidade, agência política e soberania linguística. Em Patxohã e, em certa medida, em Kariri-Xocó, a inovação lexical funciona como um recurso estratégico que permite aos falantes se adaptarem às realidades modernas sem perder sua identidade indígena. Os neologismos adquirem, assim, uma função dupla: atendem às necessidades comunicativas e, ao mesmo tempo, servem como atos simbólicos de resistência contra a contínua marginalização linguística e cultural, muito em linha com o que é discutido por Loncón Antileo e Castillo Sánchez (2019) sobre os esforços de revitalização em geral. Neste contexto, a inovação não se opõe à tradição, mas surge através de uma interação dinâmica com ela.

As línguas cujos processos de revitalização estão fortemente enraizados em iniciativas lideradas pela comunidade tendem a minimizar o empréstimo da língua majoritária circundante, como parte de uma postura mais ampla de resistência cultural. Ao mesmo tempo, o empréstimo de línguas indígenas vizinhas é generalizado em todo o corpus examinado e, a nível estrutural, o empréstimo de parâmetros tipológicos do Português está a ocorrer em todas as línguas estudadas (por exemplo, o sistema numérico decimal, bem como a abertura da ordem das palavras SVO em paralelo com o SOV existente).

Por fim, os sistemas numéricos oferecem um exemplo ilustrativo de neologismo estruturado. A maioria das línguas estudadas desenvolveu estratégias explícitas de criação de numerais, empregando comumente bases decimais ou quínarias. Na maioria dos casos, os numerais ou a estrutura numérica do Português são adotados parcialmente, em paralelo ou inteiramente. Nenhuma das línguas examinadas mostra evidências de ter possuído um sistema de contagem que se estendesse além do dez na época da colonização. Isso aponta novamente para o fato evidente de que os processos de cunhagem nessas línguas são frequentemente forçados pela pressão externa da língua dominante.

Em conjunto, as descobertas demonstram que as estratégias de neologismo nas línguas revitalizadas não são simplesmente respostas a necessidades lexicais, mas estão profundamente enraizadas na história da comunidade falante, nos padrões de perda da língua, bem como nas escolhas ideológicas impulsionadas pela comunidade. A inovação lexical surge, assim, como um processo estrutural e um ato significativo de posicionamento cultural dentro das lutas contínuas pela soberania linguística e política.

Perspectivas (Rumo a um possível modelo baseado no Chickasaw)

Os padrões identificados neste estudo sugerem que os futuros esforços de revitalização no Nordeste do Brasil provavelmente entrarão em uma fase em que a expansão lexical se tornará menos reativa e mais estruturada institucional e ideologicamente, como mostra o exemplo do Patxohã. Nesse sentido, os desenvolvimentos observados podem ser interpretados de forma produtiva através das lentes do modelo Chickasaw de revitalização (Fitzgerald & Hinson 2013), que enfatiza o planejamento linguístico de longo prazo, a autoridade da comunidade sobre a inovação lexical com base na cooperação com a comunidade de estudiosos da linguística e a normalização do neologismo como um estágio permanente, em vez de um mecanismo de transição. Este modelo é descrito com mais detalhes na subseção 2.1 acima. Os esforços do Patxohã também podem ser usados como um exemplo de como alcançar um desenvolvimento impulsionado pela comunidade.

A neologia deve ser entendida não apenas como uma solução de emergência para preencher lacunas, mas também como um componente explícito do planejamento do corpus, de preferência coordenado por meio de comitês linguísticos, instituições pedagógicas e consenso comunitário em um processo de validação. A aplicação dessa estrutura ao contexto do Nordeste sugere que línguas como o Patxohã e o Yathé — onde a revitalização já está fortemente enraizada na comunidade, embora em contextos completamente diferentes — estão bem posicionadas para avançar em direção a modelos semelhantes de criação organizada de neologismos. Nesses casos, o neologismo contínuo seria cada vez mais moldado por normas metalinguísticas compartilhadas, favorecendo estratégias como extensão semântica e composição que se alinham com a lógica morfológica interna e compromissos ideológicos com a autonomia linguística (Fitzgerald, 2017; Fitzgerland, 2021).

A experiência Chickasaw também destaca a importância da profundidade temporal no desenvolvimento lexical. À medida que o uso diário se estabiliza e os domínios de uso se expandem, as formas criadas anteriormente tendem a se naturalizar, e processos mais orgânicos — como resignificação, deriva metafórica e ampliação pragmática — começam a ocorrer. A partir dessa perspectiva, a atual ausência de resignificação semântica nas línguas nordestinas estudadas não deve ser interpretada como uma limitação estrutural, mas sim como um efeito previsível da transmissão interrompida (ver hipótese proposta anteriormente sobre esse efeito). Se os esforços de revitalização forem bem-sucedidos na criação de comunidades de falantes sustentáveis, essas línguas podem gradualmente começar a desenvolver novos significados para as palavras por conta própria novamente.

Para línguas com histórias de transmissão mais fragmentadas, como o Borum e o Xukuru, uma perspectiva baseada no Chickasaw sugere uma trajetória diferente, mas igualmente legítima. O modelo Chickasaw reconhece explicitamente as práticas híbridas de criação de palavras — incluindo empréstimos, calques e adaptações baseadas em códigos — como inevitáveis em condições de contato assimétrico intenso. Em tais contextos, a criação contínua de palavras provavelmente permanecerá pragmática, enfatizando a eficiência comunicativa e a acessibilidade dos falantes em detrimento do purismo estrutural. Em vez de representar um fracasso, essas estratégias refletem respostas adaptativas a traumas históricos e documentação limitada, paralelamente aos estágios iniciais da revitalização do Chickasaw antes do surgimento de estruturas institucionais estáveis.

De forma mais ampla, a adoção dessa perspectiva teórica também reformula os critérios de avaliação para o sucesso da revitalização. Evitar empréstimos do Português não deve ser tratado como uma referência absoluta. Em vez disso, o fator decisivo é se os processos de cunhagem de palavras são socialmente aceitáveis e autorizados, debatidos internamente e culturalmente significativos para a

comunidade linguística. A criação contínua de neologismos, nesse sentido, deve ser vista não como um problema a ser resolvido, mas como um sinal de vitalidade linguística: um marcador de que as línguas revitalizadas voltaram a fazer parte da família das línguas vivas.

Em conclusão, os futuros esforços de revitalização na região Nordeste provavelmente se beneficiarão do reconhecimento explícito do neologismo como um processo contínuo e multifásico incorporado ao planejamento linguístico, em vez de uma resposta temporária ao déficit lexical. A sustentabilidade a longo prazo dessas línguas dependerá não

da eliminação da influência externa, mas da capacidade das comunidades de controlar como novos significados entram em seus sistemas linguísticos — e como a inovação é conciliada com a memória, a identidade e a soberania.

8. Referências

- Abreu, S. F. (1929). Os índios crenaques (botocudos do rio Doce) em 1926. *Revista do Museu Paulista*, 16, 594–600.
- Adelaar, W. F. H., & Pache, M. J. (2022). Are all language isolates equal? The case of Mapudungun. In T. C. Chacón, N. H. Lee, & W. D. L. Silva (Eds.), *Language change and linguistic diversity* (pp. 164–186). Edinburgh University Press
- Algeo, J. (1991). *Fifty years among the new words: A dictionary of neologisms, 1941–1991*. Cambridge University Press.
- Alves, I. M. (1990). *Neologismo*. Ática.
- Anchieta, J. de. (1595). *Arte de grammatica da lingua mais usada na costa do Brasil*. António de Mariz.
- Batista Silva, A. J. M. (2025). Krenak legacy materials: Learning Indigenous languages from centuries-old sources. *Smithsonian Folklife Magazine / Endangered Languages Project*.
- Bauer, L. (1983). *English word-formation*. Cambridge University Press.
- Bernal, J. D. (1939). *The social function of science*. Routledge & Kegan Paul.
- Bomfim, A. B. (2017). *Patxohã: Língua de guerreiro. Um estudo sobre o processo de retomada da língua Pataxó*. [Master's thesis, Universidade Federal da Bahia].
- Bomfim, A., & Lima, S. (2020). Count and mass nouns in Patxohã. *Linguistic Variation*, 20(2), 324–335. John Benjamins Publishing Company.
- Bomfim, E. S., Durazzo, L., & Aguiar, M. S. (2021). O “levante linguístico indígena” no Nordeste, Espírito Santo e Minas Gerais: Aspectos teóricos, políticos e etnográficos. *Policromias – Revista de Estudos do Discurso, Imagem e Som*, 6(2), 398–420.
- Bradley, D. (2019). Language endangerment and revitalization in Asia. In R. B. Kaplan (Ed.), *The Oxford handbook of language policy and planning* (pp. 555–574). Oxford University Press.
- Branner, J. C. (1887). Notes upon a native Brazilian language. In *Proceedings of the American Association for the Advancement of Science (35th meeting)*. Biblioteca Digital Curt Nimuendajú.
- Brasil. (2018). *Constituição da República Federativa do Brasil: Promulgada em 5 de outubro de 1988*. Senado Federal.
- Braz da Conceição, N. (2016). *Uma reflexão sobre variação linguística na língua Patxohã do povo Pataxó* [Undergraduate Licentiate Thesis, Universidade Federal de Minas Gerais].
- Caetano de Sá, A. (1998). *Dicionário latê–Português; Português–latê*. Associação da Imprensa de Pernambuco.
- Calma, T. (2009). *Social justice report*. <https://www.humanrights.gov.au/publications/chapter-3-introduction-social-justice-report-2009>
- Calude, A.S. (2021). The history of number words in the world's languages—what have we learnt so far? *Philos Trans R Soc Lond B Biol Sci* 10 May 2021; 376 (1824)
- Calvet, L.-J. (2005). *Linguística y colonialismo: Breve tratado de glotofagia*. FICE.

- Camacho, L. (2024). Automating the proposition of neologisms for the Quechua language. *Journal of the International Phonetic Association*, 54(3), 922–938. <https://doi.org/10.1017/S0025100324000227>
- Cámara-Leret, R., & Bascompte, J. (2021). Language extinction triggers the loss of unique medicinal knowledge. *Proceedings of the National Academy of Sciences of the United States of America*, 118(24), e2103683118.
- Campani, G., & Urani, L. (2025). Spiritual resistance and cultural exchange: The Fulni-ô of Águas Belas in an intercultural context. *London Journal of Research in Humanities & Social Science*, 25(5), 449U.
- Campos, C. S. (2023). Revitalização e retomada de línguas: Consequências e desafios. *PERcursos Linguísticos*, 13(33). <https://doi.org/10.47456/pl.v13i33.40923>
- Carvalho, S. (2019). *O papel da língua no fortalecimento da identidade étnica dos Xukuru da Serra de Ororubá em Pesqueira–PE*. Universidade Federal de Pernambuco.
- Coordenação Atxôhã. (n.d.). *Dicionário Português x Patxôhã*. Coordenação de Pesquisa da Língua e Cultura Pataxó.
- Coronel-Melina, S. M., & McCarty, T. L. (2016). *Indigenous language revitalization in the Americas*. Routledge.
- Correia de Queiroz, J. M. (2012). *Um estudo gramatical da língua Dzubukuá, família Kariri* [Doctoral dissertation, Universidade Federal da Paraíba].
- Da Costa, J. F. (1999). *Ya:thê, a última língua nativa no Nordeste do Brasil: Aspectos morfofonológicos e morfo-sintáticos* [Doctoral dissertation, Universidade Federal de Pernambuco].
- Da Costa, J. (2015). Descrevendo línguas brasileiras: Yaathe, a língua dos índios Fulni-ô. *Revista do GELNE*, 17(1–2), 93–111.
- Dantas, B. G., & Dallar, D. de A. (1980). *Terra dos índios Xocó : Estudos e documentos*. Comissão Pró-Índio.
- Durazzo, L., & Costa, F. V. F. da. (2020). Línguas indígenas no Nordeste brasileiro : Esboço político-linguístico de seus processos de valorização. *Revista del CESLA: International Latin American Studies Review*. <https://doi.org/10.36551/2081-1160.2022.30.97-118>
- Eberhard, D. M., Simons, G. F., & Fennig, C. D. (2023). *Ethnologue: Languages of the world* (26th ed.). SIL International. <https://www.ethnologue.com>
- Elsen, H. (2004). *Neologismen: Formen und Funktionen neuer Wörter in verschiedenen Varietäten des Deutschen*. Günter Narr Verlag.
- Ethnologue. (2025). *Papua New Guinea: Languages, literacy, and maps*. <https://www.ethnologue.com/country/PG/>
- European Parliament, Directorate-General for Internal Policies of the Union. (2013). *Endangered languages and linguistic diversity in the European Union*. European Parliament Publications Office. <https://doi.org/10.2861/22124>

- Falco Genovez, P., & Santana, G. B. (2025). Memória, território e ancestralidade Krenak : A defesa dos direitos dos povos indígenas e os esforços da Organização dos Estados Americanos (OEA). *Revista de História da América*, 17(171).
- Feest, C. (2022). *Botocudos in Europe in the 1820s* [Manuscript]. ResearchGate. https://www.researchgate.net/publication/363885283_Botocudos_in_Europe_in_the_1820s
- Fishman, J. A. (1991). *Reversing language shift: Theoretical and empirical foundations of assistance to threatened languages*. Multilingual Matters.
- Fitzgerald, C. M. (2017). Language reclamation: Chickasaw language revitalization and the reunification of language and identity. In W. Leonard & H. De Korne (Eds.), *Language documentation and description* (Vol. 14, pp. 165–182). EL Publishing.
- Fitzgerald, C. M. (2021). A framework for language revitalization and documentation. *Language*, 97(1), e1–e11.
- Fitzgerald, C. M., & Hinson, J. D. (2013). 'littibatoksali "We work together": Perspectives on our Chickasaw tribal–academic collaboration. In M. J. Norris, E. Anonby, M.-O. Junker, N. Ostler, & D. Patrick (Eds.), *FEL XVII: Endangered languages beyond boundaries: Community connections, collaborative approaches, and cross-disciplinary research* (pp. 53–60). The Foundation for Endangered Languages.
- Florêncio, R. R. (2022). *Educação indígena e intercultural nas aldeias do Opará: Uma sociedade de sujeitos silenciados* [Doctoral dissertation, Universidade Federal da Bahia]. Repositório UFBA.
- Franchetto, B. (2020). Língua(s): Cosmopolíticas, micropolíticas, Macropolíticas. *Campos – Revista de Antropologia*, 21(1), 21–36. <https://doi.org/10.5380/cra.v21i1.70519>
- Frassetto, P. T. (2018). *Recuperação lexical: Um vocabulário Português–Krenak / Krenak–Português* [Bachelor's Thesis, Universidade Estadual de Campinas].
- Fundação Nacional dos Povos Indígenas. (2022). *Cultura indígena: Conheça o Ouricuri, ritual religioso da etnia Fulni-ô*. <https://www.gov.br/funai/pt-br/assuntos/noticias/2022-02/cultura-indigena-conheca-o-ouricuri-ritual-religioso-da-etnia-Fulni-o>
- Garfield, S. (2000). As raízes de uma planta que hoje é o Brasil: Os índios e o Estado-nação na era Vargas. *Revista Brasileira de História*, 20(39), 15–42.
- Guerrettaz, A. M., & Engman, M. (2023). Indigenous language revitalization. In P. Groves Price (Ed.), *Oxford encyclopedia of race and education*. Oxford University Press.
- Greenberg, J. H. (1987). *Language in the Americas*. Stanford University Press.
- Greenough, J. B., & Kittredge, G. L. (1901). *Words and their ways in English speech*. Macmillan.
- Grenoble, L. A., & Whaley, L. J. (2005). Language revitalization as a global issue. In *Saving languages: An introduction to language revitalization*. Cambridge University Press.
- Gruda, S., Haimovich, G. & Sullivan, J. (2023). *Lexical creativity in modern Nahuatl: An analysis of multidialectal data*. *Lingua*, 285, 103488. <https://doi.org/10.1016/j.lingua.2023.103488>
- G20 Brasil. (2024). *Indigenous languages, essential for preserving ancestral knowledge, could disappear in 20 years*. <https://www.gov.br/g20/en/news/Indigenous-languages-essential-for-preserving-ancestral-knowledge-could-disappear-in-20-years>

- Harrison, K. D. (2007). *When languages die: The extinction of the world's languages and the erosion of human knowledge*. Oxford University Press.
- Hilpert, M., Correia Saavedra, D. & Rains, J. (2023). Meaning differences between English clippings and their source words: A corpus-based study. *ICAME Journal*, 47(1), 19–37. <https://doi.org/10.2478/icame-2023-0002>
- Hinton, L., Huss, L., & Roche, G. (Eds.). (2018). *The Routledge handbook of language revitalization*. Routledge.
- Hohenthal, W. D., Jr. (1954). Notes on the Shucurú Indians of Serra de Ararobá, Pernambuco, Brazil. *Revista do Museu Paulista*, Nova Série, 8, 93–166.
- Humboldt, W. von. (1999/1836). *On language: On the diversity of human language construction and its influence on the mental development of the human species* (P. Heath, Trans.; M. Losonsky, Ed. pp. 50–51). Cambridge University Press. (Original work published 1836)
- Huylebrouck, D. (2010). Mathematics in (central) Africa before colonization. *Anthropologica et Præhistorica*, 121, 135–162.
- Instituto Brasileiro de Geografia e Estatística (IBGE). (2025). *2022 Census: Brazil has 391 Indigenous ethnicities and 295 Indigenous languages*. Agência de Notícias IBGE. <https://agenciadenoticias.ibge.gov.br/en/agencia-news/2184-news-agency/news/44866-2022-census-brazil-has-391-indigenous-ethnicities-and-295-indigenous-languages>
- Instituto Socioambiental. (n.d.). *Fulni-ô*. In *Povos Indígenas no Brasil*. <https://pib.socioambiental.org>
- Instituto Socioambiental. (n.d.). *The work of linguists*. In *Indigenous Peoples in Brazil*. https://pib.socioambiental.org/en/The_work_of_linguists
- Instituto Socioambiental. (n.d.). *Krenak*. In *Indigenous Peoples in Brazil (PIB)*. <https://pib.socioambiental.org/en/Povo%3AKrenak>
- Itchok. (n.d.). *Sobre Itchok – Casa da Língua Krenak*. <https://www.itchok.com/blank>
- Jolkesky, M. (2016). *Estudo arqueo-ecolinguístico das terras tropicais sul-americanas* [Doctoral dissertation, Universidade de Brasília].
- Kariri Xocó, I., Kariri Xocó, N., Nelson, D., & Pitman, T. (2020). A retomada da língua Kariri Xocó. *Cadernos de Linguística*, 1(3), 1–13. <https://doi.org/10.25189/2675-4916.2020.v1.n3.id254>
- Kariri-Xocó, N. (n.d.). Blog *Kariri-Xocó cultura digital, cultura popular, blogs, história*. <https://kxnhenety.blogspot.com>
- Kaufman, T. (1990). *Languages and language families of South America: A global classification*. University of Chicago Press.
- Kik, A., Adamec, M., Aikhenvald, A., Bajzekova, J., Baro, N., Bownern, C., Colwell, R. K., Drozd, P., Duda, P., Ibalim, S., Jorge, L. R., Mogina, J., Ruli, B., Sam, K., Sarvasy, H. S., Saulei, S., Weiblen, G. D., Zrzavý, J., & Novotný, V. (2021). *Language and ethnobiological skills decline precipitously in Papua New Guinea, the world's most linguistically diverse nation*. *Proceedings of the National Academy of Sciences*, 118(22), e2100096118. <https://doi.org/10.1073/pnas.2100096118>

- Kimura, L., & Counciller, I. A. G. L. (2009). Indigenous new words creation: Perspectives from Alaska and Hawaii. In J. Reyhner & L. Lockard (Eds.), *Indigenous language revitalization: Encouragement, guidance & lessons learned* (pp. 121–139). Northern Arizona University.
- Kinne, M. (1998). Der lange Weg zum deutschen Neologismenwörterbuch. In W. Teubert, *Neologie und Korpus* (pp. 63–110). Narr.
- Kirilenko, S. V. (2024). *Language attitudes: Sociolinguistic aspect. Research Result: Theoretical and Applied Linguistics*, 10(1). <https://doi.org/10.18413/2313-8912-2024-10-1-0-2>
- Kuldashova, N. B., & Safarbaeva, S. (2025). The concept of neologism and its linguistic significance. *American Journal of Education and Evaluation Studies*, 2(2). <https://semantjournals.org/index.php/AJEES/article/view/1051>
- Lapenda, G. C. (1962). *O dialecto Xukuru. Doxa*. Revista Oficial do Departamento de Cultura do Diretório Acadêmico da Faculdade de Filosofia de Pernambuco, 10, 11–23.
- Lapenda, G. C. (1968). *Estrutura da língua latê: Falada pelos índios Fulniôs em Pernambuco*. Imprensa Universitária, Universidade Federal de Pernambuco.
- Leonard, S. P., & Árnason, K. (2011). Language ideology and standardisation in Iceland. In *Standard languages and language standards in a changing Europe* (pp. 91–96). Novus Forlag
- Loncón Antileo, E., & Castillo Sánchez, S. (2019). Neologismos en mapuzugun: Palabras creadas en un proceso de enseñanza y de aprendizaje. *Literatura y Lingüística*, (38), 195–212.
- Loukotka, Č. (1968). *Classification of South American Indian languages*. UCLA Latin American Center.
- Mamiani, L. V. (1877/1699). *Arte de grammatica da língua brasilica da naçam Kiriri* (2nd ed.). Bibliotheca Nacional.
- Marlière, G. T. (1825). *Vocabulário botocudo* [Manuscript]. Arquivo Público Mineiro.
- Max Planck Institute for Evolutionary Anthropology. (n.d.). *Krenak*. In *Numeral systems of the world's languages*. <https://lingweb.eva.mpg.de/channumerals/Krenak.htm>
- Meland, D., & Meland, D. (2009). *Fulniô (Yaathe) syntax structure* (Preliminary version ; original work published 1968). SIL Brasil.
- Ministério de Minas e Energia. (n.d.). *Pataxó, Êtxawênuk ãkgaré iõ otxemã hãtxay upã níognigãhí pírikã*. <https://www.gov.br/mme/ptbr/arquivos/cartilhapataxo.pdf>
- Mirim. (n.d.). *Língua*. <https://mirim.org/pt-br/lingua>
- Museu Índia Vanuíre. (2019). *Vocabulário unificado Português–Krenak (Botocudo) / Krenak–Português, do século XIX*. https://museuindivanuire.org.br/cultura_em_casa/vocabulario-unificado-portugues-Krenak-botocudo-Krenak-portugues-do-seculo-xix/
- Nelson, J. F. (2018). *Pataxó Hãhãhãe: Race, indigeneity and language revitalization in the Brazilian Northeast* [Doctoral dissertation, University of Arizona]. <https://repository.arizona.edu/handle/10150/630222>
- Nelson, D., et al. (2023). *We most certainly do have a language*. University of Leeds. <https://doi.org/10.1215/22011919-10216239>

- Nikulin, A. (2020). *Proto-Macro-Jê: Um estudo reconstrutivo* [Doctoral dissertation, Universidade de Brasília].
- Nunes Duarte Silveira, C. (2018). *A arte de evangelizar no teatro anchetano: Memória, conversão e doutrina* [Doctoral dissertation, Universidade Estadual do Sudoeste da Bahia].
- Office of the President of the Republic of Poland. (2025, October 16). *Weto do nowelizacji ustawy o mniejszościach narodowych i etnicznych oraz o języku regionalnym*.
<https://www.prezydent.pl/prawo/ustawy-zawetowane/weto-do-nowelizacji-ustawy-o-mniejszosciach-narodowych-i-etnicznych-oraz-o-jezyku-regionalnym,108735>
- Oliveira Gonçalves, A. A. (2025). Falar em Patxôhã, cortar língua. *Tellus*, 24(53), 99–132.
- Olko, J., & Sallabank, J. (2020). Why revitalize? In J. Olko & J. Sallabank (Eds.), *Revitalizing endangered languages* (pp. 1–16). Cambridge University Press.
- Oreški, J. (2021). The end of a never-ending story of attempts to define neologisms? *SN Social Sciences*, 1, Article 170. <https://doi.org/10.1007/s43545-021-00181-y>
- Overmann, K. A. (2021). Finger-counting and numerical structure. *Frontiers in Psychology*, 12, Article 723492. <https://doi.org/10.3389/fpsyg.2021.723492>
- De Paula Camargo, J. (2020). *Os neologismos como um processo de resistência: Observações na comunidade Wakalitesu – Aldeia Três* [Master's Thesis, Universidade Federal de Mato Grosso].
- Pentland, D. H. (1982). French loanwords in Cree. *Kansas Working Papers in Linguistics*, 7, 105–117.
- Pereira Serra, F., & de Maria de Araujo Ramos, C. (2021). Falares africanos no Maranhão: Um glossário afro-maranhense. *Afluente: Revista de Letras e Linguística*, 6(19), 174–198.
- Pérez Báez, G., Vogel, R., & Patolo, U. J. (2019). Global survey of revitalization efforts: A mixed methods approach to understanding language revitalization practices. *Language Documentation & Conservation*, 13, 446–513.
- Pessoa, K. N. (2008). A língua Krenak : Fonologia não-linear e aspectos morfosintáticos. *Anais do SETA*, 2, 389-394.
- Pinto, V. B., & Monteiro, V. B. (2024). A resistência do povo indígena Fulni-ô sob a perspectiva decolonial. *História Unicap*, 11(21), 284–306.
<https://doi.org/10.25247/hu.2024.v11n21.p284-306>
- Popovich, A. H., & Popovich, F. B. (2005). *Maxakalí–English dictionary; English–Maxakalí glossary*. Sociedade Internacional de Linguística.
- Porantim. (1989, March 31). Fulni-Ô : Os “guerreiros” de Sua Majestade. *Terras Indígenas no Brasil*.
- Pompeu Sobrinho, T. (1958). Línguas tapuias desconhecidas do Nordeste: Alguns vocabulários inéditos. *Revista do Instituto do Ceará*, 72, 153–187.
- Ribeiro, E. R., & van der Voort, H. (2010). Nimuendajú was right: The inclusion of the Jabutí language family in the Macro-Jê stock. *International Journal of American Linguistics*, 76(4), 517–570.
- Rodrigues, A. D. (1986). *Línguas brasileiras: Para o conhecimento das línguas indígenas*. Edições Loyola.

- Rodrigues, A. D. (2002). Línguas indígenas: 500 anos de descobertas e perdas. In M. C. B. Lopes & C. R. Cavalcante (Eds.), *O índio no Brasil : História, direitos e cidadania* (pp. 111–122). MEC/UNESCO.
- Rudolph, B. (1909). *Wörterbuch der Botokudensprache*. Friedrichsen & Co
- Ryskina, M., Rabinovich, E., Berg-Kirkpatrick, T., Mortensen, D. R., & Tsvetkov, Y. (2020). Where new words are born: Distributional semantic analysis of neologisms and their semantic neighbourhoods. In *Proceedings of the Society for Computation in Linguistics 2020* (pp. 367–376). Association for Computational Linguistics. <https://doi.org/10.7275/1jra-8m83>
- Santana, P. M. S. de, & Magalhães, T. Q. de. (2022). Caso Xukuru e o Bem Viver do povo Fulni-ô (PE) / The Xukuru Case and the Well-Living of the Fulni-ô People (PE). *Revista Direito e Práxis*, 13(1), 607–635. <https://doi.org/10.1590/2179-8966/2022/65133>
- Schippan, T. (2002). *Lexikologie der deutschen Gegenwartssprache*. Max Niemeyer Verlag.
- Seki, L. (1987). Apontamentos para a bibliografia da língua Botocudo/Borum. *Revista de Antropologia*, 30–32, 511–535.
- Seki, L. (2004). Aspectos da morfossintaxe Krenak : Orações independentes. *LIAMES*, 4, 129–146. <https://core.ac.uk/download/201255007.pdf>
- Shavit, Z. (2017). “Can it be that our dormant language has been wholly revived?” *Israel Studies*, 22(1), 101–138.
- Shi. (2025). An analysis of word-formation processes in neologisms: A case study of the Oxford Advanced Learner’s Dictionary. *Lexikos*, 35, 26–43. <https://doi.org/10.5788/35-1-1977>
- Silva, A. J. M. B. (2024, April 1). *Breathing new life into an almost-lost language*. Davis UWC Scholars. <https://www.davisuwcscholars.org/newsroom/breathing-new-life-almost-lost-language>
- Silva, A. J. M. B. (2025, August 20). Krenak legacy materials: Learning Indigenous languages from centuries-old sources. *Smithsonian Folklife Magazine*.
- Silva, E. (2007). “Nós vencemos a guerra!” *História, memórias e leituras indígenas da Guerra do Paraguai*. *Clio: Revista de Pesquisa Histórica*, 2(25), 39–58. <https://periodicos.ufpe.br/revistas/revistaclio/article/download/24721/19996>
- Silva, E. (2015). A história contada pelos que voltaram: História e memórias indígenas no Nordeste sobre a Guerra do Paraguai. *Diálogos*, 19(3), 1043–1063.
- Silva, E. H. (2007). História, memórias e identidade entre os Xukuru do Ororubá. *Tellus*, 7(12), 89–102.
- Silva, M. (2011). *Descrição gramatical da língua Yathê* [Doctoral dissertation, Universidade Federal de Pernambuco].
- Silva, W. D. L. (2012). *A descriptive grammar of Desano* [Doctoral dissertation, University of Utah].
- Spolsky, B. (2014). *The languages of the Jews: A sociolinguistic history*. Cambridge University Press.
- Steward, J. H. (Ed.). (1946–1959). Botocudos. *Handbook of South American Indians* (Vols. 1–7). Smithsonian Institution, Bureau of American Ethnology.
- Storto, L. (2019). *Toward the survival of Indigenous languages*. Pesquisa Fapesp.

- Telles, E., & Esteve, A. (2019). Racial intermarriage in the Americas. *Sociological Science*, 6, 293–320. <https://doi.org/10.15195/v6.a12>
- De Tássio Borges da Silva, P. (2019). *Políticas linguísticas de revitalização entre os Pataxó do território Kai-Pequi* [Master's Thesis, Universidade Federal do Rio de Janeiro, Museu Nacional].
- Trudgill, P. (2020). Sociolinguistic typology and the speed of linguistic change. *Journal of Historical Sociolinguistics*, 6(2). <https://doi.org/10.1515/jhsl-2019-0015>
- UNESCO. (2010). *Atlas of the world's languages in danger* (3rd ed.). UNESCO Publishing. <https://unesdoc.unesco.org/ark:/48223/pf0000187026>
- Wied-Neuwied, M. (1989). *Viagem ao Brasil*. Brasiliana.
- Woodward, K. (2018). Indigenous language revival: The Wôpanâak language reclamation project as a case study in Indigenous identity, representation, and place-based knowledge. *James Madison Undergraduate Research Journal*, 5(1), 71–80.
- Xavier de Oliveira, C. (2018). *Função, distribuição e as alterações do morfema /-NE/ da língua indígena Yaathe* [Bachelor's Thesis, Universidade Federal de Alagoas].
- Yule, G. (2023). *The study of language* (8th ed.). Cambridge University Press.

Apêndice A

Lista de fontes para neologismos

Xukuru	Borum	Patxohã
Substantivos: Da Silva Ribeiro, 2010 Lapenda, 1962 Pompeu Sobrinho, 1958	Abreu, 1929 Frassetto, 2018 Seki, 2004	Ministério de Minas e Energia, s.d. Coordenação Atxôhã, s.d. Oliveira Gonçalves, 2025 Bomfim, 2012
Numerais: -	Instituto Max Planck de Antropologia Evolutiva, s.d.)	Wied-Neuwied, 1989 Bomfim & Lima, 2020

Kariri-Xocó	Yaathé
Kariri-Xocó, 2020 Correia de Queiroz, 2012 Kariri-Xocó, N. (s.d.) Blog 2018 - 2026	Caetano de Sá, 1998 Meland & Meland, 2009 Lapenda, 1968
Correia de Queiroz, 2012 Kariri-Xocó, N (s.d.) Blog 2018-2026	Lapenda, 1968

Apêndice B

Distribuição de neologismos por processo de cunhagem. Quantidade e porcentagens.

Um neologismo que tenha utilizado vários processos de cunhagem será contado em cada um deles. O número total de neologismos classificados é 60, com um total de 86 instâncias de processos de cunhagem utilizados, uma média de 1,4 processos de cunhagem por neologismo.

	Xukuru	Yathé	Patxohã	Borum	Kariri-Xocó	
Empréstimos do Português	2	5	3	12	1	
Empréstimos de línguas indígenas vizinhas	6		5			
Composição		3	4	2	8	
Palavra-valise			3	1	5	
Derivação		4	3		2	
Conversão		1		2		
Truncamento		5			1	
Acrônimo						
Re-significação semântica						
Extensão semântica	2	2	2		2	
Total	10	20	20	17	19	
	Xukuru	Yathé	Patxohã	Borum	Kariri-Xocó	Total
Empréstimos do Português	20	25	15	71	5	27
Empréstimos de línguas indígenas vizinhas	60	0	25	0	0	13
Composição	0	15	20	12	42	20
Palavra-valise	0	0	15	6	26	10
Derivação	0	20	15	0	11	10
Conversão	0	5	0	12	0	3
Truncamento	0	25	0	0%	5%	7
Acrônimo	0	0	0	0	0%	0
Re-significação semântica	0	0	0	0	0%	0
Extensão semântica	20	10	10	0	11	9

Apêndice C**Abreviaturas nas glosas**

Abreviação	Significado
ADJ	Adjetivo
CAUS	Causativo
COMP	Complementador
FEM	Feminino
FUT	Futuro
IMP	Imperativo
IND	Indicativo
INDEF	Indefinido
INF	Infinitivo
LIM	Limitativo
NEG	Negação / Negativo
PASS	Passivo
PL	Plural
POSS	Possessivo
PRES	Presente
PRIV	Privativo
PST	Passado
REF	Referencial
S.REF	Autorreflexivo
SG	Singular