Revista Letras Escreve Macapá, AP v. 16 – n. 3 ISSN 2238-8060
DOSSIÊ TEMÁTICO - Línguas ameaçadas: Documentação, descrição, análise, revitalização e ensino
ARTIGO

DICIONÁRIO MULTIMÍDIA DA LÍNGUA MAKURAP:

Uma Ferramenta para aprendizagem da língua

MULTIMEDIA DICTIONARY OF THE MAKURAP LANGUAGE:

A tool for language learning

Revista dos Cursos de Letras e do Programa de Pós-graduação em Letras – Unifap

FLUXO DA SUBMISSÃO
Submissão do trabalho: 23/07/2025
Aprovação do trabalho: 2/09/2026
Publicação do trabalho: 20/11/2026

Victor Siqueira Rocha
victorsrocha123hdbe@gmail.com
Museu Paraense Emílio Goeldi

Ana Vilacy Galucio
avilacy@museu-goeldi.br
Museu Paraense Emílio Goeldi

COMO CITAR
ROCHA, Victor; GALUCIO, Ana Vilacy. DICIONÁRIO MULTIMÍDIA DA LÍNGUA MAKURAP: Uma Ferramenta para aprendizagem da língua. Revista Letras Escreve, [S. l.], v. 16, n. 3, p. 106–123, 2026. Disponível em: https://periodicos.unifap.br/letrasescreve/article/view/1836.

Resumo

Este trabalho apresenta os aspectos teóricos e práticos da produção de um Dicionário Multimídia Makurap-Português, baseado em um método reaplicável concebido na área de linguística do Museu Paraense Emílio Goeldi (MPEG). O artigo descreve a macroestrutura e a microestrutura do dicionário e discute as decisões tomadas para sua produção, especialmente no que se refere aos processos de lematização de palavras homônimas, polissêmicas e verbos; inserção de palavras gramaticais e classificação semântica dos itens lexicais de fauna e flora. Discorre-se ainda sobre o processo de elaboração do Dicionário Multimídia Makurap - Português não só como aplicação de um modelo reaplicável, mas como uma tecnologia social importante para o fortalecimento da língua na comunidade.

Palavras-chave: lexicografia; línguas indígenas; estudo lexical; revitalização; tecnologia social.

Abstract:

This paper presents the theoretical and practical aspects of producing a Makurap-Portuguese Multimedia Dictionary, based on a replicable method conceived in the linguistics division of the Museu Paraense Emílio Goeldi (MPEG). The article describes the macrostructure and microstructure of the dictionary and discusses several aspects of its production, especially regarding the processes of lemmatization of verbs and homonymous and polysemous words; insertion of grammatical words and semantic classification of lexical items of fauna and flora. The article also discusses the process of creating the Makurap - Portuguese Multimedia Dictionary not only as an application of a replicable model, but also as an important social technology for strengthening the use of the language in the community.

Keywords: lexicography; indigenous languages; lexical study; revitalization; social technology.

Introdução

A elaboração de dicionários, de modo geral, envolve a compilação de palavras e significados de uma língua. Biderman (2001a) argumenta que o dicionário faz a descrição do vocabulário de uma língua, buscando registrar e definir os signos lexicais que referem a conceitos elaborados e cristalizados na cultura, tornando-se um objeto cultural de suma importância nas sociedades contemporâneas. Segundo Borba (2003), ao apresentar o léxico, os dicionários viabilizam o conhecimento e a difusão da língua entre a comunidade, e por isso não se deve tomá-lo apenas com um mero repositório de palavras, mas como um material pedagógico de primeira linha.

Os dicionários podem ser usados pelas comunidades de fala como um instrumento didático-pedagógico, de acordo com suas necessidades. Para Krieger (2007 *apud* Fargetti e Mondini, 2021, p. 26), todo dicionário é potencialmente didático, justamente pela possibilidade de ser uma ferramenta fundamental no processo de ensino e aprendizagem de uma língua, mas nem todo dicionário é pedagógico.

Para além do caráter pedagógico, a produção lexicográfica pode ser percebida como um registro do patrimônio sociocultural de uma comunidade:

> "É relativamente fácil antever que um dicionário, em geral, supera a singela função de "tira-dúvidas". Afinal, é também um tipo de repositório ou de registro de todo um patrimônio sociocultural configurado pela língua [...]. Mas, com certeza, esse "a mais" contido em um dicionário só será descoberto pelo consulente da obra, independentemente do seu tipo ou natureza, à medida que ele saiba reconhecer as diferentes nuances de informação oferecidas em um verbete." (Bevilacqua; Finatto, 2006, p. 3)

Seja para servir como um instrumento didático-pedagógico ou para representar uma parte da dimensão sociocultural, buscamos ao longo do projeto de desenvolvimento do Dicionário Multimídia Makurap-Português criar estratégias para apresentar o léxico da língua Makurap de forma mais adequada aos objetivos desse projeto, desenvolvido no âmbito de amplo projeto de elaboração de dicionários multimídia vinculado ao Museu Goeldi (cf. Galucio et al, 2025).

Questões como classificação semântica de itens de lexicais, variação e inserção de palavras gramaticais exigiram uma atenção especial para definir um modo de apresentação padronizado no dicionário. De modo geral, a solução encontrada refere-se à organização estrutural das entradas lexicais e à criação de verbetes com textos descritivos.

As soluções estruturais são possíveis graças às possibilidades de personalização incluídas na atualização da metodologia empregada por Brito, Birchal e Galucio (2025)[1], pois essas possibilidades

[1] Para uma maior compreensão das etapas metodológicas de elaboração de dicionários multimídia no formato utilizado, indicamos também o manual disponível em: GitHub - MPEG-Linguistica/CSV2RMD: Script Python que permite converter arquivos de modelo específico de CSV em dicionários multimídia.

não estavam presentes nas versões beta dos primeiros dicionários multimídia desenvolvidos no Museu Goeldi (cf. Soares; Rodrigues Junior: Galucio, 2024). A inclusão de verbetes com textos descritivos busca oferecer as informações necessárias para compreensão do significado e do funcionamento de uma palavra. Esses textos são determinantes para que um dicionário supere a função de "tira-dúvidas" e contenha em si uma parte de um patrimônio sociocultural configurado pela língua, como propõem Bevilacqua e Finatto (2006).

Nesse contexto, o objetivo central do presente artigo é apresentar os aspectos teóricos e práticos da elaboração de um dicionário multimídia para a língua Makurap. Buscaremos descrever as decisões editoriais pertinentes à organização da macroestrutura e da microestrutura do dicionário, bem como suas motivações.

Ao apresentar o processo de elaboração do dicionário multimídia Makurap-Português, procuramos destacar como esse material pode oferecer aos seus usuários as informações necessárias para compreensão de um item lexical ou gramatical da língua, especialmente no contexto de aprendizagem do Makurap como segunda língua.

Na próxima seção, discutiremos brevemente os aspectos sociolinguísticos do povo Makurap e o projeto de documentação linguística em curso.

1 A língua Makurap

A língua Makurap, uma das cinco línguas integrantes do ramo Tupari da família Tupi, é falada pelo povo homônimo que, em sua maioria, reside atualmente nas Terras Indígenas Rio Branco e Rio Guaporé no estado de Rondônia, com uma população estimada em cerca de 500 pessoas.

O número de falantes fluentes da língua varia entre 24 e 30 falantes fluentes (Wajuru et al., 2025). Desses, 83% são adultos com mais de cinquenta anos, conforme dados levantados em campo no âmbito de projeto de documentação da língua e cultura Makurap[2], coordenado por Ana Vilacy Galucio, com apoio do Programa de Documentação de Línguas Ameaçadas[3] (Wajuru et al., 2025). Essa distribuição etária ressalta uma crítica interrupção na transmissão intergeracional da língua. Cientes dessa situação, vários membros da comunidade de fala demonstram interesse em aprender sua língua de tradição, ao mesmo tempo em que colaboram com projetos que visam à documentação da língua e cultura Makurap (Wajuru et al., 2025).

Wajuru e colegas (2025) discutem o projeto de documentação da língua Makurap em curso. A iniciativa é uma ação colaborativa com a comunidade de fala, que adota como estratégia a combinação de treinamento técnico, capacitação e uma abordagem comunitária da concepção à execução do projeto, com o objetivo de obter como resultados coleção de gravações, um corpus

[2] Projeto registrado na Plataforma Brasil, Comissão Nacional de Ética em Pesquisa, e aprovado sob o número CAAE 63101422.1.0000.0173.

[3] O *Endangered Language Documentation Program* (ELDP) é um programa que tem como objetivo dar suporte globalmente a projetos que visam à documentação e preservação de línguas ameaçadas. Ver em Endangered Languages Documentation Programme.

traduzido e transcrito, um banco de dados lexical e um dicionário multimídia. A criação de um dicionário multimídia atende a uma demanda direta da comunidade Makurap por ferramentas adequadas para fomentar a aprendizagem da língua, como relatado por Maisa Macurap, professora de língua e cultura Makurap na Aldeia Ricardo Franco, em entrevista realizada em março de 2026:

> "Ele (Dicionário Multimídia Makurap-Português) vai ajudar muito as nossas crianças a aprender e a falar na língua materna, e com a ajuda desse material que vem com desenhos e sons vai ser muito bom, assim eles vão ver o desenho e a escrita e ouvir o som. Então, isso vai ajudar bastante na valorização da nossa língua materna"

É importante destacar que grande parte das decisões editoriais sobre o dicionário multimídia realizaram-se a partir de contribuições da comunidade Makurap, porque partimos do princípio defendido por Bichall (2023) de que o trabalho de elaboração de dicionários deve ser realizado em estreita colaboração com a própria comunidade, atender à sua situação sociolinguística específica e às suas necessidades educacionais.

Segundo Gonçalves (2025), o dicionário produzido para a língua Medzeniakonai (Baniwa-Koripako) tornou-se um material didático-pedagógico valioso, especialmente por oferecer os recursos multimídia que consequentemente permitem que a língua possa ser utilizada em novos espaços sociais, inclusive na escola, no uso de aparelhos digitais e na internet. A possibilidade de acesso à língua em um novo espaço social através da tecnologia como benefício para as pessoas é perceptível na fala de Jefferson Macurap, da aldeia Ricardo Franco, uma liderança política jovem do povo Makurap, ao comentar sobre o dicionário multimídia em entrevista realizada em março de 2026:

> "A juventude indígena tem muito contato com a internet e com a tecnologia física, e essa tecnologia vem pra ajudar a recuperar a nossa cultura (...) nossa geração mais jovem de poder tá utilizando e praticando nossa fala (..). Isso ajuda muito porque eu tava viajando e deu tempo de eu pegar o aplicativo (do Dicionário Multimídia Makurap-Português) (...) e tirar minha dúvida, porque não é todo tempo que a gente tá com a titia ou com a vovó falando, mas o aplicativo já fala e ensina a gente, é só apertar aqui e a gente já escuta a voz da vovó (...) a gente pode estar em outros espaços, mas o dicionário multimídia aproxima a gente da língua"

Dessa forma, considerando a disponibilidade de registros em áudio e vídeo em diferentes contextos de uso da língua Makurap e o interesse da comunidade em aprendê-la, um dicionário multimídia, contendo áudio e imagem, foi considerado mais adequado.

A seguir, apresentaremos o amplo projeto de desenvolvimento de dicionários multimídia que o Dicionário Multimídia Makurap-Português está situado.

2 O projeto de elaboração de dicionário multimídia para línguas indígenas

A produção do Dicionário Multimídia Makurap-Português, utiliza um método reaplicável, concebido por Brito, Birchall e Galucio (2025). Esse método é resultado da cooperação entre linguistas do Museu Paraense Emílio Goeldi (MPEG) e da *University of New Mexico* (UNM, EUA) para criação de produtos lexicográficos que possam funcionar como ferramenta de referência e também como instrumento de preservação e revitalização linguística, a partir da demanda direta de comunidade por ferramentas para fomentar o ensino de línguas, que em uma sua grande maioria estão severamente ameaçadas (cf. Galucio et al, 2025).

A definição de dicionário multimídia é apresentada por Galucio *et al.* (2025) como "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" (p. 5), cujo objetivo é fomentar a aprendizagem e o uso das línguas através de ferramentas digitais acessíveis e construídas de forma colaborativa com as comunidades indígenas.

Essa perspectiva de dicionário multimídia contempla os aspectos tecnológico e social por servir como uma tecnologia social voltada à documentação, preservação e revitalização de línguas indígenas ameaçadas em um processo de colaboração ativa com a comunidade. Fator importante a ser considerado na realização de trabalhos lexicográficos com línguas ameaçadas, conforme propõe Kroskrity (2015).

De modo geral, a metodologia consiste em compilar dados linguísticos e arquivos de mídia em um formato que possibilita um fácil compartilhamento e acesso. Essa metodologia inclui informações básicas sobre cada entrada lexical (lema e frase ilustrativa de uso na língua em questão, com tradução para o português), valendo-se de recursos de imagem para ilustração e de áudio para exemplificar a pronúncia. A adoção desse formato multimídia para a produção do dicionário da língua Makurap justifica-se pela acessibilidade proporcionada pela tecnologia, que possibilita a utilização do dicionário no celular e no computador, *online* e *offline,* e também permite à comunidade utilizá-lo do modo que desejar, inclusive de forma impressa no formato PDF.

Discutiremos a seguir a metodologia que norteou as escolhas no processo de elaboração do dicionário multimídia Makurap-Português.

3 Metodologia

A metodologia empregada subdivide-se em três grandes etapas que ocorrem sucessivamente: a) coleta de dados; b) tratamento de dados e c) validação final.

Para a coleta de dados, contamos com a colaboração de 07 pesquisadores indígenas e falantes da língua Makurap em trabalhos de campo realizados na Terra Indígena Rio Guaporé entre 2023 e 2026. A coleta foi realizada especificamente com a finalidade de elaborar o dicionário multimídia e inclui gravações de elicitação lexical para 14 campos semânticos, a saber: Animais, Atividades e Eventos, Comidas e Bebidas, Cores, Estados e Propriedades, Itens Gramaticais, Manufaturas e Itens

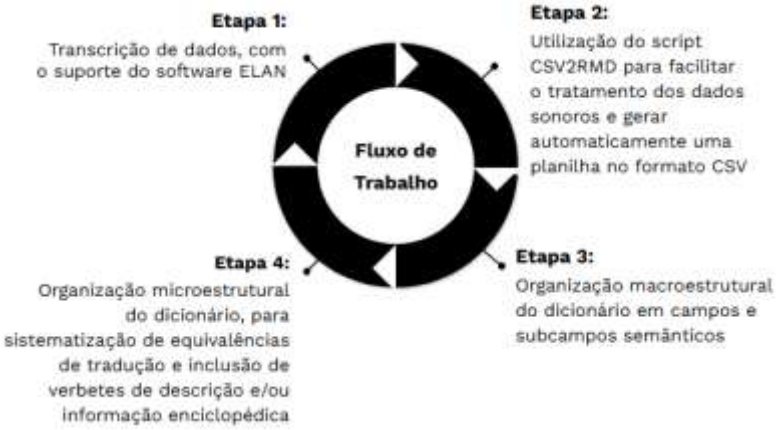
do Dia a Dia, Natureza, Números, Outros, Partes do Corpo, Pessoa, Parentesco e Relações Familiares, Plantas e Fungos e Secreções Corporais.

O registro dos dados linguísticos (entrada lexical, sentenças ilustrativas de uso de cada entrada e suas respectivas traduções para português) foi feito por campo semântico, envolvendo preferencialmente uma equipe de dois ou três pesquisadores indígenas e falantes de Makurap e um ou dois pesquisadores não-indígenas. Em alguns casos, a coleta foi feita individualmente com um pesquisador não-indígena e um pesquisador indígena falante da língua, e as informações foram posteriormente validadas com outros falantes, seguindo solicitação dos próprios pesquisadores indígenas.

Como mencionado anteriormente, para o desenvolvimento do dicionário multimídia Makurap-Português optamos pela utilização da metodologia previamente concebida por Brito, Birchall e Galucio (2025). Tal metodologia serviu de base e também norteou o desenvolvimento dos dicionários multimídia do MPEG.

De forma resumida, o fluxo de trabalho técnico para tratamento de dados linguísticos e imagéticos, voltado a elaboração do dicionário, podem ser verificadas na Figura 1:

Figura 1: Etapas do fluxo de trabalho técnico para tratamento dos dados do Dicionário Multimídia Makurap-Português.



Fonte: Elaboração própria.

As quatro sub etapas do fluxo de trabalho técnico para tratamento dos dados ocorreram concomitantemente:

(i) Na primeira etapa, realizamos a transcrição e anotação dos dados no *software* ELAN, seguindo um padrão de trilhas específico concebido por Brito, Birchall e Galucio (2025), etapa que exige também a seleção dos trechos de áudios que deverão compor o dicionário;

(ii) Na segunda etapa utilizamos um *script* que tem como função principal recortar trechos de áudios referentes às palavras e frases selecionadas para composição do dicionário; e

também gerar automaticamente uma planilha no formato .CSV[4] para trabalho manual de organização nas etapas seguintes;

(iii) A terceira e a quarta etapas, referem-se à organização macro e micro estruturais dos dicionários, o que nos exigiu um estudo apropriado para nortear as decisões metodológicas a serem tomadas, seja para organização dos itens lexicais em campos semânticos, seja para estruturação dos verbetes. Essas decisões serão discutidas na próxima seção do presente artigo.

A última etapa metodológica envolve a validação final do dicionário multimídia pela comunidade. Nessa etapa, o material é apresentado à comunidade para avaliação, teste e discussão coletiva de sua organização e conteúdo. No 2º semestre de 2025 apresentamos a primeira versão do dicionário multimídia no território do Rio Guaporé. Uma segunda apresentação geral foi realizada no 1º semestre de 2026. Durante todo o processo, acolhemos diversas sugestões dos colaboradores Makurap que influenciaram de forma fundamental uma reorganização do dicionário, uma vez que essa etapa tem como consequência uma etapa de revisão final do dicionário, com a finalidade de atender as demandas, realizar as correções necessárias e preparar a implementação e ampla divulgação do dicionário para o público em geral.

Apresentaremos na próxima seção os resultados e as discussões acerca da organização estrutural do dicionário multimídia Makurap-Português.

4 O dicionário multimídia Makurap-Português

De acordo com Welker (2004), a macroestrutura da obra lexicográfica refere-se, em diferentes acepções, à organização do dicionário como um conjunto de entradas ou à forma como o corpo do dicionário é organizado. Já o planejamento interno das entradas no dicionário e a decisão sobre o conjunto de informações que acompanham cada um dos verbetes que compõem a obra lexicográfica é uma tarefa referente à microestrutura.

Assim subdividimos essa seção em duas subseções, uma com foco na macroestrutura e outra com foco na microestrutura.

4.1 Macroestrutura

Como pode ser visualizado na Figura 2, abaixo, a *interface* do Dicionário Makurap inclui um conjunto de *hyperlinks*. Os *hyperlinks* referentes a cada campo semântico, em que é possível clicar e se direcionar à lista de palavras desejada; e o *hyperlink* de introdução do dicionário multimídia Makurap, em que é possível clicar e visualizar um texto que apresenta as informações essenciais para familiarizar o usuário com o dicionário. Ele inclui detalhes sobre a concepção do dicionário, dados sobre a língua e o povo Makurap e informações relevantes sobre o projeto de pesquisa no qual

[4] Uma planilha CSV (Valores Separados por Vírgula) é um arquivo de texto que armazena dados tabulares em um formato simples, onde cada linha representa um registro (uma linha da tabela) e os valores de cada coluna são separados por vírgulas, facilitando o uso entre diferentes programas de planilha e bancos de dados.

este produto foi elaborado.

Figura 2: Corpo do Dicionário Multimídia Makurap-Português no formato digital.



Fonte: Dicionário Multimídia Makurap-Português (Versão 1.0).

Para facilitar o manuseio do dicionário pelo usuário, a macroestrutura organiza as entradas lexicais em campos semânticos[5]. Dessa forma, caso o usuário tenha interesse em aprender os números da língua Makurap, por exemplo, ele pode selecionar essa categoria e encontrar todas entradas referentes. Além disso, para alguns campos semânticos há subcampos semânticos[6]:

Tabela 1: Lista de campos semânticos que possuem subcampos semânticos.

Campo Semântico	Sub Campo semântico
Animais	Aves, Peixes, Mamíferos, Anfíbios e Répteis, Insetos e Outros Animais
Plantas e Fungos	Amendoins, Tubérculos, Bananas, Castanhas, Milhos, Palmeiras, Outros Frutos Comestíveis, Plantas de Rama, Árvores da Floresta e Outras Plantas
Natureza	Astros e Eventos Naturais, Itens Relativos a Árvore, Itens Relativos a Água, Terra e Rochas e Outros
Manufaturas e Itens do Dia-a-Dia	Adornos Corporais, Instrumentos Musicais, Itens de Caça e Pesca, Itens de Casa, Itens de Cozinha, Itens da Roça e Outros Materiais
Pessoas	Termos para Pessoas e Pronomes Pessoais
Parentesco e Relações Familiares	Fala Masculina e Fala Feminina

Fonte: Elaboração Própria.

[5] Há a possibilidade do usuário ordenar os itens lexicais por ordem alfabética ou no formato de lista e ocultar as imagens utilizadas para ilustração.

[6] É possível que futuramente, com a chegada de novos dados, os demais campos semânticos (Comidas e Bebidas, Atividades e Eventos, Estados e Propriedades, Cores, Partes do Corpo, Secreções Corporais, Números e Itens Gramaticais) também sejam divididos em subcampos semânticos, bem como os itens categorizados como "Outros" podem formar um novo campo semântico.

Cada campo semântico agrupa as entradas lexicais internamente por ordem alfabética, por *default*. Porém, seguindo orientação dos pesquisadores indígenas, em alguns casos, como no campo semântico de animais, a organização dentro do campo semântico segue uma organização específica.

Durante a validação, os colaboradores indígenas sugeriram que determinadas espécies de animais fossem agrupadas. A análise dessas sugestões revelou a existência de categorias semanticamente organizadas que orientam a forma como os falantes classificam os animais. Em resposta a essas observações, optou-se por adaptar a macroestrutura do dicionário para refletir essas relações taxonômicas.

Um dos aspectos observados foi a existência de lexemas considerados genéricos capazes de ordenar uma série de hipônimos, por exemplo, *potkap* 'peixe' trata-se de hiperônimo para *mbalera* 'tucunaré', *kotixa* 'peixe quebra-galho', *mbiria* 'traíra', *towa* 'tamuatá', etc. É possível observar que essas relações de hiperonímia e hiponímia estão diretamente ligadas a um conjunto de características que aproximam e diferenciam itens dentro de um mesmo campo semântico e os classificam dentro de uma hierarquia taxonômica.

Segundo Cruse (2000), as hierarquias taxonômicas são sistemas classificatórios que refletem a maneira como os falantes de uma língua categorizam o mundo e a experiência, sendo possível observar duas características importantes nas hierarquias taxonômicas: (i) uma taxonomia bem estruturada oferece um conjunto ordenado de categorias em diferentes níveis de especificidade, partindo do mais geral (hiperônimo) para o mais específico (hipônimo); (ii) a hiperonímia e a hiponímia são relações semânticas extremamente relevantes para demonstrar traços em comum da taxonomia observados de diversas formas por diferentes culturas (cf. Cruse, 2000, p. 168).

Em relação à língua Makurap, essa classificação hierárquica pode não seguir necessariamente a taxonomia das ciências biológicas, mas pode ser moldada pelo conhecimento empírico do povo Makurap e pelo uso cultural de determinado animal ou planta. Como observa Biderman (2001a), a geração do léxico resulta de sucessivos atos de cognição da realidade e de categorização da experiência, cristalizados em palavras e termos. Essas experiências encontram-se codificadas na língua Makurap e constituem uma importante fonte de conhecimento sobre o ambiente e sobre a cultura. Quando solicitados a explicar o significado de um animal ou de uma planta, os colaboradores Makurap geralmente não fornecem apenas um equivalente de tradução, mas mobilizam informações relacionadas à aparência física, ao habitat, ao comportamento, ao uso cultural e, por vezes, às associações cosmológicas desses referentes.

Um exemplo é a classificação dos jacus na subcategoria semântica de aves (categoria de animais) do dicionário multimídia. Esses dispõem de conjunto de atributos considerados essenciais e suficientes para filiação de determinado animal como ave e seguidamente como jacu. Esses atributos são descritos pelos sabedores Makurap:

Tabela 2: Exemplo de descrição das aves classificadas como jacus.

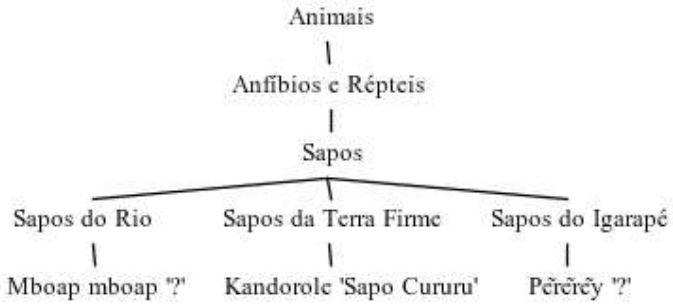
Animal	Descrição
Wakopem	"Se chama 'jacu' em português. É um animal de terra firme, preto, tem dois brincos vermelhos, o pescoço é vermelho, a cabeça é preta, o bico é vermelho, as penas na parte de baixo do corpo são avermelhadas. Ele marca o amanhecer."
Wako axiit	"Se chama 'cujubim' em português. É um pássaro que vive no baixo do igarapé, tem brinco debaixo do peito, a cabeça é branca, a pena é preta no corpo e branca nas asas. Esse pássaro adivinha o amanhecer."
Wako uat mbaat	"Não se sabe o nome em português. É um pássaro que anda de bando, vive na beira do campo, as cores são as mesmas do wakopem, pena preta e brinco vermelho. Eles juntam muitos ovos e os criam juntos, vários pássaros chocam o mesmo ovo."
Wako ndandali	"Se chama 'jacuzinho' em português. É um pássaro que vive em todo canto, ele é todo vermelho, porém nas costas é mais escuro. O canto é semelhante ao nome 'ndandali'. Ele adivinha as mortes, o seu canto acompanha o choro e agoura as pessoas."

Fonte: Elaboração própria.

Como observado na Tabela 2, a categoria de jacus designa um conjunto de aves percebidas pelos Makurap como pertencentes a uma mesma classe semântica e reúne aves da família *Cracidae*, mas não inclui outras aves classificadas na mesma família, como os mutuns e aracuãs. Essa classificação parece fundamentar-se principalmente em características anatômicas compartilhadas, como a aparência geral do corpo e a coloração das penas. Ao mesmo tempo, os sabedores mencionam, simultaneamente: aspectos relacionados ao habitat, ao comportamento, à vocalização e às associações cosmológicas atribuídas a cada ave. Portanto, as descrições evidenciam critérios que permitem também distinguir as diferentes espécies agrupadas nessa categoria.

Um outro exemplo está na categoria de sapos, durante o trabalho de campo registramos 14 nomes de sapos, embora estes animais fossem classificados dentro de uma mesma categoria, o nosso colaborador observou o habitat como atributo crucial para classificação em hierarquias taxonômicas, como pode ser verificado na figura abaixo:

Figura 3: Exemplo de classificação em hierarquia taxonômica dos sapos.



Fonte: Elaborado pelos autores.

A Figura 3 também evidencia uma das dificuldades encontradas na elaboração do dicionário: a ausência ou o desconhecimento dos equivalentes em português para determinados animais nomeados em Makurap. Essa limitação dificulta o reconhecimento preciso da espécie a que cada lexema se refere. As informações descritivas apresentadas na Tabela 2, por exemplo tornam-se fundamentais, pois fornecem características que auxiliam na identificação dos organismos e contribuem tanto para a organização macroestrutural do dicionário, ao possibilitar o agrupamento dos itens em sub categorias semanticamente relacionadas, quanto para a microestrutura, ao subsidiar a elaboração de verbetes mais informativos – questão que veremos na seção posterior.

Nesse contexto, a decisão metodológica tomada foi conceber uma organização não padrão do dicionário para demarcar essas relações de taxonomia. Para isso, seguindo uma sugestão e solicitação de pesquisadores indígenas participantes do projeto de pesquisa e falantes da língua Makurap, a macroestrutura do dicionário foi organizada de forma a apresentar essas relações de hiperonímia e hiponímia como subgrupos dentro de campos semânticos mais amplos.

À guisa de ilustração, observe a apresentação do hiperônimo *potkap* 'peixe' (figura 4), ao qual estão associados outros termos que identificam os peixes específicos e que também seguem uma taxonomia específica: peixes de escama, peixes de couro sem escamas, peixes pequenos do igarapé etc.

Assim, na definição da macroestrutura do dicionário Makurap, optou-se por uma personalização da ordem das entradas lexicais. Essa personalização permite realizar manualmente um reordenamento dos dados linguísticos na planilha CSV dispondo os itens lexicais na ordem desejada, como pode ser visualizado na figura 4 abaixo:

Revista Letras Escreve Macapá, AP v. 16 – n. 2 ISSN 2238-8060 - **VOLUME ATEMÁTICO - ARTIGO**

Figura 4: Exemplo de preenchimento da planilha CSV com nota descritiva e ordenamento personalizado de entradas lexicais.

potkap	peixe	Termo geral para nomear qualquer espécie de peixe
2.5.1 Peixes de couro		
anole	surubim	
anole wop	pirarara	
i'kiri	mandi da beira do rio	
mõkowa	mandi sapo	
ãrangã	cangati	
atiti wop	peixe sabão	
ndoko	poraquê	
iwotxon	peixe cachorro grande, peixe morcego	
xao	arraia	
2.5.2 Peixes de escama grande		
mbalera	tucunaré	
mbilia	traíra	
mbilia xurui	traíra do igarapé	
mbilia'i	traírinha do igarapé	
ite	piau grande	

Fonte: Elaboração própria

Observe que na figura acima, o subcampo semântico de peixes apresenta as espécies de peixe de couro (*anole* 'surubim', *anole wop* 'pirarara', *i'kiri* 'mandí da beira do rio' e *mõkowa* 'mandí sapo') seguidas das espécies de escama grande (*mbalera* 'tucunaré' e *mbilia* 'traíra'). Esse formato de organização das entradas lexicais está disponível para a categoria semântica de animais como um segundo nível de subcampo semântico. A Tabela 3, abaixo, sintetiza os critérios semânticos utilizados pelos falantes para agrupar e distinguir os organismos dentro dessa categoria.

Tabela 3: Critérios semânticos de classificação dos animais em Makurap.

Características	Conhecimentos envolvidos
Físicas	Envolve a observação da aparência física e características anatômicas dos animais que permitem distinguir uns dos outros. Por exemplo, aves podem ser descritas e categorizadas a partir do seu tamanho, formato do bico, fisionomia, presença de papo, coloração das penas, etc.
Ecológicas	Envolve a observação do habitat e do comportamento do animal. Por exemplo, os sapos podem ser categorizados como animais que vivem na beira do rio, na beira do igarapé ou na terra firme. Por outro lado, aves podem ser descritas a partir de seu modo de se reproduzir, criar filhotes, construir ninhos, etc.
Culturais	Envolve a prática de produção de cultura material e alimentos com partes do corpo do animal. Por exemplo, os mamíferos podem ser categorizados e descritos por terem suas carnes usadas na alimentação. No caso específico da queixada, há referência ao uso de seus ossos para criação de instrumentos de caça.
Cosmológicas	Envolve a experiência espiritual com os animais. Alguns animais podem ser descritos por estarem associados a uma narrativa tradicional, a uma crença, etc. Um exemplo específico é o pássaro *akaxkiang*, tipo de maracanã

	conhecido como "papagaio do diabo", em que se acredita que uma pessoa ao olhá-lo terá uma dor de cabeça imediata e poderá ficar doente

Fonte: Elaboração própria

A classificação semântica dos itens lexicais de fauna no dicionário multimídia Makurap-Português foi construída a partir da integração de conhecimentos empíricos compartilhados pelos próprios falantes, fundamentados na observação e na experiência acumulada em suas interações com os animais. Os dados analisados indicam que características físicas e ecológicas desempenham papel central na organização das hierarquias taxonômicas identificadas até o momento, permitindo o agrupamento de espécies percebidas como semelhantes. Outros aspectos, como uso cultural e associações cosmológicas, também se mostram relevantes, sobretudo para a caracterização mais detalhada dos referentes e para a elaboração de verbetes mais informativos.

Consideramos essa organização indispensável, pois, a partir dela, foi possível organizar todo o dicionário, com a finalidade de sistematizar e organizar os itens lexicais em campos e subcampos semânticos que representassem, da forma mais fidedigna possível, o saber empírico dos falantes. Afinal, conforme explicitado por Biderman (2001b), o saber empírico não difere essencialmente do saber científico, uma vez que ambos se estruturam por meio de processos de categorização e de nomeação subsequentes. Entretanto, ainda são necessários estudos mais aprofundados para compreender de forma mais precisa os critérios de classificação empregados pelos Makurap e sua aplicação às demais categorias semânticas representadas no dicionário. Atualmente apenas a categoria de animais dispõe de um segundo nível de subcampo semântico.

4.2 Microestrutura

A microestrutura do Dicionário Multimídia Makurap-Português inclui, principalmente: lema, transcrição fonética, tradução para o português, frases contextuais (em Makurap com tradução para o português), imagem para ilustrar o verbete e áudios dos verbetes e das frases ilustrativas de uso e, geralmente, apenas um equivalente de tradução para o português. Porém, para alguns casos específicos, foi necessário usar alguns recursos opcionais da planilha CSV, como descrição e itens relacionados.

A seguir, ilustramos questões identificadas na elaboração do dicionário Makurap para as quais determinadas soluções analíticas influenciaram a composição da microestrutura dos verbetes e discutimos as necessidades específicas apresentadas pelos dados da língua e como os recursos opcionais do método utilizado podem evidenciar no dicionário as relações semânticas e nuances lexicais identificadas.

Nos dados analisados para compor o dicionário Makurap, foram identificados casos de variação morfofonológica, palavras que possuem o mesmo significado ou função gramatical, mas com formas diferentes para cada ambiente fonológico, como os listados nos exemplos da Tabela 3

abaixo:

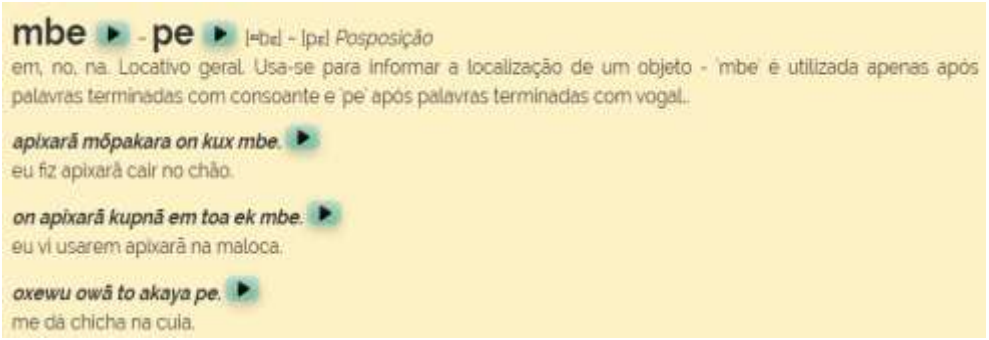
Tabela 4: Exemplo de variação.

mbe pe	'Locativo (ocorre após consoantes)' 'Locativo (ocorre após vogais)'	ẽl̃i ẽr̃i	'Ablativo' (variação livre)

Fonte: Elaboração própria

Nos exemplos da Tabela 4, a posposição locativa possui uma forma quando ocorre após palavras terminadas em vogal e outra quando ocorre após palavras terminadas em consoante, ou seja, trata-se de uma mesma palavra, que assume uma forma específica motivada por processos morfofonológicos (Braga, 2005); e a posposição ablativa varia livremente.

Figura 5: Exemplo de entrada lexical com variação



Fonte: Dicionário Multimídia Makurap-Português

Observe que incluímos as duas formas em uma mesma entrada lexical e, como mostrado na imagem acima, é possível acrescentar um pequeno texto descritivo para esses casos. Além disso, por identificarmos essa palavra como um item gramatical, elaboramos uma nota explicativa para indicar suas funções

A inclusão das palavras gramaticais (pronomes, posposições, partículas e conectores), como ilustrado na Figura 5, acima, representou uma questão importante no processo de elaboração do dicionário. Embora não necessariamente possuam um equivalente de tradução em português, essas palavras são essenciais para compreensão e produção de qualquer texto em Makurap. Sua inclusão exigiu decisões específicas quanto à organização das entradas e à apresentação de informações sobre suas funções e usos, conforme ilustrado pelos exemplos da Tabela 5.

Tabela 5: Exemplo de palavras 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.

Fonte: Elaboração própria

Na Tabela 5, primeiramente, observa-se a distinção entre os pronomes de primeira pessoa do plural, que se apresentam em duas formas: exclusiva *tetyã* e inclusiva *kitetyã* (Braga, 2005). Em seguida, temos dois morfemas que fazem referência a tempo futuro, diferenciando entre futuro genérico *eya* e imediato *ouga* (Albuquerque, 2025).

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. Justamente porque partimos do pressuposto de que devemos apresentar na microestrutura, para fins pedagógicos, as informações que permitam a compreensão do significado da palavra apresentada, além da simples equivalência (cf. Fargetti; Mondini, 2021).

Portanto, a criação de verbetes descritivos tornou-se fundamental não só para o campo semânticos de itens gramaticais, mas também para determinados campos semânticos, como de animais, variando de acordo com os objetivos gerais do projeto e com o público-alvo.

Por exemplo, na Tabela 2 acima, apresentamos o lexema *Wako uat mbaat* que se refere a um pássaro do subcampo semântico dos jacus. E, embora não tenha sido possível identificar o nome preciso para se referir a esse pássaro em português, os pesquisadores Makurap foram capazes de fornecer informações detalhadas sobre sua aparência física, comportamento, habitat, cosmologia e outras características relevantes para sua identificação. Essas descrições revelam um conjunto de conhecimentos associados ao referente que ultrapassam a simples tradução lexical e que podem ser compreendidos como informações enciclopédicas relevantes para a compreensão do significado da palavra.

A solução encontrada para preencher o 'vazio' dos nomes de animais para os quais não identificamos, até o momento, o nome equivalente em português foi incluir essas informações enciclopédicas nos verbetes descritivos do dicionário multimídia.

Conclusão

O dicionário multimídia da língua Makurap conta atualmente com 1.075 entradas lexicais distribuídas em 15 grandes campos semânticos. A elaboração deste material permitiu ampliar a documentação da língua e iniciar a construção de uma base de dados lexicais. Para cada lexema que exige informações gramaticais, é apresentado um pequeno texto descritivo que explica seu significado ou função. Além disso, recursos como "itens relacionados" e a possibilidade de incluir duas formas de uma mesma palavra em uma única entrada lexical foram utilizados para tornar a consulta mais clara e funcional para os usuários.

Essas decisões foram tomadas considerando o uso do dicionário como ferramenta de aprendizagem da língua Makurap no contexto de segunda língua. Apresentar informações gramaticais sobre esses itens justifica-se não só por uma questão de tradução, mas principalmente para facilitar a compreensão do significado dessas palavras pelos usuários a quem esse dicionário se destina.

Para além dessas contribuições, no atual estágio de elaboração do dicionário multimídia, estamos aplicando a classificação definida em conjunto com os membros da comunidade, somente para os animais e plantas comestíveis. Uma possível etapa futura inclui a expansão e o aprofundamento do estudo de classificação semântica do léxico, com foco na descrição dos sistemas de categorização semântica estabelecidos pelo povo Makurap, tendo em vista a elaboração de árvores de domínio que permitam representar esses sistemas na composição do dicionário. Esse estudo teria como consequência uma reorganização geral dos campos semânticos em uma 2ª versão do Dicionário Multimídia Makurap-Português.

O dicionário multimídia Makurap-Português foi apresentado e testado pela comunidade Makurap em uma atividade de campo realizada em março de 2026, durante essa atividade observamos a sua recepção e acolhemos opiniões e sugestões, e após uma última etapa de revisão de pontos que necessitam de melhoria, será publicado no site Dicionários Multimídias para Línguas Indígenas[7] do MPEG, a fim de que estejam disponíveis para uso dos falantes, pesquisadores de diferentes áreas de conhecimento e o público em geral.

Com este trabalho, esperamos contribuir para a consolidação e divulgação de um modelo replicável de elaboração de dicionários multimídia voltados a línguas ameaçadas, reaplicável a qualquer língua e reprodutível de forma autônoma por pesquisadores interessados. No caso do Dicionário Multimídia Makurap-Português, ele auxilia na documentação, fortalecimento e revitalização da língua Makurap, aumentando a motivação dos falantes, fluentes e em potencial, e promovendo o prestígio da língua. Como tecnologia social, esse dicionário poderá servir futuramente como material pedagógico de referência e recurso tecnológico de consulta, configurando-se como uma ação dinâmica que ajuda a preservar a rica herança linguística e cultural

[7] Ver em: Dicionários de Línguas Indígenas | Museu Paraense Emílio Goeldi.

dos Makurap, além de oferecer um recurso valioso para a aprendizagem da língua.

Referências

ALBUQUERQUE, Frida Natália Lobato de. **Estudo Tipológico-Comparativo de Marcação de Tempo nas Línguas do ramo Tupari da família Tupi**. Dissertação (Mestrado). Programa de Pós-Graduação em Letras - Universidade Federal do Pará. Belém, 2025.

BEVILACQUA, Cleci Regina; FINATTO, Maria José Bocorny. Lexicografia e Terminografia: alguns contrapontos fundamentais. **ALFA: Revista de Linguística**, v. 50, n. 2, 2006.

BIDERMAN, Maria Tereza Camargo. As Ciências do Léxico in: OLIVEIRA. A. M; ISQUERDO. A. N. (Org.). **As ciências do léxico**: lexicologia, lexicografia, terminologia. Campo Grande: Editora UFMS, p. 13-22, 2001ª.

BIDERMAN, Maria Tereza Camargo. Terminologia e lexicografia. **TradTerm**, v. 7, p. 153-181, 2001b.

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.

BORBA, Francisco da Silva. **Organização de dicionários**: Uma introdução à lexicografia. São Paulo: Editora Unesp, 2003.

BRAGA, Alzerinda de Oliveira. **Aspects morphosyntaxiques de la langue makurap**. 2005. Tese (Doutorado em Ciências da Linguagem) – Universidade de Toulouse-Le Mirail, Toulouse, 2005.

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.

CRUSE, Alan. **Meaning in Language**: An Introduction to Semantics and Pragmatics. Oxford Linguistics. New York: 2000.

FARGETTI, Cristina Martins; MONDINI, Juliana Nazatto. Dicionários de Línguas Indígenas e Lexicografia Pedagógica. **Entrepalavras**, Fortaleza, v. 11, n. esp., p. 19-40, 2021.

GALUCIO, Ana Vilacy *et al*; NOGUEIRA, Antonia Fernanda; COSTA, Carla Daniele. 2023. **Makurap**: Documentation of Language and Culture. Endangered Languages Archive. Handle: http://hdl.handle.net/2196/73b9o01s-1c78-2246-5187-0fm978700a80. Acesso em 13 ago. 2024.

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

GONÇALVES, Artur Garcia. A construção de dicionário bilíngue e bidialetal para a língua

Medzeniakonai (Baniwa-Koripako). **Cadernos de Linguística**, v. 6, n. 3, p. e788-e788, 2025.

KROSKRITY, Paul V. Designing a Dictionary for an Endangered Language Community: Lexicographical deliberations, language ideological clarifications. **Language documentation and Conservation**, Vol. 9, 2015, p.140–157, 2015.

KRIEGER, M. G. O dicionário de língua como potencial instrumento didático. In: ISQUERDO, A. N.; ALVES, I. M. (orgs.). **As Ciências do Léxico:** lexicologia, lexicografia, terminologia. Campo Grande: Ed. UFMS; São Paulo: Humanitas, 2007, p. 295-309.

SOARES, Matheus Augusto Ribeiro; RODRIGUES JUNIOR, Douglas da Costa; GALUCIO, Ana Vilacy. A dicionarização como ferramenta de revitalização de línguas ameaçadas: dois relatos de experiência. **Cadernos de Linguística**, [S. l.], v. 5, n. 1, p. e704, 2024.

WAJURU, Jaqueline; MACURAP, Josiclei; NOGUEIRA, Antonia Fernanda; GALUCIO, Ana Vilacy; COSTA, Carla Daniele Nascimento. Engaging Wajuru /Wayoro and Makurap Communities in Collaborative Documentation: Recording, Learning, and Communicating. **Dutkansearvvi dieà111alaà161 áigeà10Dála**, v. 09, p. 54-74, 2025.

WELKER, Herbert Andreas. **Dicionários**: Uma pequena introdução à Lexicografia. 2ed. revista e ampliada. Brasília: Thesaurus, 2004.