



OPEN ACCESS

EDITADO POR

- Maria Filomena Spatti Sandalo
(Unicamp)

AVALIADO POR

- Kristina Balykova (UT)
- Filomena Sandalo (Unicamp)

SOBRE OS AUTORES

- Ana Vilacy Galúcio
Conceitualização, pesquisa e interpretação dos dados, redação e revisão do artigo.

- Joshua Birchall
Conceitualização, pesquisa e interpretação dos dados, redação e revisão do artigo.

- Ivan Rocha
Conceitualização, redação e revisão do artigo.

- Saulo Rodrigo Teixeira Brito
Escrita - rascunho original, Metodologia, Software.
- Matheus Augusto Ribeiro Soares

Pesquisa e Redação do artigo.

- Israel Nathan Marcon
Pesquisa e Redação do artigo.

- Douglas da Costa Rodrigues Junior

Pesquisa e Redação do artigo.
- Ellison Santos

Investigação e redação parcial do artigo.

- Victor Siqueira Rocha
Pesquisa e Redação do artigo.

- Antônia Fernanda de Souza Nogueira

Pesquisa e Escrita - rascunho original.

- Flávio Henrique Ferreira Pinheiro

Escrita - rascunho original.

- Juliana Ketelen Souza Solano
Pesquisa e Redação do artigo

- Bruno Pinto Silva
Escrita - rascunho original.

ENSAIO TEÓRICO

Dicionários multimídia: uma tecnologia social para auxiliar a aprendizagem e revitalização de línguas minorizadas

Ana Vilacy GALÚCIO

Museu Paraense Emílio Goeldi (MPEG)

Joshua BIRCHALL

University of New Mexico (UNM)

Ivan ROCHA

Museu Paraense Emílio Goeldi (MPEG)

Saulo Rodrigo Teixeira BRITO

Museu Paraense Emílio Goeldi (MPEG)

Matheus Augusto Ribeiro SOARES

Universidade Federal do Pará (UFPA)

Israel Nathan MARCON

Universidade do Estado do Pará (UEPA)

Douglas da Costa RODRIGUES JUNIOR

Universidade Federal do Pará (UFPA)

Ellison SANTOS

Museu Paraense Emílio Goeldi (MPEG)

Victor Siqueira ROCHA

Universidade Federal do Pará (UFPA)

DATAS

- Recebido: 13/10/2025

- Aceito: 15/12/2025

- Publicado: 16/06/2026

COMO CITAR

Galúcio, Ana Vilacy *et al.* (2026).
Dicionários multimídia: uma
tecnologia social para auxiliar a
aprendizagem e revitalização de
línguas minorizadas. *Revista da
Abralín*, v. 24, n. 2, p. 116-144,
2025.

Antônia Fernanda de Souza NOGUEIRA 

Universidade Federal do Pará (UFPA)

Flávio Henrique Ferreira PINHEIRO 

Universidade Federal do Pará (UFPA)

Juliana Ketelen Souza SOLANO 

Universidade Federal do Pará (UFPA)

Bruno Pinto SILVA 

University of New Mexico (UNM)

RESUMO

Este artigo aborda a importância dos dicionários multimídia como um produto multifuncional e multiuso desenvolvidos como uma tecnologia social voltada para a aprendizagem de línguas ameaçadas. Os dicionários multimídia para línguas indígenas constituem uma metodologia reaplicável, de uso livre e fácil acesso, que pode ser usada pelas comunidades em larga escala e sem necessidade de acesso à internet. O uso de tecnologia como smartphones, tablets ou computadores, ainda que incipiente, é uma realidade em muitas comunidades indígenas, o que favorece a implementação de dicionários em formatos digitais, cujas entradas lexicais podem ser ilustradas com recursos multimídias. O artigo apresenta inicialmente o conceito, a metodologia e a tecnologia envolvidas na elaboração dos dicionários e descreve as macro e microestruturas adotadas nessas obras, decididas conforme a necessidade dos falantes ou potenciais falantes da língua. O trabalho discute ainda o papel dos dicionários multimídia no planejamento linguístico em práticas de revitalização de línguas minorizadas e ameaçadas, a partir da experiência de elaboração de dicionários de oito línguas indígenas em desenvolvimento no setor de linguística do Museu Paraense Emílio Goeldi (MPEG) em colaboração com a Universidade do Novo México (UNM, EUA): Kanoé, Sakurabiat, Oro Win, Djeoromitxi, Puruborá, Makurap, Wayoro e Salamã. No contexto de vulnerabilidade e ameaça de desaparecimento vivenciado pela maioria das línguas indígenas no Brasil, os dicionários multimídia surgem como uma ferramenta para auxiliar na revitalização de uma língua e podem ser úteis tanto no processo de transmissão das línguas quanto em seus processos de documentação e descrição.

ABSTRACT

This article discusses the importance of multimedia dictionaries as multi-functional and multipurpose products developed as a social technology aimed at the learning of endangered languages. Multimedia dictionaries for Indigenous languages represent a reusable, freely available, and easily accessible methodology that can be used by communities on a large scale without the need for internet access. The use of technology such as smartphones, tablets, or computers, though still incipient, is already a reality in many Indigenous communities, which favors the implementation of dictionaries in digital formats whose lexical entries can be illustrated with multimedia resources. The article first presents the concept, methodology, and technology involved in creating these dictionaries and describes the macro- and microstructures adopted in them, which are determined according to the needs of speakers or potential speakers of the language. It also discusses the role of multimedia dictionaries in language planning within the context of revitalization practices for minoritized and endangered languages, based on the experience of developing dictionaries for eight Indigenous languages at the Linguistics Division of the Museu Paraense Emílio Goeldi (MPEG) in collaboration with the University of New Mexico (UNM, USA): Kanoé, Sakurabiat, Oro Win, Djeoromitxi, Puruborá, Makurap, Wayoro and Salamã. In the context of the vulnerability and threat of extinction faced by most Indigenous languages in Brazil, multimedia dictionaries emerge as a tool to support language revitalization and can be useful both in the transmission of languages and in their documentation and description.

PALAVRAS-CHAVE

Línguas indígenas. Lexicografia. Revitalização de línguas.

KEYWORDS

Indigenous languages. Lexicography. Language revitalization.

RESUMO PARA NÃO ESPECIALISTAS

O cenário de desaparecimento experimentado pela ampla maioria das línguas indígenas faladas no Brasil vem mobilizando cada vez mais esforços, seja das comunidades de fala destas línguas ou da comunidade científica, em torno de práticas de revitalização linguística, isto é, ações voltadas para manter uma língua em uso. Este tipo de prática, muitas vezes, requer ferramentas que auxiliem no processo para que essas ações sejam mais eficazes e alcancem os objetivos desejados. O presente trabalho se propõe

a discutir uma dessas ferramentas, os dicionários multimídias de línguas indígenas, uma ferramenta digital que pode ser utilizada nas ações voltadas para a manutenção do uso de uma língua indígena. O trabalho apresenta a metodologia utilizada para a elaboração de dicionários multimídias digitais e discute a experiência de oito dicionários de línguas indígenas em elaboração no setor de linguística do Museu Paraense Emílio Goeldi em colaboração com a Universidade do Novo México, focando nos usos e nas características estruturais de cada um. Além disso, o trabalho também discute a inserção desta ferramenta no planejamento linguístico das comunidades e o papel desempenhado pelos dicionários multimídias digitais como ferramentas de suporte à revitalização das línguas indígenas.

Introdução

Este artigo aborda a importância dos dicionários multimídia como um produto multifuncional e multiuso, desenvolvido como uma tecnologia social replicável, de uso livre e fácil acesso, que pode ser usada pelas comunidades em larga escala e sem necessidade de acesso à internet. O termo *tecnologia social*¹ refere-se a um conjunto de saberes, metodologias, técnicas, conhecimentos científicos e tradicionais integrados para buscar solução para um ou vários problemas sociais. No caso do tema deste artigo, reverter ou reduzir as perdas linguísticas. O trabalho discute as propostas de elaboração de dicionários multimídia de línguas minorizadas em uma perspectiva voltada para a aprendizagem e revitalização das línguas, tendo como enfoque o planejamento linguístico necessário para a definição do escopo dos dicionários desenvolvidos no contexto de uma metodologia de base comunitária (*community-based methodology*) em que as comunidades de falantes estejam envolvidas no processo. De modo geral, a metodologia para a elaboração dos Dicionários Multimídia para Línguas Indígenas (DMLI) abordada neste trabalho consiste em compilar dados linguísticos e arquivos de mídia em um formato que possibilita um fácil compartilhamento e acesso (Brito; Birchall; Galucio, 2025).

Os Dicionários Multimídias para Línguas Indígenas aliam documentação de línguas indígenas e tecnologia digital de comunicação e informação, e podem ser usados em aplicativos móveis (celulares, tablets, computadores) sem necessidade de internet. A tecnologia é de utilização gratuita, com aplicação de softwares de acesso livre aberto, e pode ser utilizada pelas comunidades, nas escolas e em ambientes fora do espaço escolar. Essa Tecnologia Social envolve desenvolvimento de metodologias reaplicáveis e de produtos digitais, visando contribuir para a documentação de línguas indígenas, bem como fomentar sua aprendizagem e revitalização. O uso de tecnologia como smartphones, tablets ou computadores, ainda que incipiente, é uma realidade em muitas comunidades

1 A Rede de Tecnologia Social (RTS) define tecnologias sociais como “produtos, técnicas ou metodologias reaplicáveis, desenvolvidas na interação com a comunidade e que representem efetivas soluções de transformação social”.

indígenas, o que favorece a implementação de dicionários em formatos digitais, cujas entradas lexicais podem ser ilustradas com recursos multimídias (além de informações conceituais, definições, transcrição, campos semânticos, exemplos de uso etc.).

O trabalho apresenta a metodologia para elaboração dessa modalidade de dicionários, desenvolvida no setor de linguística do Museu Paraense Emílio Goeldi (MPEG) em colaboração com a Universidade de Novo México (UNM, EUA) e discute o papel desse tipo de dicionário no planejamento linguístico associado a práticas de revitalização de línguas indígenas minorizadas e ameaçadas, a partir da experiência de elaboração de dicionários multimídia para oito línguas indígenas, a saber: Kanoé, Sakurabiat, Oro Win, Djeoromitxi, Puruborá, Makurap, Wayoro e Salamã.

O restante do artigo está organizado da seguinte forma. Na seção 1, apresentamos a metodologia e a tecnologia envolvidas na elaboração dos dicionários multimídia, concebidos como uma tecnologia social para apoio à aprendizagem e revitalização de línguas. A seção 2 apresenta o contexto sociolinguístico das línguas Kanoé, Sakurabiat, Oro Win, Puruborá, Djeoromitxi, Wayoro, Makurap e Salamã e a metodologia, rotinas de trabalhos e tecnologia utilizadas no processo de criação dos dicionários multimídias para cada uma dessas línguas, que foram as primeiras aplicações da metodologia de desenvolvimento dos DMLI. Na seção 3, abordamos o papel dos dicionários multimídias, enquanto ferramenta didática no processo de transmissão linguística. Apresentamos as principais contribuições do artigo na seção 4.

1. Metodologia: Dicionários multimídia para línguas indígenas (DMLI) - uma tecnologia social com foco na aprendizagem e revitalização de línguas

Os dicionários são um dos principais instrumentos para descrever uma língua, além das coleções de textos e das gramáticas. A sua elaboração parte do registro do conhecimento lexical dos falantes, ou seja, das palavras e expressões que compõem o vocabulário de uma comunidade. Um dicionário pode ter muitos usos: servir como apoio no ensino, como fonte para estudos sobre o léxico ou como referência para a própria comunidade linguística. Seu alcance também pode variar: pode ser mais restrito, abordando apenas um tema específico, como nomes de plantas ou objetos da cultura material, ou mais amplo, reunindo informações de diferentes áreas do léxico. Para uma discussão mais detalhada sobre lexicografia e a produção de dicionários de línguas indígenas, ver Birchall (2023).

Um dos papéis principais de um dicionário é servir como referência para as comunidades linguísticas. Quando essas comunidades contam com poucos falantes ativos capazes de modelar a pronúncia e o uso das palavras, os recursos multimídia tornam-se especialmente valiosos. Dicionários multimídia são aqueles que incorporam recursos como clipes de áudio e vídeo em suas entradas, oferecendo exemplos de pronúncia das palavras e ilustrações de seu uso em frases mais complexas. Esses recursos multimídia permitem registrar e transmitir de forma mais fiel aspectos orais da

língua, como pronúncia e contexto de uso, funcionando não apenas como ferramenta de referência, mas também como instrumento de preservação e revitalização linguística, em um contexto de alta vulnerabilidade das línguas indígenas brasileiras.

A criação da tecnologia social Dicionários Multimídia de Línguas Indígenas foi motivada por uma demanda direta de comunidades indígenas (inicialmente em Rondônia) que, embora engajadas em movimentos autônomos de retomada linguística, careciam de ferramentas adequadas para fomentar o ensino e aprendizagem de suas línguas. Nesse cenário, em 2019, o Museu Paraense Emílio Goeldi, instituição que atua há décadas na documentação e estudo das línguas dos povos indígenas no Brasil, iniciou o desenvolvimento dos dicionários multimídia para línguas indígenas (DMLI), respondendo à demanda de alguns povos indígenas, por materiais que pudessem ser utilizados para ajudar na manutenção e aprendizagem de suas línguas de tradição. O objetivo dos dicionários multimídia é servir como uma tecnologia social voltada à documentação, preservação e revitalização de línguas indígenas ameaçadas, fomentando a aprendizagem e o uso dessas línguas através de ferramentas digitais acessíveis e construídas de forma colaborativa com as comunidades.

Ao integrar a elaboração desses dicionários ao próprio fluxo de trabalho de documentação linguística, busca-se aproveitar de materiais comumente presentes em acervos de línguas, como os acervos citados neste artigo (Acervo de Línguas Indígenas do Museu Paraense Emílio Goeldi-ALIM e Endangered Languages Archive -ELAR), de modo a facilitar a construção do dicionário e potencializar o aproveitamento do trabalho já realizado. Os formatos digitais de dicionários apresentam vantagens em relação aos formatos impressos, como a distribuição e o fácil acesso dos *apps* entre usuários de smartphones ou tablets, redução de custos e flexibilidade na organização da macroestrutura. Outra vantagem dessa metodologia é o versionamento, que permite o lançamento de atualizações graduais à comunidade de usuários do dicionário. Assim o dicionário pode passar por várias etapas de revisão e a comunidade pode ter acesso a versões atualizadas de forma imediata.

A seguir descrevemos a metodologia aplicada no desenvolvimento e elaboração dos dicionários multimídia, conforme o modelo que será apresentado na próxima seção.

Para apoiar e guiar o processo de elaboração de DMLI, estabelecemos algumas práticas gerais, como a estrutura de bancos de dados lexicais e convenções de nomear arquivos multimídia, e desenvolvemos alguns recursos computacionais, como o aplicativo CSV2RMD². Essa metodologia faz parte da documentação do aplicativo CSV2RMD (Brito; Birchall; Galucio 2025) e será detalhada a seguir³.

O aplicativo CSV2RMD permite gerar dicionários digitais a partir de arquivos CSV (*Comma Separated Values*) padronizados e preenchidos com dados lexicais e um conjunto organizado de arquivos de mídia, que podem ser convertidos tanto para o formato HTML (*HyperText Markup Language*) quanto para PDF (*Portable Document Format*). Um dicionário em formato HTML permite uma

²Estamos no processo de migrar o nome do aplicativo para CSV2DMLI, a fim de refletir o abandono do formato RMD (R Markdown) nas versões mais recentes do programa. Uma versão atualizada, com novas funcionalidades, está prevista para lançamento até o final de 2026, no mesmo repositório do GitHub.

³ Outro modelo de elaboração de dicionários multimídia para línguas indígenas no Brasil é aquele aplicado nos dicionários disponibilizados no Portal Japiim (<https://japiim.museudoindio.gov.br/>), desenvolvido no âmbito do Projeto de Documentação de Línguas Indígenas (Pro-Doclin) do Museu dos Povos Indígenas em parceria com a Unesco (<http://prodoclin.museudoindio.gov.br/>).

organização mais dinâmica do conteúdo, com vantagens como a possibilidade de reestruturação dinâmica e a inclusão de mídias, como áudio e vídeo, que não poderiam ser utilizadas diretamente em outros formatos mais tradicionais, além do encapsulamento desse formato em APK para dispositivos Android. Já o documento em PDF, embora não possibilite a manipulação dinâmica de sua estrutura ou a adição de mídias em áudio ou vídeo, permite criar dicionários em um formato mais similar aos tradicionais, que podem ser impressos e distribuídos onde a utilização de dispositivos eletrônicos não está disponível.

As ferramentas computacionais utilizadas no processo de elaboração desses recursos multimídia são: os softwares de transcrição e anotação linguísticas (ELAN, 2025); um programa elaborado em Python para a automatização na extração de dados segmentados (texto e áudio) de arquivos ELAN (.eaf) para planilhas CSV e posterior conversão para HTML; uma distribuição do LaTeX chamada MiKTeX, necessária para a produção de arquivos PDF. O fluxo de trabalho utilizado abrange desde a organização dos dados até a exportação dos produtos e pode ser sumarizado em algumas etapas: **preparação de dados e metadados, configuração e verificação, e criação de produtos.**

A **preparação de dados e metadados** envolve a organização dos materiais necessários para a criação dos produtos, seguindo modelos previamente definidos. Essa etapa inclui organização das mídias em pastas padronizadas e preenchimento do banco de dados lexical em formato CSV, que pode ser elaborado manualmente ou gerado a partir de informações transcritas em arquivos no formato .eaf utilizando uma função de CSV2RMD desenvolvido para essa finalidade. Cada entrada no banco de dados lexicais deve conter, no mínimo, três campos obrigatoriamente preenchidos: item lexical, tradução/significado e campo semântico. Além disso, podem ser incluídos campos opcionais, tais como: imagem, arquivo sonoro, transcrição fonêmica, transcrição fonética, classe gramatical, descrição, arquivo sonoro do exemplo, transcrição do exemplo, tradução do exemplo, arquivo de vídeo, subcampo semântico e itens relacionados. A macro e microestrutura de cada dicionário são definidas conforme a necessidade dos falantes ou potenciais falantes de cada língua.

Com os dados devidamente organizados, o usuário pode realizar a **configuração e verificação**, na qual, a partir do primeiro uso do programa principal, pode validar o preenchimento da tabela CSV, buscando erros e corrigindo-os, ou ainda adicionar arquivos de configuração que permitem a customização e o preenchimento dos documentos multimídia gerados. Através desses arquivos de configuração, é possível, por exemplo, customizar a ordem alfabética ou adicionar uma introdução textual ao documento.

Uma vez validado e configurado na etapa anterior, o ambiente de trabalho pode ser utilizado para a etapa final de **criação de produtos**, que permite gerar (através do programa principal) documentos HTML ou PDF (figura 1). Esses documentos podem, então, ser revisados e analisados para divulgação ou, ainda, ajustados, caso sejam necessárias modificações, reiniciando assim um novo ciclo de atividades.

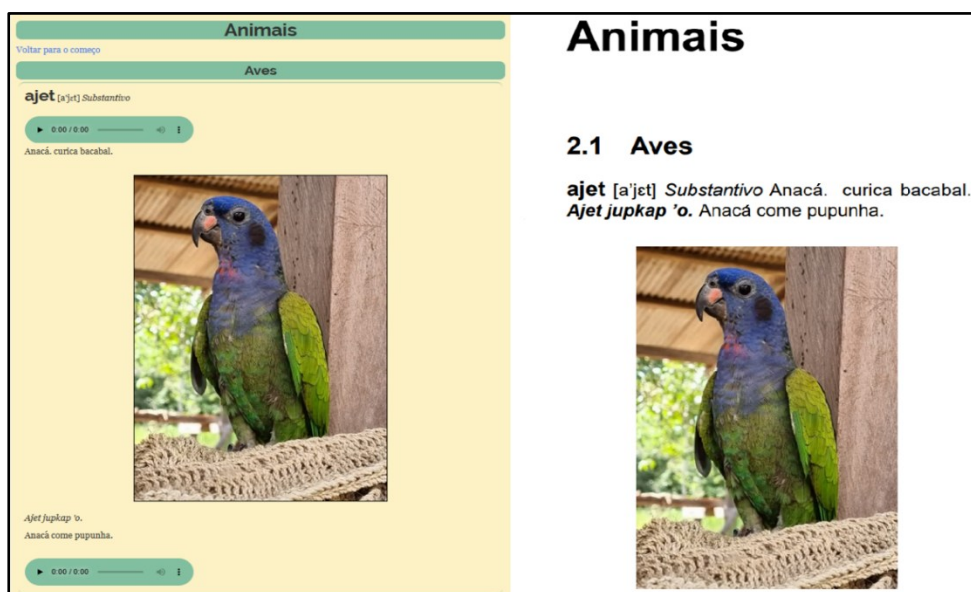


FIGURA 1 – Amostras da versão HTML online (esquerda) e PDF (direita) do Dicionário Puruborá.
Fonte: Soares et al. (2024).

Os produtos HTML gerados podem conter arquivos de mídia embutidos ou vinculados externamente. A principal vantagem da mídia embutida é permitir o acesso ao conteúdo sem necessidade de conexão à internet. Por outro lado, isso aumenta significativamente o tamanho do arquivo final no dispositivo do usuário. A produção de aplicativos Android, com o HTML do dicionário integrado, ainda não está totalmente consolidada, pois depende do uso de um script externo online para sua geração. Atualmente, esses aplicativos podem ser compartilhados de maneira informal, mas não podem ser publicados em lojas oficiais de aplicativos. Essa etapa do fluxo de trabalho segue em desenvolvimento, visando automatizar a criação de arquivos APK que atendam aos requisitos técnicos necessários para serem disponibilizados em plataformas como a Google Play Store a partir da versão HTML do dicionário.

Todas as ferramentas e práticas aqui descritas foram utilizadas na produção de dicionários e essas atividades, por sua vez, possibilitaram diversas melhorias nas próprias práticas, permitindo hoje a distribuição em livre acesso de manuais e ferramentas que viabilizam a construção simplificada de DMLIs, com baixa necessidade de conhecimento técnico. As ferramentas estão disponíveis e podem ser acessadas na plataforma GitHub que contém um repositório destinado ao registro e desenvolvimento do sistema⁴, permitindo a colaboração entre os interessados na utilização e melhoria desses processos (cf. Brito; Birchall; Galucio, 2025).

A aplicação dos procedimentos metodológicos de elaboração dos DMLIs aqui descritos permite ao usuário gerar tabelas de dados e recortar áudios correspondentes a itens lexicais e seus exemplos

⁴ O conteúdo do repositório <https://github.com/MPEG-Linguistica/CSV2RMD> inclui programas que podem ser baixados em formato para Windows (.exe) ou em Python e um manual de utilização em formato HTML e PDF.

de uso a partir de anotações padronizadas no ELAN, realizar ajustes em imagens e áudios, bem como a produção dos arquivos em PDF e HTML. A aplicação dessa metodologia será ilustrada na próxima seção, através de oito exemplos de projetos de dicionários multimídia, desenvolvidos no âmbito da cooperação entre o Museu Paraense Emílio Goeldi e a Universidade de Novo México, para o desenvolvimento dos DMLIs.

2. Dicionários multimídia: aplicação da metodologia

Nesta seção, apresentamos oito estudos de caso de aplicação da metodologia de desenvolvimento dos dicionários multimídia (Brito; Birchall; Galucio, 2025), a saber: os dicionários das línguas Kanoé, Sakurabiat, Oro Win, Puruborá, Djeoromitxi, Makurap, Wayoro e Salamã. Inicialmente, descreve-se brevemente a caracterização do povo e a situação sociolinguística de cada uma das línguas indígenas discutidas no artigo e, em seguida, apresenta-se o processo de produção e a caracterização do dicionário de cada língua, destacando a abordagem colaborativa com as respectivas comunidades de fala.

Todas as oito línguas abordadas neste trabalho são faladas por povos indígenas que habitam na região do atual estado de Rondônia (Brasil) desde tempos imemoriais. Em todos os casos, o contato desses povos com a sociedade não-indígena intensificou-se no início do século XX, durante o ciclo da borracha, provocando grandes mudanças culturais. Há registros de epidemias de doenças com resultados catastróficos para essas populações, escravidão por empresas seringalistas e o deslocamento forçado de seus territórios tradicionais. A situação sociolinguística desses povos, com relação à manutenção e vitalidade de suas línguas tradicionais, reflete os resultados dessa combinação de fatores históricos violentos, consequentes do contato com não-indígenas. O Quadro 1, a seguir, apresenta o panorama de cada uma das línguas dos dicionários elaborados, detalhando sua classificação e situação atual quanto à vitalidade e ao contexto de uso.

Língua	Código Iso	Família/Ramo Linguístico	Localização/Terras Indígenas Principais	Situação
Kanoé	kxo	Isolada	TI Rio Guaporé, TI Rio Branco, TI Sagarana e TI Omerê.	Severamente ameaçada. Deixou de ser transmitida há 2 gerações. Apenas 4 falantes fluentes conhecidos (3 na TI Omerê + 1 na TI Rio Guaporé).
Sakurabiat	skf	Tupi (Tupari)	TI Rio Mequéns (RO).	Severamente ameaçada. Apenas 8 falantes fluentes (Mar/2025) e cerca de 10 parciais. Transmissão interrompida há 3 décadas.
Oro Win	orw	Txapakura	TI Uru-Eu-Wau-Wau (Guajará-Mirim, RO).	Severamente ameaçada. 6 falantes com fluência nativa (mais de 60 anos), menos de 20 com alguma competência produtiva.

Puruborá	pur	Tupi (Rama-rama-Puruborá)	Aldeia Aperi (Seringueiras, RO) e outros municípios de Rondônia.	Quebra de transmissão. 1 ancião que lembra palavras + 1 novo falante/professor aprendendo.
Djeoromixi	jbt	Macro-Jê (Jabutí)	TI Rio Guaporé e TI Rio Branco (Sul de RO).	Severamente ameaçada. 35 falantes fluentes e 50-60 com fluência passiva (Censo INDL/IPHAN, 2018).
Wayoro	wyr	Tupi (Tupari)	Porto Rolim de Moura do Guaporé e TI Rio Guaporé (RO).	Severamente ameaçada. Não é mais falada no dia a dia. 1 falante fluente (idoso) e cerca de 20 semi-falantes.
Makurap	mpu	Tupi (Tupari)	TI Rio Guaporé e TI Rio Branco (RO).	Severamente ameaçada. Número de falantes reduzido (24 a 30 pessoas). Transmissão intergeracional interrompida.
Salamái	mnd	Tupi (Mondé)	Região entre os rios Piamenta Bueno e Barão de Melgaço (RO).	Severamente ameaçada. Em 2021, apenas duas mulheres idosas lembravam palavras.

QUADRO 1 – Panorama das Línguas Indígenas Abordadas.

Fonte: Elaborado pelos autores.

A elaboração dos dicionários multimídia é resultado de um processo de pesquisa colaborativa entre os pesquisadores não-indígenas e as comunidades e embasado em pesquisas linguísticas realizadas ao longo de anos. Para o registro das informações constantes nos dicionários, a documentação das línguas é realizada por pesquisadores indígenas e não-indígenas em parceria com falantes nativos das línguas específicas. Alguns DMLIs produzidos seguindo a metodologia e usando a tecnologia descritas neste artigo aproveitaram da documentação linguística disponível no Acervo de Línguas Indígenas do Museu Paraense Emílio Goeldi (ALIM), como o dicionário multimídia Kanoé-Português. O ALIM é uma coleção de valor inestimável, caracterizando-se como um acervo multimídia que reúne uma vasta documentação sobre a diversidade linguística da Amazônia. Este acervo é um pilar fundamental para as pesquisas e a preservação do patrimônio linguístico e cultural dos povos indígenas da região⁵. Para os outros dicionários, foi possível coletar dados com membros falantes das respectivas comunidades em trabalho colaborativo, seja em visitas nas aldeias, seja através de visita dos falantes ao laboratório de Linguística do Museu Paraense Emílio Goeldi, conforme será apresentado a seguir.

A quantidade de entradas, a macroestrutura e a microestrutura são definidas de modo particular para cada dicionário, dependendo em parte da disponibilidade de dados disponíveis e/ou da priorização definida em conjunto com cada comunidade de fala. Todos os dicionários são organizados por campos e, em alguns casos, subcampos, semânticos. Avaliou-se, em conjunto com as comunidades de fala, que essa organização era mais apropriada que uma organização em ordem alfabética⁶. Porém,

⁵ Atualmente o ALIM/MPEG possui registros de 80 línguas.

⁶ Nas versões mais atuais, é possível gerar também uma listagem completa em ordem alfabética em português ou na língua vernácula do dicionário, além da organização por campo semântico.

a definição dos campos semânticos utilizados é específica de cada dicionário, não há uma padronização nesse sentido. Com relação à microestrutura dos verbetes, todos os dicionários apresentam Lema ou Entrada Lexical, acompanhado de áudio e imagem ilustrativa (sempre que possível), transcrição fonética e/ou fonêmica, informação sobre classe gramatical, tradução para português e exemplo ilustrativo contextualizado de uso da entrada lexical em uma sentença, seguido de tradução e acompanhado de áudio, sempre que possível. Porém, os dados que compõem a microestrutura são preenchidos conforme a disponibilidade de informações para cada língua, havendo, por conta disto, itens lexicais que não possuem todos estes campos preenchidos. O Quadro 2, abaixo, sintetiza as principais características e o status atual de cada um dos oito dicionários discutidos.

Língua	Nº de Entradas Lexicais	Macroestrutura	Microestrutura	Status	Data/Previsão
Kanoé	457	13 Campos Semânticos.	Lema ou Entrada lexical na ortografia da língua (áudio quando disponível), Transcrição Fonética, Classe Gramatical, Tradução (português), Imagem, Exemplo de uso com tradução (áudio quando disponível).	Concluído, em uso pela comunidade.	2020 (Versão 1.0)
Sakurabiat	646	11 Campos Semânticos (com subdivisões).	Lema ou Entrada lexical na ortografia da língua (áudio quando disponível), Transcrição Fonética, Classe Gramatical, Tradução (português), Imagem, Exemplo de uso com tradução (áudio quando disponível).	Concluído, em uso pela comunidade.	2024 (Versão 3.0)
Oro Win	858	11 Campos Semânticos.	Lema ou Entrada lexical na ortografia da língua (áudio quando disponível), Transcrição Fonêmica e Fonética, Classe Gramatical, Observações Culturais, Exemplo Contextualizado (texto e áudio).	Concluído, em uso pela comunidade.	2025 (Versão 3.0)
Puruborá	727	11 Campos Semânticos (com subdivisões).	Lema ou Entrada lexical na ortografia da língua (áudio quando disponível), Transcrição Fonética, Classe Gramatical, Tradução (português), Imagem, Exemplo de uso com tradução (áudio quando disponível).	Concluído, em uso pela comunidade.	2024 (Versão 3.0)

Salamã	114	3 Campos Semânticos.	Lema ou Entrada lexical na ortografia da língua (áudio quando disponível), transcrição fonética, Classe Gramatical, Tradução (português).	Concluído, em uso pela comunidade.	2024
Djeoromitxi	800	11 Campos Semânticos.	Lema ou Entrada lexical na ortografia da língua (áudio quando disponível), Transcrição Fonética, Classificação Gramatical, Tradução (português), Imagem, Exemplo de uso com tradução (áudio quando disponível).	Em processo de Revisão/validação.	2026
Wayoro	643	11 Campos Semânticos.	Lema ou Entrada lexical na ortografia da língua (áudio quando disponível) Classe de Palavra, Tradução (português), Imagem, Exemplo de uso com tradução (áudio quando disponível).	Em processo de finalização.	2026
Makurap	1.075	15 Campos Semânticos.	Lema ou Entrada lexical na ortografia da língua (áudio quando disponível), Transcrição Fonética, Classe Gramatical, Tradução (português), Imagem, Exemplo de uso com tradução (áudio quando disponível).	Em processo de finalização.	2026

QUADRO 2 – Principais características dos dicionários abordados.

Fonte: Elaborado pelos autores.

A elaboração dos dicionários segue os procedimentos metodológicos descrito na seção 1 acima e pode ser descrita em seis eixos principais: (i) coleta de dados em campo com falantes da língua em questão e/ou pesquisa e uso de listas de palavras do Acervo ALIM/MPEG; (ii) segmentação e transcrição fonética e ortográfica de mídias de áudio (iii) tradução para o português conforme as interpretações dos falantes (iv) estruturação da base lexicográfica em formato digital por campos semânticos; (v) seleção de imagens e (vi) processamento técnico para a geração do produto final.

Os cinco primeiros dicionários apresentados nesta seção (Kanoé, Sakurabiat, Oro Win, Puruborá e Salamã) já estão concluídos e em uso pelas respectivas comunidades e podem ser acessados através do portal de Dicionários Multimídia do Museu Paraense Emílio Goeldi⁷.

⁷ Ver em <https://dicionarios.museu-goeldi.br/>

O dicionário Kanoé foi o primeiro dicionário produzido utilizando a metodologia dos DMLIs, respondendo a uma demanda específica da comunidade Kanoé da Terra Indígena Rio Guaporé, através do cacique José Augusto Kanoé, por uma ferramenta que auxiliasse no processo de aprendizagem da língua e que possibilitasse escutar a pronúncia das palavras. De modo geral, a transmissão intergeracional da língua tradicional do povo Kanoé foi interrompida há cerca de duas gerações. Apenas na Terra Indígena Rio Omerê ainda há um pequeno grupo de pessoas da etnia, composto por 3 adultos, que utilizam a língua na comunicação diária. Entretanto, a comunidade Kanoé está interessada em resgatar sua língua tradicional e vem desenvolvendo estratégias voltadas para sua revitalização (Aquino; Kanoé, 2024). Nesse contexto é que foi desenvolvido o Dicionário Multimídia Kanoé – Português (Marcon; Galucio; Brito, 2023), iniciado em 2019 e pensado para responder à demanda da comunidade uma vez que devido ao contexto sociolinguístico, materiais somente impressos não satisfaziam a necessidade da comunidade. O dicionário Kanoé foi elaborado com base no acervo da língua disponível no Arquivo de Línguas Indígenas do Museu Goeldi – ALIM (versão 1.0 foi produzida em 2020)⁸. Foram transcritas gravações de itens lexicais e frases realizadas pelo linguista Denny Moore (MCTI/Museu Goeldi), com D. Teresa Piraguê, na década de 1990, e por Ana Vilacy Galucio (MCTI/Museu Goeldi), com o Sr. Francisco Kanoé, em 2008, na Terra Indígena Rio Guaporé. Devido à limitação de dados disponíveis, é um dos menores dicionários produzidos, com menos de 500 entradas lexicais e apenas uma pequena porcentagem das entradas apresenta frases ilustrativas do contexto de uso. Os dados disponíveis permitiram compor o dicionário com entradas nos campos semânticos referentes a animais, corpo, cores, eventos, localizações, manufaturas, natureza, números, parentescos, pessoas, plantas, propriedades e outros.

O dicionário Sakurabiat foi iniciado em 2022. Embora a língua tradicional do povo Sakurabiat⁹ não esteja sendo transmitida para as novas gerações há pelo menos três décadas (Galucio, 2021), há interesse entre os jovens adultos de fortalecer a língua e retomar sua aprendizagem (Guaratira; Costa, 2020; Galucio, 2020). Neste contexto de interesse pela aprendizagem da língua, vinculado ao ensino da mesma no sistema escolar, os professores indígenas e outros membros da comunidade solicitaram consultoria e auxílio dos linguistas para a elaboração de material didático-pedagógico para alfabetização e ensino-aprendizagem da língua dentro e fora do espaço escolar (Costa, 2020a). Foi inicialmente produzido um livro de apoio à alfabetização e aprendizagem da língua (Costa, 2020b) e iniciada a elaboração do dicionário multimídia Sakurabiat-Português em colaboração direta com a comunidade (Soares; Rodrigues Junior; Galucio, 2024). O dicionário foi elaborado a partir de corpus da coleção digital Sakurabiat hospedada no ALIM/MPEG¹⁰ e de registros complementares de áudios

⁸ Houve tentativa por parte da equipe de acessar o acervo Kanoé depositado no Museu dos povos indígenas/FUNAI, no Rio de Janeiro, mas devido a dificuldades advindas durante a pandemia COVID-19, não foi possível. Para a ampliação do dicionário, pretende-se incluir dados desse acervo.

⁹ Na literatura e em documentos pessoais e oficiais o nome da etnia também ocorre registrado como Sakyrabiar e Saquirabiar. Na década de 1980, o povo adotou a autodenominação Sakurabiat, abrangendo todos os subgrupos ou clãs, porém mais recentemente tem retomado a referência aos subgrupos específicos e suas variedades da língua, como as variedades Guaratira e Sakurabiat.

¹⁰ A coleção da língua Sakurabiat no Acervo de Línguas Indígenas do Museu Goeldi (ALIM) vem sendo criada desde a década de 1990, a partir de projetos coordenados por Ana Vilacy Galucio e possui atualmente cerca de 150 horas de gravação audiovisual, distribuídas em diferentes gêneros de fala.

para as entradas lexicais e frases ilustrativas de uso dos verbetes, especialmente com os falantes Olimpio Saquirabiar, Rosalina Guaratira Sakyrabiar, e Severino Guaratira Sakyrabiar. Após revisão e validação pelos falantes e outros membros da comunidade, o dicionário foi publicado e disponibilizado para uso pela comunidade em junho de 2024 (Galucio et al., 2024). Este dicionário tem a especificidade de contemplar as duas variedades da língua ainda em uso na comunidade, a variedade Sakurabiat e a variedade Guaratira.

A elaboração do dicionário Oro Win iniciou em 2019, através de um projeto financiado pelo programa *Documenting Endangered Languages* (DEL) dos Estados Unidos¹¹. A população Oro Win atual é de aproximadamente 150 pessoas, a maioria das quais reside na Terra Indígena Uru-Eu-Wau-Wau. Apenas seis pessoas falam a língua com fluência nativa, e menos de vinte possuem alguma competência produtiva (Birchall, 2018). Todos os falantes fluentes têm mais de 60 anos, enquanto os demais falantes estão na faixa etária entre 30 e 60 anos. A língua ainda é usada como meio de comunicação entre os anciãos e, ocasionalmente, entre estes e os outros falantes. Nos últimos anos, a comunidade Oro Win implementou um programa de ensino da língua tradicional nas escolas locais (Birchall; Oro Win Cabixi, 2021).

O objetivo principal do projeto do dicionário Oro Win foi capacitar pesquisadores Oro Win em métodos de documentação linguística, de modo que eles pudessem registrar, em formato audiovisual, o conhecimento léxico dos falantes da língua e elaborar um dicionário multimídia em parceria com o linguista¹². Dois pesquisadores Oro Win, Silvânia Oro Eo' Cabixi e Luciano Oro Win, foram indicados pela comunidade, após demonstrarem proficiência na compreensão e escrita da língua, e receberam treinamento em técnicas de condução de entrevistas linguísticas para elicitação de dados lexicais e no uso do programa ELAN para transcrição. Com a chegada da pandemia de COVID-19, no início de 2020, a presença do linguista na aldeia tornou-se inviável até o lançamento da primeira versão do dicionário em 2022. Nesse período, os pesquisadores indígenas assumiram a continuidade da documentação linguística dentro da própria comunidade, demonstrando como uma metodologia de pesquisa de base comunitária pode gerar resultados não apenas relevantes do ponto de vista científico, mas também fortalecedores do protagonismo indígena na preservação e transmissão de sua própria língua e cultura. A versão mais recente do Dicionário Oro Win (3.0) foi entregue à comunidade em junho de 2025 (Birchall; Oro Eo' Cabixi; Oro Win, 2025). O dicionário reúne exemplos extraídos de narrativas do acervo Oro Win disponíveis no ALIM, bem como de trabalhos de campo voltados à elicitação de itens lexicais, realizados pelos autores do dicionário. Após uma revisão conduzida com falantes da língua em julho de 2025, está em preparação uma versão final para publicação formal impressa por uma editora científica de acesso aberto, bem como para a atualização da próxima versão do dicionário multimídia online.

¹¹ Concessão de uma bolsa de pesquisa ao linguista Joshua Birchall.

¹² A coleção de documentação Oro Win no ALIM vem sendo desenvolvida em parceria com a comunidade linguística desde 2009. Atualmente essa coleção inclui cerca de 58 horas de produtos audiovisuais, distribuídos entre narrativas, diálogos e elicitaciones. Os dados coletados entre 2019 e 2022 como parte do projeto DEL também foram depositados no *Archive of Indigenous Languages of Latin America* (AILLA) na Universidade do Texas, disponível em <https://ailla.utexas.org/collections/1721/>.

O dicionário multimídia Puruborá foi elaborado a partir de corpus da coleção digital Puruborá, que está hospedada no ALIM/MPEG¹³ e de registros complementares de áudios para as entradas lexicais e frases ilustrativas de uso dos verbetes. A pesquisa para a elaboração do dicionário Puruborá centrou-se inicialmente nas gravações realizadas entre os anos 2001 e 2007 com os dois últimos anciãos mais fluentes em Puruborá, o Sr. Paulo Aporete Filho e o Sr. José Evangelista Puruborá. No início da elaboração do dicionário, em 2021, o professor Mário Oliveira Neto Puruborá compôs a equipe de pesquisa e participou de todas as etapas, inclusive contribuindo com dados linguísticos para a composição do dicionário. Devido ao longo e violento processo de exploração vivenciado pelos Puruborá, a língua do povo sofreu uma profunda quebra de transmissão intergeracional e interrupção em seu uso e não foi falada durante muitos anos, chegando quase a desaparecer (Galucio, 2005). Com o falecimento de três anciãos, entre 2020 e 2023, há somente um ancião que ainda lembra algumas palavras da língua e o professor Mário Oliveira Neto Puruborá que vem aprendendo a língua nas últimas duas décadas a partir dos produtos oriundos da documentação linguística. No contexto de fortalecimento e revitalização cultural desenvolvido pelos Puruborá (Oliveira Neto, 2020; Galucio, 2020), o professor Mário Puruborá tem ensinado a língua com o apoio do dicionário. O dicionário Puruborá é um dos mais completos, no sentido de que todas as entradas lexicais possuem os respectivos áudios dos falantes pronunciando a palavra. Após validação e verificação pelos usuários, em 2023, a versão mais recente do dicionário Puruborá (Soares et al., 2024) foi disponibilizada para uso da comunidade Puruborá, em 2024, e está sendo utilizada pelos membros da comunidade, tanto no contexto escolar na aldeia Aporoi, nas aulas de aprendizagem da língua e cultura tradicional Puruborá, quanto nas aulas online de revitalização da língua e nos espaços externos à escola (cf. Mário Oliveira Neto, comunicação pessoal, 19 de setembro de 2025).

A documentação linguística que serve de base ao Dicionário Salamã inseri-se no contexto do projeto *A Composição Etnolinguística do Sudoeste de Rondônia*,¹⁴ desenvolvido no âmbito do programa internacional DoBeS (*Dokumentation Bedrohter Sprachen*), financiado pela Fundação Volkswagen, sob a coordenação de Hein van der Voort. Esse projeto produziu um registro multimídia abrangente das línguas e culturas indígenas do sudeste de Rondônia – especialmente dos povos Aikanã, Kwaza e seus vizinhos imediatos, entre eles os Salamã, incluindo narrativas tradicionais, histórias pessoais, diálogos, cantos e dados linguísticos. As gravações utilizadas especificamente na elaboração do Dicionário Salamã foram realizadas por Joshua Birchall, em 2016 e 2019 com Dona Peridalva Wanzerip, uma das últimas pessoas que falam a língua Salamã. As gravações consistem principalmente em sessões de elicitación de palavras isoladas e formas verbais flexionadas, além de registros de músicas e histórias pessoais¹⁵. Durante o trabalho de campo, Birchall restabeleceu

¹³ Resultado de um projeto de documentação da língua Puruborá, coordenado por Ana Vilacy Galucio, no período de 2001 a 2007, a Coleção Puruborá no Acervo do Museu Goeldi possui cerca de 91 horas de gravação audiovisual, distribuídas entre elicitaciones, relatos pessoais e tradicionais do povo.

¹⁴ Disponível em https://dobes.mpi.nl/projects/southeastern_rondonia.

¹⁵ Todo o material processado até o momento encontra-se depositado no Acervo de Línguas Indígenas do Museu Paraense Emílio Goeldi (ALIM) e no acervo DoBeS como parte do The Language Archive.

contato com familiares de Dona Peridalva que vivem na região do Guaporé, que expressaram o desejo de que os materiais coletados fossem utilizados para apoiar o estudo de sua língua.

O Dicionário Salamã foi elaborado a partir dos dados coletados por Joshua Birchall e trabalhados posteriormente por Bruno Pinto Silva, doutorando em Linguística pela Universidade do Novo México (Pinto Silva; Birchall, 2024). O dicionário reúne 114 itens lexicais provenientes de cinco arquivos de áudio de duração total de 2 horas e 12 minutos, abrangendo os principais campos semânticos: plantas e alimentos (25 itens), animais (73) e parentesco (16). A macroestrutura organiza os verbetes segundo esses domínios temáticos e a microestrutura reúne as informações internas de cada entrada, conforme Quadro 2, acima. Uma particularidade deste dicionário foi a inclusão, sempre que possível, de informações sobre variação morfológica e, nos casos de verbos, formas flexionadas. Convenções ortográficas foram estabelecidas de modo a manter relativa proximidade com o português, favorecendo a legibilidade e o uso por falantes bilíngues. A definição do sistema gráfico e a análise fonológica preliminar basearam-se em referências como Hanke (1950), Galucio et al. (2015) e Felzke e Moore (2019). Embora o corpus ainda não permita uma análise detalhada relativa à distinção tonal, a duração das vogais foi identificada e marcada (vogais longas são indicadas pela duplicação da letra correspondente). O resultado é um produto lexicográfico bilíngue, multimodal e acessível, que alia rigor linguístico e relevância comunitária.

Os três dicionários apresentados a seguir, das línguas Djeoromitxi, Makurap e Wayoro, estão em fase de finalização, com previsão de conclusão e publicação na plataforma digital de dicionários em 2026.

O dicionário multimídia Djeoromitxi está sendo elaborado por Juliana Solano e Ivan Rocha, com a comunidade Djeoromitxi, a partir de corpus da coleção digital Djeoromitxi, que está hospedada no Arquivo de Línguas Indígenas do Museu Goeldi (ALIM/MPEG) e no Endangered Languages Archive (ELAR)¹⁶. Os Djeoromitxi (230 indivíduos, segundo registro do SIASI/SESAI do ano de 2020) estão distribuídos na Terra Indígena Rio Guaporé e na Terra Indígena Rio Branco (onde há parte da população, mas não há registro de falantes). A língua Djeoromitxi é considerada severamente ameaçada pela UNESCO devido à falta de transmissão da mesma entre gerações. O censo do INDL/IPHAN realizado em 2018 pela equipe do Museu Goeldi mostrou que 35 pessoas falam fluentemente o idioma, além disso, identificou que entre 50 e 60 pessoas possuem fluência passiva, grupo que poderia ser público alvo para o processo de revitalização da língua (Moore, 2018). Em 2022, iniciou-se o projeto de documentação da língua Djeoromitxi, financiado pelo Endangered Languages Documentation Programme/ELDP (Rocha, 2022), originando uma ampla coleção audiovisual, incluindo 5 horas de textos anotados (transcritos e traduzidos) passíveis de integrar o dicionário multimídia com exemplos de uso contextual. O dicionário Djeoromitxi e uma gramática pedagógica (ambos em fase de revisão) são produtos resultantes do trabalho de documentação linguística.

¹⁶ A coleção Djeoromitxi, no Acervo de Línguas Indígenas do Museu Goeldi (ALIM), vem sendo constituída desde a década de 1990 pelos linguistas Denny Moore, Hein van der Voort, Nádia Pires e Thiago Castro. O projeto de documentação Djeoromitxi (Rocha, 2022) ampliou esse acervo, com aproximadamente 170 horas de produtos audiovisuais, organizados em diferentes gêneros discursivos e elicitacoes de léxico e gramática, depositados no ALIM/MPEG e no ELAR.

A versão atual do dicionário encontra-se em fase de revisão e validação pela comunidade e deve estar disponível para uso em 2026. A metodologia adotada para a elaboração do dicionário segue as etapas já descritas no início desta seção. Cada entrada lexical do dicionário multimídia foi organizada para integrar informações linguísticas, pedagógicas e multimodais. Como exemplo, apresenta-se o item lexical *hibzi* ‘chicha’, conforme figura 2 abaixo. A entrada contempla: a forma ortográfica padronizada da palavra em Djeoromitxi, conforme a proposta de Pires (1992); a transcrição fonética [ˈhi.bzi], que registra sua realização sonora; a classificação gramatical como “nome”, por designar uma entidade concreta; e a tradução para o português, possibilitando a identificação do equivalente mais próximo. Além disso, inclui uma frase exemplificadora, que demonstra o uso do vocábulo em contexto e evidencia sua função sintática e semântica na oração. O material também é composto por um arquivo sonoro, gravado por falantes Djeoromitxi, e por uma ilustração, a fim de reforçar a associação visual e apoiar o processo de aprendizagem. Por fim, a palavra é inserida no campo semântico “alimentação”, categoria responsável por organizar o vocábulo dentro de um grupo temático específico.

hibzi [hi.ˈbzi] *nomes*

0:00 / 0:00

chicha

hoo *hibzi* nõ dje.
vamos tomar chicha.

0:00 / 0:00

hũ *hibzi* nõ.
eu tomei chicha.

0:00 / 0:00

adje *hibzi* nõ.
você foi tomar chicha.

FIGURA 2 – Verbetes *hibzi* ‘chicha’ no dicionário multimídia Djeoromitxi

Fonte: elaborada pelos autores.

Esse processo metodológico permitiu não apenas a construção de um dicionário com cerca de 800 entradas lexicais com informações multimídia, mas também a integração entre documentação

linguística, desenvolvimento tecnológico e devolução do conhecimento à comunidade, bem como o envolvimento de membros da comunidade no processo.

O processo de elaboração do Dicionário Multimídia Wayoro iniciou em 2023, no contexto da documentação sistemática da língua e da cultura Wayoro, através do projeto “*Documentation of the severely endangered languages Makurap and Wayoro (Brazil): material and non-material traditional culture, and its associated knowledge*”, apoiado pelo Endangered Languages Documentation Programme (Wajuru et al., 2025; Nogueira, Wajuru; Wajuru, 2023)¹⁷. Atualmente, o povo Wajuru habita duas regiões do estado de Rondônia: Porto Rolim de Moura do Guaporé (município de Alta Floresta d'Oeste) e a Terra Indígena Rio Guaporé (município de Guajará-Mirim). A população atual é de cerca de 340 indivíduos. A língua Wayoro não é mais falada no dia a dia. As crianças Wajuru se comunicam em português. Há apenas uma senhora de 77 anos que pode ser considerada falante fluente de Wayoro, além de cerca de vinte semi-falantes (Nogueira, 2019).

Atendendo à solicitação da comunidade Wajuru, foram elaborados, impressos e distribuídos um material de apoio à alfabetização na língua Wayoro (2017), um material de ensino da língua Wayoro (2015) e um vocabulário Wayoro-Português (Nogueira; Wajuru; Sousa, 2021). Além do dicionário multimídia Wajuru-Português, em 2022, uma equipe de pesquisadores indígenas e não-indígenas iniciou a elaboração de um livro de ensino da língua Wayoro, com o apoio do Museu dos Povos Indígenas/UNESCO.

A versão atual do dicionário multimídia Wayoro-Português (versão 1.0) está em processo de finalização e conta, atualmente, com 643 entradas lexicais, organizadas em 11 campos semânticos. A previsão de ser disponibilizada para validação e teste pela comunidade Wajuru é para o segundo semestre de 2026.

O trabalho de elaboração de um dicionário traz muitos desafios, que podem ser amplificados em situações de línguas severamente ameaçadas, como é o caso do Wayoro. Por exemplo, o processo de coleta lexical, especialmente em domínios especializados como fauna e flora, revela complexas relações entre o signo linguístico (a palavra na língua indígena), seu significado (a tradução) e o referente (a espécie ou objeto no mundo). A experiência com o dicionário Wayoro-Português ilustra bem esses desafios, conforme exemplificado a seguir:

- A entrada *nangoro* foi traduzida pela falante como ‘camaleão’, mas a imagem associada era a de um lagarto comum na região. No caso específico deste exemplo, a informação de que o termo ‘camaleão’ é o nome popular na região para essa e outras espécies de lagartos deve ser inserida no dicionário;
- A palavra *akui* foi inicialmente identificada a partir da imagem de uma barata-do-mato, mas a mesma colaboradora, em um contexto comunitário, corrigiu a tradução para ‘uma espécie de grilo’, evidenciando como mesmo estímulo visual pode levar a identificações equivocadas.

¹⁷ Os primeiros registros de áudio da língua Wayoro na Coleção Wayoro do ALIM foram realizados nos anos finais da década de 1990, por pesquisadores do Museu Paraense Emílio Goeldi. Em 2008, Antônia Fernanda de Souza Nogueira iniciou uma parceria com a etnia Wajuru da Terra Indígena Rio Guaporé, realizando documentação linguística, de maneira intermitente. Atualmente, como resultado do projeto de documentação, o Acervo Wayoro possui cerca de 148 horas de gravação audiovisual (elicitacoes, relatos pessoais e tradicionais).

Esses exemplos demonstram alguns dos desafios na elaboração de dicionários e como informações imprecisas, seja na tradução ou na associação com uma imagem, podem comprometer o objetivo central da transmissão linguística. Para evitar tais riscos, a documentação lexical exige uma metodologia rigorosa e, idealmente, uma colaboração interdisciplinar. Como sugerem Fleck (2007) e Berto (2010), a parceria com biólogos, o uso de guias de campo específicos da fauna e flora locais, estímulos auditivos, como o canto de aves e anfíbios, são importantes e até essenciais para garantir a correta identificação dos referentes.

O processo de elaboração do Dicionário Multimídia Makurap também iniciou em 2023, no contexto do projeto “*Documentation of the severely endangered languages Makurap and Wayoro (Brazil): material and non-material traditional culture, and its associated knowledge*”, mencionado anteriormente¹⁸. O povo Makurap é formado por uma população estimada em cerca de 500 pessoas, porém a sua língua tradicional, assim como as demais línguas abordadas neste artigo, o Makurap, encontra-se severamente ameaçada. O número de falantes está reduzido para cerca de 30 pessoas e a transmissão intergeracional da língua foi interrompida (Wajuru et al., 2025). Entretanto, um forte interesse da comunidade em aprender a língua Makurap foi despertado, em parte motivado por esse projeto colaborativo de documentação da língua e cultura Makurap (Wajuru et al., 2025). Nesse contexto, a comunidade Makurap tem feito atividades autônomas de revitalização da língua e da cultura tradicionais Makurap. Entre essas atividades podemos citar a realização de oficinas de fortalecimento cultural, curso de aprendizagem da língua e retomada de festas tradicionais. Um exemplo é a “Potkap Xako – Festa do Peixe” que teve sua terceira edição no contexto dessas retomadas, em setembro de 2025. A retomada dessa festa, que envolve diversos componentes culturais e reúne várias aldeias, visa valorizar a língua e a cultura Makurap, especialmente através das músicas cantadas na língua Makurap. Iniciativas como essa demonstram o empenho dos Makurap em participar de projetos que possam contribuir positivamente para a revitalização de sua língua de tradição e é nesse contexto que o dicionário multimídia Makurap está sendo produzido.

O dicionário multimídia Makurap está sendo elaborado integralmente a partir de registros específicos (áudios para as entradas lexicais e frases ilustrativas de uso dos verbetes) para este fim, envolvendo diversos falantes da língua, durante pesquisa de campo na Terra Indígena Rio Guaporé. Até o momento, o dicionário multimídia Makurap-Português é o mais amplo dos DMLIs desenvolvidos, contém, atualmente, 1.075 entradas lexicais e está em fase de ampliação. A macroestrutura e a microestrutura do dicionário seguem a organização apresentada no início desta seção e no Quadro 2, definidas de acordo com a orientação apresentada pela comunidade de fala e as especificidades da língua.

Questões específicas da língua Makurap, tais como casos de variação, lematização de verbos, homonímia, polissemia e classificação semântica preliminar dos itens lexicais de fauna e flora, demandaram estudos lexicais específicos e soluções analíticas influenciaram a composição dos

¹⁸ A coleção Makurap no ALIM vem sendo constituída desde a década de 1990, através de ações esporádicas de documentação pelos linguistas Denny Moore e Ana Vilacy Galucio. O projeto de documentação, financiado pelo Endangered Languages Documentation Programme/ELDP (Galucio; Nogueira; Costa, 2023) ampliou a coleção Makurap, para 160 horas de gravação audiovisual, com diferentes gêneros discursivos e elicitções, depositadas no ALIM e no ELAR.

verbetes e disposição das entradas lexicais. Esse estudo foi fundamental para estabelecer critérios específicos sobre como apresentar, no dicionário multimídia, as relações e nuances lexicais e gramaticais identificadas no trabalho lexicográfico de forma consistente e padronizada (Rocha; Galucio, no prelo). Uma das questões consideradas neste estudo foi a inclusão de palavras gramaticais (posposições, partículas, pronomes e conectivos) no dicionário, em que se buscou decidir como incluir nos verbetes as informações necessárias para que o usuário possa compreender a função gramatical dessas palavras na estrutura da língua utilizando os recursos opcionais da planilha CSV, a saber: descrição e itens relacionados, conforme parâmetros possíveis na metodologia definida por Brito, Birchall e Galucio (2025). O Quadro 3, abaixo, ilustra a apresentação de dois grupos de verbetes que incluem informações gramaticais.

tetyã	‘nós - 1ª pessoa plural exclusivo’	(não inclui o ouvinte, ou seja, quando se usa a forma tetyã, está se fazendo referência ao falante e outra(s) pessoa(s), mas não inclui o seu interlocutor: eu + outro(s))
kitetyã	‘nós - 1ª pessoa plural inclusivo’	(inclui o ouvinte, faz referência ao falante e seu interlocutor com a possibilidade de incluir ou não outras pessoas: eu + tu (+ outro(s)))
eya	‘futuro genérico’	(indica que o evento descrito pelo verbo da sentença ocorrerá no futuro, sem mais especificações)
ouga	‘futuro imediato’	(indica que o evento descrito pelo verbo ocorrerá em um futuro próximo ou imediato)

QUADRO 3 - Exemplos de lematização de palavras gramaticais no dicionário multimídia Makurap-Português.

Fonte: Rocha; Galucio. (no prelo)

Nesses casos, optou-se por registrar a tradução acompanhada de uma nota explicativa, de modo a evidenciar a diferença de significado e o uso de cada forma, como ilustrado no quadro acima, para possibilitar ao usuário do dicionário a compreensão da função/significado de cada palavra gramatical. Para isso, utilizamos as ferramentas de personalização de macro/microestrutura do dicionário multimídia, especialmente os campos de descrição (para criar verbetes com conteúdo explicativo sobre significado, função gramatical e/ou modo de uso da palavra) e itens relacionados (para relacionar palavras que podem ter um mesmo equivalente de tradução em português e criar um tipo de hipertexto, mas sem um link direto). Essas parametrizações visam criar mecanismos para tornar o material mais acessível e funcional para aprendizes de Makurap como segunda língua. A versão atual do dicionário Makurap (versão 1.0) está em processo de finalização, e está em fase de validação e teste pela comunidade de fala desde março de 2026.

3. Discussão: Dicionários multimídia e transmissão linguística

Esta seção discute o papel dos dicionários multimídias enquanto ferramenta didática no processo de transmissão linguística, incluindo casos específicos de retomada e/ou revitalização linguística. Ressalta-se sua importância como recurso pedagógico, sem presumir que a simples elaboração de um dicionário para uma língua ameaçada seja suficiente para documentar, preservar ou revitalizar tal idioma.

Do total de línguas indígenas faladas atualmente no Brasil, cerca de 40% a 50% está ameaçada e a outra parte em situação de vulnerabilidade. Nos últimos anos, muitos povos indígenas cujas línguas encontram-se em situação de vulnerabilidade, com poucos falantes, espaços de uso restritos e ruptura na transmissão intergeracional, iniciaram ou planejam iniciar movimentos autônomos de retomada, ou revitalização de suas línguas tradicionais. Nesse contexto, uma questão essencial é o planejamento linguístico, que inclui a avaliação e a descrição do status (situação), do corpus (material existente, documentado e descrito) e da aquisição (modo de transmissão linguística e produtos disponíveis) (Sallabank, 2012; McCarty, 2018; Amaral, 2020). Por exemplo, a preparação de material de apoio didático utilizado no processo de transmissão da língua é exemplo de planejamento da aquisição.

A revitalização linguística exige ações integradas de status, corpus e aquisição, orientadas pelo protagonismo das comunidades afetadas e apoiadas por políticas sensíveis ao contexto linguístico, histórico e sociocultural. Nestas etapas do planejamento, os dicionários multimídia surgem como uma ferramenta para auxiliar na revitalização de uma língua e podem ser úteis tanto no processo de transmissão da língua quanto no processo documental e descritivo, através da integração entre o trabalho dos linguistas e das comunidades indígenas para a documentação das línguas e produção de materiais acessíveis para quem usa essas línguas.

Os dicionários multimídia apresentados neste artigo foram planejados como ferramentas para contribuir com ações comunitárias de revitalização de línguas, aplicação de técnicas linguísticas e o uso de tecnologia. O desenvolvimento dessa tecnologia resulta de uma estreita colaboração entre linguistas e as comunidades indígenas para a formulação e elaboração de produtos que possam contribuir para o fortalecimento e revitalização de suas línguas tradicionais.

Ao combinar documentação de línguas indígenas e tecnologia digital de comunicação e informação, a Tecnologia Social dos Dicionários Multimídia envolve metodologias reaplicáveis de produtos digitais que podem contribuir para a aprendizagem e revitalização das línguas. O uso de tecnologia como smartphones, tablets ou computadores favorece a implementação de dicionários em formatos digitais, cujas entradas lexicais podem ser ilustradas com recursos multimídias. Assim, a prática da dicionarização e o formato multimídia, que pode ser usado *online* ou *offline*, contribui para aumentar a motivação dos falantes e potenciais falantes, inclusive crianças e jovens, uma vez que o uso do dicionário possibilita a ampliação do vocabulário da língua e a familiaridade com os sons e a pronúncia das palavras. Os oito dicionários multimídia apresentados neste artigo estão sendo elaborados com esse alinhamento, ou seja, os membros das comunidades estão envolvidos em todo o

processo de produção dos dicionários, desde o planejamento do corpus à validação do produto e ao uso da ferramenta nas ações em suas comunidades.

Essa tecnologia social é projetada para apoiar ativamente os processos de transmissão e aquisição linguística na comunidade. Esta é a função concreta dos dicionários. Além disso, os dicionários multimídia fomentam a promoção ou melhoria, de certa forma, da autoestima da comunidade e da percepção positiva dos falantes ou futuros aprendizes com relação a suas línguas. Ver suas línguas dentro dessa tecnologia, conforme temos experienciado, muitas vezes causa emoção nos usuários, especialmente quando a tecnologia permite trazer de volta as vozes de falantes que já não estão mais no meio da comunidade. Isto graças aos minuciosos trabalhos de documentação e à preservação de longa-duração dos acervos digitais, como o Acervo de Línguas Indígenas do Museu Goeldi (ALIM), que permitiu o resgate do material sonoro (também audiovisual) que foi utilizado para o desenvolvimento dos dicionários das línguas Kanoé, Puruborá, Djeoromitxi, Oro Win, Sakurabiat e Salamã.

A tecnologia e o trabalho incansável de documentar as línguas também permitiram colocar nesse tipo de mídia as vozes de falantes que ainda estão vivos na comunidade. Isto também gera outro tipo de motivação, trazendo à tona uma discussão muito importante no processo de revitalização, que é aproveitar a presença e disponibilidade dos falantes como uma fonte para a transmissão linguística, enquanto eles estão aqui. Concretamente, os usuários entendem que as ferramentas multimídia são fontes importantes de consultas para o aprendizado, mas que quando é possível consultar um falante e obter o contexto de uso daquele conhecimento linguístico direto de um avô, avó, tia etc., através da interação social, o aprendizado se torna mais eficaz.

Por outro lado, é fundamental reconhecer que a simples existência de um dicionário, por mais tecnologicamente avançado que seja, não garante por si só a preservação ou revitalização de uma língua. Para que cumpra seu papel de forma eficaz, seu desenvolvimento deve ser intencionalmente alinhado às três dimensões do planejamento linguístico: corpus, status e aquisição. No contexto das línguas abordadas, os dicionários nascem de um minucioso planejamento de **corpus**, que envolve a coleta, documentação e estruturação do léxico e contextos de uso. Simultaneamente, eles visam fortalecer o **status** da língua ao apresentá-la em um formato moderno, digital e acessível (via aplicativos em smartphones), o que aumenta sua visibilidade e prestígio, especialmente entre os jovens. O objetivo final, no entanto, é o planejamento da **aquisição**: servir como um instrumento prático que facilite o aprendizado e a transmissão intergeracional.

Para alcançar esse objetivo, a concepção do dicionário deve ir além da base de dados do linguista. É preciso um planejamento pedagógico que considere o público-alvo. Isso implica em decisões cruciais, como a seleção de campos lexicais relevantes para os aprendizes, a formulação de exemplos de uso claros e contextualmente ricos, e a incorporação estratégica de recursos multimídia (áudio, imagens, vídeos), como ilustrado na descrição dos dicionários apresentados na seção 2, acima. Esses elementos transformam o dicionário de uma lista de palavras e textos naturalísticos em uma experiência de imersão que apoia a aprendizagem da pronúncia, do significado e do uso culturalmente apropriado da língua.

Quatro dos dicionários apresentados neste artigo (Kanoé, Sakurabiat, Oro Win e Puruborá) já estão sendo utilizados pelas respectivas comunidades de fala, como ferramenta de auxílio para aprendizagem e revitalização das línguas. O dicionário da língua Kanoé foi distribuído para uso pelos membros da comunidade no ambiente familiar e residencial como uma ferramenta de consulta e aprendizagem, tanto na Terra Indígena Rio Guaporé quanto na Terra Indígena Rio Branco. Está sendo discutida a possibilidade de ampliação do dicionário, a partir de dados da língua armazenados no arquivo do Museu Nacional dos Povos Indígenas/FUNAI, no Rio de Janeiro, e também do registro com os falantes da língua na Terra Indígena Rio Omerê. No caso do dicionário Sakurabiat, além do uso do aplicativo pelos membros da comunidade no ambiente familiar e residencial, também está sendo usado como ferramenta de aprendizagem nas duas escolas atuais da Terra Indígena Rio Mequéns. O dicionário multimídia da língua Puruborá vem sendo incluído gradativamente nas atividades relacionadas à revitalização da língua. Inicialmente, o dicionário havia sido distribuído para os membros da comunidade, os quais vinham utilizando este produto no ambiente familiar e residencial como uma ferramenta de consulta. Mais recentemente, está ocorrendo um planejamento para um uso mais sistemático desta ferramenta no ambiente escolar como suporte ao processo de ensino-aprendizado da língua e alinhada a outros materiais didáticos da língua. O uso do dicionário Oro Win já foi implementado na sala de aula de três escolas da comunidade Oro Win desde 2022. Também foram produzidos quatro cartazes a partir do dicionário para auxiliar os professores indígenas no ensino da língua, abrangendo temas como o uso do alfabeto e os nomes de vários tipos de plantas e animais.

Conforme mencionado anteriormente, os resultados diretos da aplicação e do uso das ferramentas e procedimentos metodológicos para o desenvolvimento dos dicionários multimídia para línguas indígenas (DMLI) descritos neste artigo podem ser conferidos no site de dicionários do Museu Paraense Emílio Goeldi (<https://dicionarios.museu-goeldi.br/>), criado para divulgar os produtos gerados a partir da aplicação da metodologia junto às comunidades estudadas e ao público em geral. O site contém atualmente seis dicionários publicados. Nele, é possível visualizar o conteúdo online ou baixar versões de cada um desses produtos em HTML, PDF ou para Android (APK), quando disponíveis.

4. Considerações finais

No cenário atual no Brasil de forte ameaça à diversidade linguística e considerando a motivação e o interesse de vários povos indígenas pela retomada, revitalização e fortalecimento das suas respectivas línguas tradicionais, os dicionários multimídia se consolidam como uma poderosa ponte entre a documentação linguística e a revitalização de base comunitária.

Como apresentado nas seções anteriores, a tecnologia e a metodologia para elaboração dos DMLI apresentadas neste artigo utilizam softwares de uso aberto, reaplicável, de uso livre e fácil acesso e pode ser usada pelas comunidades em larga escala e sem necessidade de acesso à internet, podendo também ser aplicada a qualquer língua. Sua utilização pelas comunidades cujos dicionários

estão em uso demonstram que esses recursos multimídia podem funcionar como ferramenta didática no processo de transmissão linguística e que, nesse caso, não podem funcionar como meros repositórios lexicais, mas podem ir além, cumprindo um papel de ferramenta (com outros materiais de apoio didático) nas ações de revitalização. Nesse sentido é que a abordagem metodológica e a aplicação da tecnologia na produção dos dicionários são concebidas como uma tecnologia social e como uma ferramenta pedagógica dinâmica que pode ser atualizada constantemente devido à abordagem colaborativa entre comunidade, linguistas e profissionais da área de tecnologia. Entendemos que documentação, descrição e análise linguística, bem como a elaboração de materiais de ensino de língua constituem um fluxo de trabalho cíclico retroalimentar, colaborativo, permanente e dinâmico (Nogueira; Massaka; Wajuru, 2023).

Considerando que intervenções de fortalecimento ou retomada linguística devem partir do desejo e objetivos das comunidades, que podem ser apoiadas por agentes institucionais quanto ao uso de tecnologia e metodologias sistematizadas e replicáveis, para facilitar a transmissão, o sucesso dos dicionários multimídia não reside apenas na tecnologia empregada, mas na integração cuidadosa entre o planejamento de corpus, status e aquisição, e em um compromisso contínuo com a precisão e a relevância cultural do conteúdo, garantindo que ele seja uma ferramenta confiável e eficaz para os atuais e futuros falantes.

Informações complementares

Avaliação e resposta dos autores

Avaliação: <https://doi.org/10.25189/rabralin.v24i2.2384.R>

Editora

Maria Filomena Spatti Sandalo

Afiliação: Universidade Estadual de Campinas

ORCID: <https://orcid.org/0000-0003-4595-7765>

RODADAS DE AVALIAÇÃO

Avaliador 1: Kristina Balykova

Afiliação: University of Texas at Austin

ORCID: <https://orcid.org/0000-0001-6194-8821>

Avaliador 2: Maria Filomena Spatti Sandalo

Afiliação: Universidade Estadual de Campinas

ORCID: <https://orcid.org/0000-0003-4595-7765>

AVALIADOR 1

O artigo é altamente relevante para o dossiê, pois discute uma ferramenta (dicionários multimídia) que serve como a ponte entre os resultados de documentação linguística e as necessidades de retomada ou fortalecimento de línguas ameaçadas. O artigo cumpre o propósito de divulgar essa ferramenta, oferecendo um relatório sobre os produtos finalizados / em construção e argumentando a favor de ampliar o uso dessa tecnologia. No entanto, o artigo tem um desafio estrutural, que é unir a discussão geral com os 8 "estudos de caso", que são apresentados um tanto isolados um do outro. Nesse sentido, a estrutura do artigo poderia ser melhorada se algumas informações e considerações comuns a todos os dicionários (e que muitas vezes aparecem repetidas em cada subseção) fossem reunidas a fim de proporcionar uma visão mais geral sobre os dicionários. Por favor, veja minhas sugestões no arquivo anexado.

AVALIADOR 2

Texto muito atual sobre a necessidade de elaboração de dicionários que alimentem a pesquisa e educação em línguas originárias do Brasil.

Conflito de Interesse

Os autores não têm conflitos de interesse a declarar.

Link para *Preprint*

<https://doi.org/10.1590/SciELOPreprints.13693>

Protocolo e Pré-Registro de Pesquisa

Avaliando os roteiros propostos pela Equator Network, consideramos que nenhum deles se mostra relevante para a pesquisa em tela. Também informamos que a pesquisa desenvolvida não foi pré-registrada em repositório institucional independente.

Declaração de Disponibilidade de Dados

O compartilhamento de dados não é aplicável a este artigo, pois nenhum dado novo foi criado ou analisado neste estudo.

Pesquisa com seres humanos

As pesquisas reportadas neste artigo contaram com o consentimento informado de todos os participantes. Os projetos de pesquisa específicos que incluíram parte das pesquisas reportadas neste artigo foram aprovados pela Comissão Nacional de Ética em Pesquisa, sob os números CAAE 63101422.1.0000.0173 e CAAE 62625822.8.0000.0173.

Agradecimentos

Os autores agradecem às comunidades indígenas envolvidas nos projetos de produção de dicionários multimídia pela parceria e colaboração. Agradecem também às seguintes instituições que apoiaram as pesquisas que possibilitaram o desenvolvimento do presente artigo: National Endowment for the Humanities (EUA), programa Documenting Endangered Languages (projeto FN-266285-19 Documentation and Dictionary of Oro Win (orw), coordenado por Joshua Birchall). Endangered Languages Documentation Programme (projetos MDP0435 Documentation of the severely endangered languages Makurap and Wayoro (Brazil): material and non-material traditional culture, and its associated knowledge, coordenado por Ana Vilacy Galucio e Antonia Fernanda Nogueira; e projeto IPF0386 Documentation and description of the Djeoromitxi language from Amazonia, Brazil, coordenado por Ivan Rocha da Silva). Conselho Nacional de Pesquisa e Desenvolvimento/CNPq: Processo 445415/2024-0; Bolsa de produtividade em pesquisa (PQ)- 304766/2022-4 para Ana Vilacy Galucio; bolsa de capacitação institucional (PCI) no MPEG/MCTI-314025/2025-1 para Saulo Brito; bolsa de Iniciação Científica no CNPq/MPEG para Douglas Rodrigues Junior, Israel Marcon, Flávio Pinheiro, Juliana Solano, Matheus Soares e Victor Rocha. Fundação Nacional dos Povos Indígenas/FUNAI pelo apoio logístico durante as pesquisas de campo e autorizações de ingresso em área indígena, processos 08620.002095/2009-57, 08620.010349/2022-88 e 08620.009160/2022-42.

Fontes de financiamento

National Endowment for the Humanities (EUA), programa Documenting Endangered Languages; Endangered Languages Documentation Programme; Conselho Nacional de Pesquisa e Desenvolvimento/CNPq

REFERÊNCIAS

- AMARAL, Luiz. Estratégias para a revitalização de línguas ameaçadas e a realidade brasileira. **Cadernos de Linguística**, [S. l.], v. 1, n. 3, p. 01-44, 2020. DOI: 10.25189/2675-4916.2020.v1.n3.id251. Disponível em: <https://cadernos.abralin.org/index.php/cadernos/article/view/251>. Acesso em: 8 out. 2025.
- AQUINO, Letícia de Souza; KANOE, M. **Compartilhando dados e experiências de Revitalização Linguística com a comunidade Kanoé do Guaporé**. 2024. (Apresentação de Trabalho/Comunicação).
- BERTO, Flávia de Freitas. **Nomes de aves em Jurunas: um estudo lexicológico**. Monografia (Bacharel em Letras) - Universidade Estadual Paulista "Júlio de Mesquita Filho", Campus de Araraquara. 2010. Disponível em: <https://repositorio.unesp.br/server/api/core/bitstreams/0344b232-4211-407d-a7e6-df145259567e/content>. Acesso em: 30 sep. 2025.
- BIRCHALL, Joshua. Lexicografía y elaboración de diccionarios. In: P. ARANDIA (org.) **Introducción a la Lingüística: Curso para investigadores de lenguas indígenas de Bolivia**, Cochabamba: Página y Signos, 2023. p. 141-159.
- BIRCHALL, Joshua. **Oro Win**. Inventário Nacional da Diversidade Linguística. Relatório técnico apresentado ao IPHAN/Museu Paraense Emílio Goeldi, 2018. (manuscrito).
- BIRCHALL, Joshua; ORO EO' CABIXI, Silvana; ORO WIN, Luciano. **Dicionário Oro Win**. Versão 3.0. 2025. Disponível em: https://dicionarios.museu-goeldi.br/dic/oro_win/. Acessado em: 07/10/2025
- BIRCHALL, Joshua; ORO WIN CABIXI, Olivia. Passos na retomada da língua do povo Oro Win. In: TONDINELI, Patrícia Goulart. (ed.), **(Re)vitalizar línguas minorizadas e/ou ameaçadas: teorias, metodologias, pesquisas e experiências**, Porto Velho: EDUFRO, 2021. p. 69-82.
- BRITO, Saulo; BIRCHALL, Joshua; GALUCIO, Ana Vilacy. **Csv2rmd**: Um programa python para produzir dicionário multimídia com Markdown. Versão 1.0. 2025. DOI: <https://doi.org/10.5281/zenodo.14894549>.
- COSTA, Carla Nascimento da. **Material Didático Sakurabiat**. (Desenvolvimento de material didático ou instrucional - Material Didático). MPEG: Belém. 2020a.
- COSTA, Carla Nascimento da. **Proposta de Material Didático para a Língua Sakurabiat**. 160 f. Dissertação (Mestrado) - Programa de Pós-graduação em Letras, Universidade Federal do Pará, Belém, 2020b.
- ELAN (Versão 7.0) [Software de computador]. Nijmegen: Instituto Max Planck de Psicolinguística, **The Language Archive**, 2025. Disponível em: <https://archive.mpi.nl/tla/elan>. Acesso em: 07 out. 2025
- FELZKE, Lediane Fani; MOORE, Denny. Kinship terminology of the groups of the Mondé language family. **Boletim do Museu Paraense Emílio Goeldi**, Ciências Humanas, v. 14, n. 1, p. 15-32, 2019. DOI: 10.1590/1981.81222019000100003.
- FLECK, David. Field linguistics meets biology: how to obtain scientific designations for plant and animal names. **STUF - Language Typology and Universals**, Berlin, vol. 60, no. 1, 2007, pp. 81-91. <https://doi.org/10.1524/stuf.2007.60.1.81>
- GALUCIO, Ana Vilacy. Documentação e Revitalização Linguística: uma interseção possível, necessária e desejável. In: Patrícia Goulart Tondineli. (Org). **(Re)vitalizar línguas minorizadas e/ou ameaçadas: teorias, metodologias, pesquisas e experiências**. 1ed. Porto Velho: EDUFRO, 2021, v.1, p. 20 - 43.

GALUCIO, Ana Vilacy. Heritage languages, linguistic obsolescence and motivations for learning in the context of indigenous school education: reflections, doubts, and challenges. **Cadernos de Linguística**, [S. l.], v. 1, n. 3, p. 01-20, 2020. DOI: 10.25189/2675-4916.2020.v1.n3.id235. Disponível em:

<https://cadernos.abralin.org/index.php/cadernos/article/view/235>. Acesso em: 30 sep. 2025.

GALUCIO, Ana Vilacy. Puruborá: notas etnográficas e linguísticas recentes. *Boletim do Museu Paraense Emílio Goeldi. Ciências Humanas*, Belém, v. 1, n.2, p. 159-192, 2005. 2005

GALUCIO, Ana Vilacy; MEIRA, Sérgio; BIRCHALL, Joshua; MOORE, Denny; GABAS JÚNIOR, Nilson; DRUDE, Sebastian; STORTO, Luciana; PICANÇO, Gessiane; RODRIGUES, Carmen Reis. Genealogical relations and lexical distances within the Tupian linguistic family. **Boletim do Museu Paraense Emílio Goeldi, Ciências Humanas**, v. 10, 2015. Disponível em: <https://doi.org/10.1590/1981-81222015000200004>. Acesso em: 30 sep. 2025

GALUCIO, Ana Vilacy; NOGUEIRA, Antonia Fernanda; COSTA, Carla Daniele da. **Makurap: Documentation of Language and Culture**, 2023. Endangered Languages Archive. Handle: <http://hdl.handle.net/2196/73b9o01s-1c78-2246-5187-0fm978700a80>. Acesso em: 07 out. 2025.

GALUCIO, Ana Vilacy; RODRIGUES JUNIOR, Douglas; BRITO, Saulo (com Manoel Sakyrabiar, Olimpio Sakyrabiar, Rosalina Guaratira, Severino Guaratira e Matheus A. R. Soares). 2024. **Dicionário multimídia Sakurabiat-Português**. Versão 3.0. Belém: Museu Paraense Emílio Goeldi. Disponível em <http://dicionarios.museu-goeldi.br/dic/sakurabiat/>. Acesso em 28 set. 2025.

GUARATIRA, Silvana da Silva Cunha; COSTA, Carla Nascimento da. Experiences of Sakurabiat language and culture rescue. **Cadernos de Linguística**, [S. l.], v. 1, n. 3, p. 01-14, 2020. DOI: 10.25189/2675-4916.2020.v1.n3.id249. Disponível em: <https://cadernos.abralin.org/index.php/cadernos/article/view/249>

HANKE, Wanda. Breves notas sobre os índios Mondé e o seu idioma. **Dusenía** (Curitiba), vol. I, n. 4, 1950, pp. 215-228.

MARCON, Israel.; GALUCIO, Ana Vilacy; Brito, Saulo. 2023. **Dicionário multimídia Kanoe-Português**. Versão 2.0. Belém: Museu Paraense Emílio Goeldi. Disponível em <http://dicionarios.museu-goeldi.br/dic/kanoe/>. Acesso em 09 out. 2025.

MCCARTY, Teresa. Community-based language planning: Perspectives from indigenous language revitalization. In : HINTON, L.; HUSS, L. M.; ROCHE, G. (Ed.). **The Routledge handbook of language revitalization**. New York and London: Routledge, 2018, p. 22 -35.

MOORE, Denny. **Djeoromitxi**. Inventário Nacional da Diversidade Linguística. Relatório técnico apresentado ao IPHAN/Museu Paraense Emílio Goeld, 2018. (manuscrito).

NOGUEIRA, Antônia Fernanda de Souza. **Predicação na língua Wayoro (Tupi)**: propriedades de finitude. 189 f. Tese (Doutorado em Semiótica e Linguística Geral) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2019. Disponível em: <https://www.teses.usp.br/teses/disponiveis/8/8139/tde-14082019-101006/en.php>. Acesso em: 25 sep. 2025.

NOGUEIRA, Antonia Fernanda de Souza; MASSAKA, Gilson Wajuru; WAJURU, Adão. 'We Didn't Wear Skirts': Extralinguistic Factors in the Development of Wayoro Pedagogical Materials. In: VALIJÄRVI, Riitta-Liisa; KAHN, Lily (Eds.). **Teaching and Learning Resources for Endangered Languages**. Leiden: Brill, 2023. p. 190 - 202.

NOGUEIRA, Antonia Fernanda de Souza; WAJURU, Comunidade; SOUSA, Clenilson M. Apresentação da primeira versão do vocabulário Wayoro-Português In: TONDINELI, Patrícia Goulart (Org.). **(Re)vitalizar línguas minorizadas e/ou ameaçadas: teorias, metodologias, pesquisas e experiências**. Porto Velho: EDUFRO, 2021. p. 83 - 108.

OLIVEIRA NETO, Mario. Revitalization and teaching of the Puruborá language. **Cadernos de Linguística**, [S. l.], v. 1, n. 3, p. 01-16, 2020. DOI: 10.25189/2675-4916.2020.v1.n3.id246. Disponível em: <https://cadernos.abralin.org/index.php/cadernos/article/view/246>. Acesso em: 30 sep. 2025.

PINTO SILVA, Bruno; BIRCHALL, Joshua. **Dicionário multimídia Salamã-Português**. Versão 1.0. 2024. Disponível em <http://dicionarios.museu-goeldi.br/dic/salamai/>. Acesso em: 07 out 2025.

PIRES, Nádia Nascimento. **Estudo da gramática da língua Jeoromitxi (Jabutí)**: Aspectos sintáticos das cláusulas matrizes. 157 f. Dissertação (Mestrado em Linguística), Universidade Estadual de Campinas, Campinas, 1992.

ROCHA, Ivan. Documentation and description of the Djeoromitxi language from Amazonia, Brazil. **Endangered Languages Archive**, 2022. Disponível em: <http://hdl.handle.net/2196/2ef60f9c-796c-4b6f-c50c-417343074e7g>. Acesso em: 02 out. 2025.

ROCHA, Victor Siqueira; GALUCIO, Ana Vilacy. Dicionário Multimídia para a língua Makurap: Uma Ferramenta para Aprendizagem da Língua. **Letras Escreve**. No prelo.

SALLABANK, Julia. From language documentation to language planning: Not necessarily a direct route. In : SEIFART, F. et al. (ed.). Potentials of language documentation: Methods, analyses, and utilization. Honolulu: University of Hawaii Press, **Language Documentation & Conservation**, v. 3. p. 118 – 125, 2012.

SOARES, Matheus Augusto Ribeiro.; GALUCIO, Ana Vilacy; OLIVEIRA NETO, Mario (com Paulo Aporete Filho, José Evangelista Puruborá e Saulo Brito). 2024. **Dicionário multimídia Puruborá-Português**. Versão 3.0. Belém: Museu Paraense Emílio Goeldi. Disponível em <http://dicionarios.museu-goeldi.br/dic/purubora/>. Acessado em 28/09/2025.

SOARES, Matheus Augusto Ribeiro; RODRIGUES JUNIOR, Douglas da Costa; GALUCIO, Ana Vilacy. Dictionary Making as a Tool for the Revitalization of Endangered Languages: Two Case Studies. **Cadernos de Linguística**, [S. l.], v. 5, n. 1, p. e704, 2024. DOI: 10.25189/2675-4916.2024.v5.n1.id704. Disponível em: <https://cadernos.abralin.org/index.php/cadernos/article/view/704>. Acesso em: 1 oct. 2025

WAJURU, Jacqueline; MACURAP, Jociclei; NOGUEIRA, Antonia Fernanda; GALUCIO, Ana Vilacy; COSTA, Carla Nascimento. Engaging Wajuru /Wayoro and Makurap Communities in Collaborative Documentation: Recording, Learning, and Communicating. **Dutkansearvvi diedalaš áigečála**, v. 09, p. 54-74, 2025.