TPBCI Tendências da Pesquisa Brasileira em Ciência da Informação

MODELAGEM NAVEGACIONAL PARA GESTÃO DO LÉXICO DA LÍNGUA GUARANI-MBYÁ: CAMINHOS METODOLÓGICOS[1]

NAVIGATIONAL MODELING FOR MANAGEMENT OF THE GUARANI-MBYÁ LANGUAGE LEXICON: METHODOLOGICALS PATHWAYS

Micheline Maria Costa de Azevedo [2]
Bruna Lessa[3]
Ivana Pereira Ivo[4]

Resumo: A situação de vulnerabilidade enfrentada pelos povos indígenas evidencia a urgência de estratégias para a organização, preservação e difusão de seu conhecimento tradicional. Nesse sentido, iniciativas voltadas para esses aspectos podem desempenhar um papel essencial na promoção de uma sociedade mais equitativa e inclusiva. Esse artigo descreve os aspectos iniciais para a classificação de uma língua indígena, destaca os seus elementos informacionais, assim como alguns pontos-chave da cultura Guarani-Mbyá. Ressalta os primeiros procedimentos metodológicos para o desenvolvimento de uma estrutura de classificação da língua Guarani-Mbyá, uma das línguas indígenas brasileiras, no âmbito da proposta da modelagem navegacional para o desenvolvimento do Protótipo para Gestão do Léxico da língua indígena. A trilha metodológica foi construída sob uma perspectiva de diálogos interdisciplinares, de natureza qualitativa e do tipo descritiva, documental e aplicada, o que contribuiu para a compreensão e representação dos elementos essenciais do léxico Guarani-Mbyá e sua estrutura classificatória. Os dados foram coletados do Dicionário Bilíngue Guarani-Mbyá/Português, isolando-se as palavras com sentido básico, polissêmicos, sinônimos, antônimos e variedades dialetais. A estrutura classificatória, parte dos resultados da pesquisa, categoriza os termos nos aspectos semânticos e sintáticos, identificando as

[1] Artigo submetido, avaliado, aprovado, apresentado e premiado no premiado no XXIV Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB2024–Vitória-ES).

[2] Mestra em Ciência da Informação pelo Programa de Pós-Graduação em Ciência da Informação – PPG-CI/UFBA. Universidade Federal do Rio Grande do Norte (UFRN). E-mail: michelineazevedo@gmail.com. ORCID: https://orcid.org/0000-0001-7964-2169

[3] Doutora em Ciência da Informação e Professora Adjunta do Departamento de Documentação e Informação, do Instituto de Ciência da Informação, da Universidade Federal da Bahia (UFBA). E-mail: brunalessa@ufba.br. ORCID: http://orcid.org/0000-0003-4485-203X.

[4] Doutora em Linguística pela Universidade Estadual de Campinas (UNICAMP). Professora do Departamento de Linguística do Instituto de Estudos da Linguagem (IEL). Universidade Estadual de Campinas (UNICAMP). E-mail: ivanaivo@unicamp.br. ORCID: https://orcid.org/0000-0002-5180-7483.

relações entre eles. Os resultados obtidos com a modelagem confirmam a viabilidade da proposta, apresentando um modelo conceitual para o desenvolvimento de um Sistema de Recuperação da Informação no futuro.

Palavras-Chave: Sistema de Organização do Conhecimento Indígena; Guarani-Mbyá; Lexicografia.

Abstract: *The vulnerability faced by indigenous peoples highlights the urgency of strategies for the organization, preservation, and dissemination of their traditional knowledge. In this sense, initiatives focused on these aspects can play an essential role in promoting a more equitable and inclusive society. This article describes the initial aspects of the classification of an indigenous language, highlights its informational elements, as well as some key points of the Guarani-Mbyá culture. It highlights the first methodological procedures for the development of a classification structure for the Guarani-Mbyá language, one of the Brazilian Indigenous languages, within the scope of the navigational modeling proposal for the development of the Prototype for Management of the Lexicon of the indigenous language. The methodological path was constructed from a perspective of interdisciplinary dialogues, of a qualitative nature and of the descriptive, documentary, and applied type, which contributed to the understanding and representation of the essential elements of the Guarani-Mbyá lexicon and its classificatory structure. The data were collected from the Guarani-Mbyá/Portuguese Bilingual Dictionary, isolating words with basic meaning, polysemic words, synonyms, antonyms, and dialectal varieties. The classification structure, based on the research results, categorizes the terms in semantic and syntactic aspects, identifying the relationships between them. The results obtained from the modeling confirm the viability of the proposal, presenting a conceptual model for the development of an information retrieval system in the future.*

***Keywords*:** *Indigenous Knowledge Organization Systems; Guarani-Mbyá; Lexicography.*

1 INTRODUÇÃO

A perda de conhecimentos indígenas, processo que inspira essa proposta, inclui os diferentes aspectos, sejam culturais ou linguísticos. Este processo se inicia com a colonização e permanece até hoje. O retorno ao passado, especialmente ao Séc. XVIII – período da publicação do Diretório dos Índios de 1755 (Almeida, 1997), revela os aspectos da política linguística de proibição das línguas indígenas, no contexto escolar, o que demonstra ser um dos aspectos que auxilia, no processo de apagamento das línguas indígenas, como se vê no texto a

seguir:

> § 6º – [...] para desterrar esse perniciosíssimo abuso, será um dos principais cuidados dos Diretores estabelecer nas suas respectivas Povoações o uso da Língua Portuguesa, não consentindo por modo algum, que os Meninos, e as Meninas, que pertencerem às Escolas, e todos aqueles Índios, que forem capazes de instrução nesta matéria, usem da língua própria das suas Nações, ou da chamada geral; mas unicamente da Portuguesa [...] (Diretório dos Índios, 1755 *apud* Almeida, 1997, p. 198).

Ao longo dos séculos, pela não aplicação de políticas linguísticas específicas, as línguas indígenas do Brasil têm sido invisibilizadas o que contribui com essa perda que igualmente reforça o não conhecimento sobre a cultura e sociedades indígenas. No âmbito acadêmico, crescem os debates em como promover e disponibilizar as ações em torno da manutenção e disseminação dos conhecimentos ancestrais, ainda mantidos pelos povos indígenas.

Nesse cenário, no âmbito da Ciência da Informação, e no domínio da Organização do Conhecimento (OC), é importante o desenvolvimento de uma abordagem que alie os estudos teóricos aos aplicados e que colaborem com a representação e organização do conhecimento em contextos informacionais e culturais diversos.

No artigo, ora retratado, o povo Guarani-Mbyá, caracterizando-se como comunidades bilíngues (Guarani-Mbyá-Português). Na atualidade, o povo Guarani, está localizado em diferentes países da América Latina, no Brasil, Paraguai, Uruguai, Argentina e Bolívia, sendo a língua Guarani-Mbyá a mais falada da família Tupi-Guarani, segundo Rodrigues (1945). No Brasil, essa parcialidade localiza-se nos estados do Espírito Santo, Rio de Janeiro, São Paulo, Paraná, Santa Catarina, Rio Grande do Sul, além de famílias residentes no Pará e em Tocantins. As parcialidades de povos Guarani, no Brasil, são divididas em quatro etnias, a saber: os “Guarani-Mbyá, Nhandeva, Nhandewa e Guarani-Kaiowá”, conforme

explica Ivo (2018, p. 30). A despeito do significativo número de falantes do Guarani-Mbyá, no Brasil, a língua é praticamente desconhecida pela maior parte da sociedade brasileira.

A complexidade e a particularidade das línguas indígenas demandam uma abordagem que integre conhecimentos linguísticos e culturais para garantir a manutenção dos saberes e práticas das comunidades indígenas, promovendo o respeito ao seu legado de forma ética e inclusiva. Assim, como parte da pesquisa, no âmbito do Mestrado em Ciência da Informação, foi levantada a questão que norteia o trabalho: como desenvolver uma estrutura de classificação, sensível e culturalmente relevante para organizar o conhecimento da língua Guarani-Mbyá, considerando-se as especificidades linguísticas, culturais e sociais dessa etnia? O objetivo deste trabalho é, pois, descrever os passos tomados para desenvolver a estrutura de classificação da língua Guarani-Mbyá, utilizando o Dicionário Bilíngue Guarani-Mbyá/Português – versão digital, como corpus de análise, para o desenvolvimento da modelagem navegacional que represente, de forma adequada, a linguagem e a cultura do povo Guarani-Mbyá, por meio do Protótipo para Gestão do Léxico da Língua Guarani-Mbyá (doravante, Prombyá). Este trabalho pretende ressaltar, ainda, a importância de uma abordagem interdisciplinar nos estudos da Organização do Conhecimento, para compreender e representar adequadamente o léxico e a organização do conhecimento indígena, levando-se em consideração a cosmovisão e os saberes da comunidade Guarani-Mbyá, contribuindo para a manutenção e difusão do seu patrimônio linguístico e cultural.

Na próxima seção, o artigo irá dedicar-se a apresentar uma síntese sobre o povo Guarani-Mbyá, destacando os aspectos relevantes de sua cultura e de seu modo de vida.

2 DESENVOLVIMENTO

A denominação Guarani-Mbyá, surge na literatura do século XVIII – segundo Meliá (1983). A depender da localidade, recebiam uma terminologia específica. No entanto, genericamente eram conhecidos como 'Caaguá' e 'Monteses' e que, genericamente designavam "[...]os habitantes da mata"[5] , conforme (Meliá, 1983, p. 52).

Os povos Guarani-Mbyá são conhecidos por cultivar uma relação harmônica com o outro, com a natureza e, principalmente, os seres espirituais que os orientam. A cultura do povo Guarani-Mbyá é rica, sua vida sociocultural se pauta pela relação harmônica com o outro, com a natureza e, principalmente, com os seres espirituais que os orientam, bem como a ideia de reencarnação. A religiosidade é muito presente e atuante na sociedade Guarani-Mbyá, por meio da qual se observa a presença de Nhanderu — pai criador ou 'nosso primeiro pai' — com frequência, presente e devotado nas músicas, nos textos e na literatura Mbyá, é o pai dos demais deuses. Machado e Ferreira (2016) apresentam detalhes sobre como isso de delineia:

> De acordo com a cosmogênese guarani, Nhanderu (Pai de toda criação), nesse caso representado pelo zênite, criou quatro deuses principais para a criação da Terra e dos seres viventes, estes quatro deuses são representações dos quatro pontos cardeais. O Norte é representado por Jakaira, deus da neblina vivificante e das brumas que abrandam o calor, origem dos bons ventos. O Leste é Karai, deus do fogo e do ruído do crepitar das chamas sagradas. No Sul, Nhamandu, deus do Sol e das palavras, representa a origem do tempo-espaço primordial. No Oeste, Tupã, é deus das águas, do mar e de suas extensões, das chuvas, dos relâmpagos e dos trovões (Machado; Ferreira, 2016, p. 43).

[5] Desde el siglo XVIII a estos grupos, en cuanto no colonizados, se les designaba con el nombre genérico de Caaguá o Monteses, es decir, habitantes de la selva (Meliá, 1983, p. 52).

Ivo (2018) acrescenta que os deuses Guarani[6] possuem um nome correlato feminino, ou seja, Nhanderu 'nosso pai' (masculino)/Nhandexy 'nossa mãe' (feminino), assim como as demais divindades, Karaí (masculino)/Kerexu (feminino), Tupã (masculino)/Pará (feminino), Jakairá (masculino)/Tataxi(feminino).

As casas de reza, de acordo com o exposto por Ivo (2018, p. 55), são chamadas de Opy, onde se realizam as "[...] as cerimônias, ritos e outras atividades, como a cura dos enfermos." Acrescenta a autora, que os não-indígenas podem ter acesso a esse ambiente, entretanto, determinadas cerimônias ocorrem apenas entre nativos Mbyá, informação corroborada por Litaiff (1996).

O Tekó (o modo de ser e viver do Guarani-Mbyá) é outro símbolo deste povo de grande importância, conforme retratado por Ivo (2018):

> [...] é ele que conecta a sociedade como um todo. O Tekó conecta as pessoas, une a comunidade, mesmo quando ela se encontra sob divisões políticas. É o Tekó que dá ao indivíduo o pertencimento ao grupo, é o que dá a interpretação da vida diária, do sobrenatural, da vida pós-morte e da atuação dos deuses. É a descrição do Tekó que mostra, por exemplo, a organização do calendário anual Guarani em dois ciclos. Nos termos Mbyá, o Ará Pyau, o ciclo novo e o Ará Yma, o ciclo velho, que dão as regras sobre o tempo do plantio, o tempo da colheita, o tempo da caça e da pesca, o tempo para a realização dos batismos etc. Isso é ensinado costumeiramente na casa de reza (Ivo, 2018, p. 55).

Há, na cosmovisão do Guarani-Mbyá, uma noção diferente de posse sobre os recursos naturais, como partes da natureza (sol, plantas etc.) não podem ser possuídos e são, na realidade, empréstimos provenientes de Nhanderu.

[6] Nesse trabalho, adotaram-se as normas da Associação Brasileira de Antropologia (ABA) que dispensam a flexão de gênero e número nos etnônimos (Reunião Brasileira de Antropologia, 1955).

A língua é outro símbolo de origem divina que contribui para a harmonização da comunidade indígena, ao mesmo tempo que estabelece os parâmetros de distinção em relação a grupos não indígenas, considerados, por essa perspectiva, como não receptores das palavras sagradas.

Na próxima seção, apresenta-se o diálogo interdisciplinar necessário à compreensão das fronteiras necessárias ao delineamento e planejamento do Prombyá

3 POR UMA ABORDAGEM INTERDISCIPLINAR PARA A ORGANIZAÇÃO DO CONHECIMENTO INDÍGENA

Pensar em uma organização da informação indígena é um desafio, principalmente, ao se agregar os aspectos que envolvem a representação de uma cultura diversa e não hegemônica, aliado aos desafios em torno de léxico de línguas orais.

Por vezes, o envolvimento de múltiplas visões auxilia no processo de uma modelagem complexa. Nesse momento, percebe-se que o diálogo interdisciplinar, segundo Pombo (2005), é o que possibilita a união dessas visões, quais sejam: [1] a Organização da Informação que apresente uma forma coerente de classificá-lo esse saber ancestral que dialogue e acolha os elementos da sociedade e da cosmovisão indígena sua Classificação (Ranganathan, 1967); [2] a Linguística, apresentada por Cabré (1998, 1999 e 2008) e a descrição do léxico língua Guarani – Mbyá e suas formas de expressão, elucidados por Ivo (2018) e Ivo *et al*. (2024) que articuladas, possibilitam essa construção.

A atuação da Organização do Conhecimento Indígena se concentra em compreender o seu alcance e os seus valores ancestrais. Neste sentido, a compreensão da cosmologia indígena apresenta sinalizações preciosas.

Para atuar na perspectiva da organização, a pesquisa emprega a Teoria da Classificação Facetada (TCF) (Ranganathan, 1967), aborda os "[...] principais fundamentos teórico-metodológicos utilizados pelos classificacionistas para modelagem e representação de um domínio do conhecimento[...]", conforme apresenta Lima (2020, p.68).

Esta Teoria, como corrente teórica necessária para nortear um processo de classificação, tem como cerne a ideia de faceta, que para Dahlberg (1978):

> Os sistemas de conceitos, cujos elementos foram até agora ordenados segundo princípios formais, são chamadas classificações facetadas. Cada faceta com os respectivos elementos constitui uma categoria. A expressão "faceta" foi introduzida na teoria da classificação por Ranganathan e indica que os elementos da descrição de uma classe (por exemplo, do tema de um livro) se compõem de vários elementos da classificação com os quais, de acordo com regras próprias de cada disciplina (fórmulas das facetas), podem constituir um tema [...] No entanto a Colon Classification de Ranganathan mostra a possibilidade da construção de uma classificação facetada universal, lista apresenta a vantagem de, com apenas alguns elementos, tornar possíveis inúmeras combinações. Assim como expressar novos assuntos com conceitos já existentes, como aliás acontece com as linguagens naturais (Dahlberg, 1978, p. 9).

A ideia de uma classificação facetada, desenvolvida por Ranganathan (1967), é, sem dúvida, uma abordagem inovadora que procura proporcionar maior flexibilidade e adaptabilidade nos sistemas de classificação. A estrutura classificatória de Ranganathan se adequa ao léxico de uma língua ao respeitar a cosmovisão dos falantes, garantindo que a organização do conhecimento seja tanto lógica quanto culturalmente relevante.

O método analítico – sintético, proposto por Shiyali Ranganathan, é uma abordagem para a classificação e organização do conhecimento que combina duas etapas principais: a análise e a síntese. O método oferece direcionamentos e sinalizações significativas sobre o mapeamento dos elementos singulares e os componentes com propensão a formar agrupamentos.

Na etapa de Análise, os conceitos são decompostos em suas partes constitutivas. Isso envolve a identificação dos diferentes aspectos ou facetas de um assunto. A análise permite que os classificadores compreendam a complexidade de um conceito, separando suas características e suas relações.

Na etapa de síntese, após a análise, há a recomposição dos elementos analisados em uma nova estrutura que reflita a inter-relação entre eles. Essa etapa busca organizar os conceitos de forma que eles possam ser facilmente acessados e compreendidos. A síntese permite a criação de uma classificação que não apenas lista os conceitos, mas também mostra como eles se conectam e se relacionam entre si.

A classificação das palavras em Guarani-Mbyá e o agrupamento das suas características carecem de uma proposta que contemple, ao mesmo tempo, a diversidade e a pluralidade das definições: os termos *anguja USE guaki*, têm o sentido 'rato'[7] , na língua portuguesa (Ivo *et al*., 2024, p. 93). No entanto, esse sentido é ancorado pelo componente empírico contextual. Enquanto *anguja* é rato não comestível, *guaki* o é. Além de inovar em conteúdo semântico, traz o desafio em como representá-lo, em que o componente pragmático lhes impõe implicações distintas, conforme se ilustra no quadro 1:

Quadro 1 - Características semânticas e pragmáticas incorporadas ao Prombyá

Léxico	Anguja	Guaki
Significado (aspecto semântico)	Rato	Rato
Aplicação do léxico em um contexto (aspecto pragmático)	Não comestível	Comestível

Fonte: Ivo *et al.* (2024), adaptado pelas autoras.

[7] Doravante, os sentidos em língua portuguesa serão apresentados apenas com aspas simples, logo após o termo na língua indígena.

Além disso, há a diversidade regional (variação linguística) nas palavras: *Ovy USE Hovy* 'verde ou azul', conforme apresenta Ivo *et al*. (2024, p. 144), e criação de novas expressões que definem os velhos ou novos objetos ao seu redor, como *Ayvu rapyta* 'celular' (fonte de origem da fala), apresenta Ivo *et al*. (2024).

Pelos exemplos, anteriormente elencados conclui-se a necessidade de uma teoria dinâmica que contemple as situações acima descritas. Nesse sentido, percebeu-se que a Teoria da Classificação facetada – TCF, apresentada por Ranganathan (1967) e o seu método analítico - sintético dialogam com o Prombyá no sentido de promoverem a organização da informação indígena.

No decorrer da pesquisa, foi observado que, as diferenças culturais fazem com que algumas expressões indígenas possam apresentadas para o português e outras não. Existem elementos, na cultura indígena, sem o correlato na língua portuguesa. Nesse sentido, para se equalizar as diferenças mencionadas, cabe ao sistema abordar e esclarecer essas assimetrias, por meio de atributos específicos. A esse tipo de informação, o protótipo emprega as notas de escopo e as notas explicativas.

Uma disfunção pode ser originada a partir de distorções (principalmente quando se busca representar uma cultura indígena espelhando-se na cultura ocidental). Esse processo dialoga com o outro aspecto detectado, no decorrer do estudo, em muitas matérias e instrumentos digitais, em que se evidenciava a falta de uma perspectiva indígena como base do processo de organização das informações. Para suprir essa lacuna, o Prombyá emprega a "[...]garantia literária[...] (Barité *et al*., 2010, p.3)" por meio da qual se pode ter o respaldo das escolhas e das palavras presentes na Organização do Conhecimento (OC). Como exemplo, cita-se o tesauro (Littletree; Metoyer, 2015), cuja proposta é servir de instrumento para atualização dos cabeçalhos de assuntos da Biblioteca do

Congresso Nacional, aceito pelos indígenas, uma vez que, conta com sua contribuição e igualmente aceito pelos bibliotecários.

Em termos de representações indígenas, tem-se ainda o viés apresentado por Olson (1999), em que as expressões indígenas, bem como os diversos grupos não hegemônicos, são reduzidos e adaptados com o intuito de serem igualados aos conhecimentos de uma cultura de origem indo-europeia. O problema é que, ao se aleijar dessa forma o saber, promove-se sua alteração em que a correção se torna difícil e, uma vez que, pode dar vida às disfunções sobre o povo e a sua cultura. Chegando a promover reflexos na recuperação da informação.

Em conjunto com a Teoria da Classificação Facetada, apresentada por Ranganathan (1967), destaca-se o diálogo com a Linguística, que oferece os fundamentos formais para a estruturação dos processos comunicativos indígenas. Nesse contexto, insere-se a Teoria Comunicativa Terminológica (TCT), apresentada por Maria Teresa Cabré (1998), amplamente difundida na Espanha, e cujas bases semasiológicas auxiliam no tratamento de léxicos.

A perspectiva de Cabré (1999) está ancorada em um arcabouço teórico, prático e metodológico. Em que a perspectiva do fortalecimento de línguas minoritárias se alinha a recepcionalidade da diversidade, acolhendo os elementos empíricos do contexto de fala e a cientificidade necessária na organização e representação dos seus elementos comunicativos. A autora apresenta o conceito como elemento mais complexo, fruto de uma “construção mental” (Cabré, 1998, p.42). O termo é o elemento mais simples que se comporta como uma propriedade do conceito, elucida Cabré (2008). Entre os princípios apresentados pela autora, encontra-se o caráter poliédrico do termo que dá origem a Teoria das Portas (Cabré, 1998). Para ancorar a dimensão tridimensional que este elemento assume em sua abordagem, são apresentadas as dimensões

linguísticas, cognitivas e sociais, cuja importância das características contextuais e semântica identificam a comunidade representada, presentes desde a concepção dos termos. Cabré (1998) apresenta as três portas:

> A Teoria Comunicativa Terminológica - TCT surge na Espanha, tendo como expoente Maria Teresa Cabré. A perspectiva de Cabré (1999) está ancorada em um arcabouço teórico, prático e metodológico na perspectiva de fortalecimento de línguas minoritárias, recepcionando a diversidade, aliando elementos empíricos do contexto de fala e a cientificidade necessária na organização e representação dos seus elementos terminológicos. A autora apresenta o conceito como elemento mais complexo, fruto de uma "construção mental" (Cabré, 1998, p.42).

Para Cabré (1998), a compreensão sobre a diferença entre palavras e termos é tênue. É possível identificar semelhanças entre palavras e termos, uma vez que, se destinam a "[...]designar objetos do mundo real", informa Cabré (1998, p. 80 - 81). No entanto, ao considerar as unidades pragmáticas e comunicativas, as palavras e os termos diferem com base na perspectiva funcional observadas em cada unidade informacional. A seguir, apresenta-se o processo de construção metodológica na elaboração do Prombyá.

4 DESENVOLVIMENTO METODOLÓGICO

Nessa proposta, os diversos elementos metodológicos são empregados de forma integrada, visando uma compreensão holística da temática. Assim, o caminho metodológico adotado inclui diálogos interdisciplinares, envolvendo áreas como a Ciência da Informação, a Linguística, a Lexicografia e as questões antropológicas. Essa abordagem multidisciplinar enriqueceu a análise e a compreensão da complexidade da língua e cultura Guarani-Mbyá.

A modelagem navegacional proposta, fruto desse processo de análise, caracteriza a pesquisa como uma investigação aplicada, uma vez que visa

preencher as lacunas, sobre a representação da informação, em pesquisa cuja temática abordam as línguas indígenas.

Os aspectos subjetivos, semânticos e contextuais da comunidade de fala, norteiam os aspectos sócio informacionais importantes que conduzem a pesquisa por um viés qualitativo, conforme ensina Minayo (2009).

Quanto aos procedimentos, é uma pesquisa documental, ancorada no conteúdo lexicográfico Guarani-Mbyá, presente no Dicionário Bilíngue Guarani-Mbyá/Português (Ivo *et al*., 2024). É uma pesquisa, cuja ancoragem bibliográfica, apresenta elementos importantes na busca pela compreensão das características socioculturais e linguísticas do povo Guarani-Mbyá.

Quanto aos objetivos, é uma pesquisa descritiva, em que a busca-se aprofundar o entendimento sobre as características dos elementos lexicográficos, suas relações, sua ordenação e, por fim, sua classificação, conforme elucida Triviños (1987), aplicando a interdisciplinaridade entre a Ciência da Informação, por meio do legado apresentado por Ranganathan (1967), a Linguística, por meio da TCT (Cabré, 1998) e a Linguística para línguas indígenas, conforme apresenta Ivo *et al*. (2024) e Ivo(2018).

O corpus da pesquisa foi extraído do Dicionário Bilíngue Guarani-Mbyá/Português, construído com metodologias de pesquisa colaborativa. Esse dicionário apresenta uma representatividade lexical abrangente, pois contempla as informações coletadas em diversas regiões do Brasil, onde os Guarani-Mbyá[8] estão localizados, incluindo os estados do Rio de Janeiro, São Paulo, Paraná, Mato Grosso do Sul e Rio Grande do Sul. A escolha da fonte documental baseou-se nas

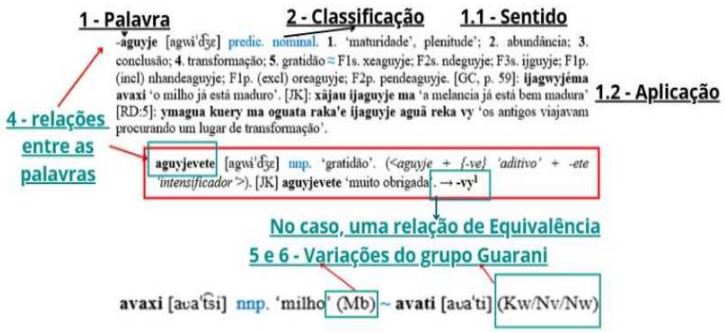
[8] Nesse trabalho, são adotadas as normas da Associação Brasileira de Antropologia – ABA (1957), que dis-pensam a flexão de gênero e número nos etnônimos. Além disso, os nomes das etnias são grafados com a inicial maiúscula.

premissas, previamente mencionadas, para a concepção da modelagem proposta. Nessa fase, a proposta empenhava-se em atingir uma classificação que representasse da forma mais fidedigna possível a cultura Guarani-Mbyá, apoiando-se na literatura especializada da área, assim como, na garantia literária.

O processo de mapeamento tem início com a escolha do léxico. A definição sobre a quantidade de elementos lexicais adotou um critério empírico, uma vez que o objetivo seria identificar os atributos suficientes para respaldar um sistema de organização de informações.

Definida a amostra empregada para a modelagem, o próximo passo seria partir para um modelo análise das informações significativas. Após a identificação das informações-chave, conforme observa-se na figura 1, cumpria se isolar aquelas úteis ao SOC, conforme cada palavra utilizada. Nesse momento da pesquisa, a construção assemelha-se a uma concepção linear, no entanto, para um melhor entendimento, nesta fase, os pontos-chave (classes, facetas, atributos, relações entre outros) serão revisitados com o intuito de encontrar uma organização que facilite a encontrabilidade e a representatividade da informação (Figura 1).

Figura 1 - Entrada do Dicionário Bilíngue Guarani - Mbyá/Português[9]

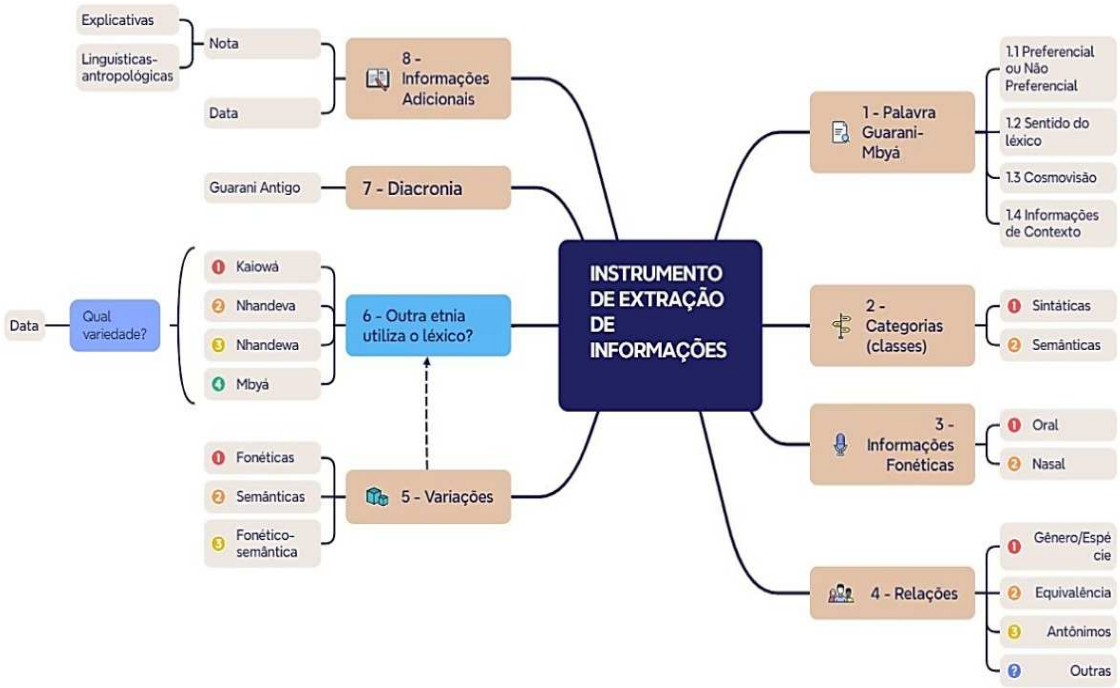


Fonte: Das autoras, baseado em Ivo *et al.* (2024).

[9] Na entrada, JK: se refere a Joel Kuaray (professor e falante Guarani-Mbyá). A sigla MB: se refere aos po-vos Guarani-Mbyá. Kw: se refere aos povos Guarani-Kaiowá. Nv: se refere aos povos Guarani-Nhandeva e Nw: se refere aos povos Guarani-Nhandewa, conforme apresenta Ivo et al. (2024).

Para a extração e análise dos dados, foi desenvolvido um instrumento com a finalidade de evidenciar as correlações entre as variáveis e as suas características lexicográficas, entre os quais pode-se citar : atributos que auxiliam no entendimento da palavra ou das expressões (1), categorias sintáticas e semânticas (2), informações fonéticas (3), relações entre as palavras (4), variações linguísticas (5), variedades dialetais pelos grupos guaranis (6), informações sobre variação conforme o tempo – diacronia (7) e, por fim, informações adicionais (8), como se pode observar na Figura 2:

Figura 2 - Instrumento de extração de informações



Fonte: Elaborado pelas autoras.

A etapa computacional da pesquisa, compreende a fase aplicada, que envolveu a utilização de um ambiente de programação específico para a elaboração do protótipo responsável pela gestão e análise dos dados coletados, e a configuração do ambiente de programação envolveu as seguintes tecnologias:

[a] Instalação do WAMP: pacote para Windows que contém (Apache, Mysql e PHP) e fornece um servidor local, responsável pela interpretação dos códigos na linguagem de programação PHP (*Hypertext Preprocessor* – Pré-Processador de Hipertexto). O servidor MySQL e Apache (é o ambiente que interpretará os scripts, entregando o conteúdo ao usuário, seja na forma de armazenamento de informações no Banco de Dados – MYSQL–, seja na forma de um Sistema de Recuperação da Informação, buscando alguma informação no Banco de Dados – MYSQL);

[b]PHP: é uma linguagem de programação de código aberto. Originalmente criada para a construção de páginas dinâmicas, ela roda no servidor e é embutida no HTML, permitindo a criação de sites interativos e aplicações web.

[c] Banco de Dados – MYSQL: é um sistema de gerenciamento de banco de dados relacional (SGBD) de código aberto, amplamente utilizado para armazenar, gerenciar e recuperar dados de forma eficiente. Ele utiliza a linguagem SQL (*Structured Query Language*) para manipulação de dados e é conhecido por sua rapidez, escalabilidade e confiabilidade;

[d] Desenvolvimento de Scripts PHP: os Scripts PHP foram desenvolvidos para facilitar a gestão, a consistência e a análise dos dados, empregando o *Visual Studio Code* como IDE – Ambiente de Desenvolvimento Integrado;

[e] Integração com *Bootstrap*: integrado aos *scripts* PHP, para criar interfaces mais amigáveis ao usuário.

Na próxima seção, os resultados e as discussões sobre as estruturas classificatórias desenvolvidas para a proposta de modelagem navegacional para organização da língua Guarani-Mbyá, ainda em desenvolvimento.

5 RESULTADOS E DISCUSSÃO

A diversidade dos povos e línguas indígenas do Brasil é um fator desafiante. É um fato que os pesquisadores do domínio da OC precisam compreender e empreender formas éticas para lidar com a temática. E, para tanto, faz-se necessário o desenvolvimento de propostas que envolvam a consideração de aspectos socioculturais como base, a cosmovisão e a sociedade indígena como pilares fundamentais. Assim, apresenta-se a estrutura classificatória desenvolvida para a modelagem navegacional do corpus de análise desta pesquisa, para aplicação futura do Prombyá.

A diversidade lexical, muitas vezes, reflete as diferentes perspectivas e abordagens representativas e identitárias na sociedade atual. A normatização de um único conjunto de palavras pode levar à exclusão de vozes não hegemônicas ou à supressão de conhecimentos locais e tradicionais. O corpus Guarani-Mbyá que embasou o Prombyá apresentou alguns elementos que demandaram uma teoria capaz de fornecer os subsídios para o tratamento das seguintes situações: [1] variações linguísticas; [2] homonímias; [3] polissemias, dentre outras situações que inspiraram o diálogo interdisciplinar entre a CI, a Linguística e a Antropologia.

Para se alcançar a proposta da modelagem, percebeu-se que um primeiro movimento seria a classificação. Esse procedimento, além de auxiliar na organização da informação, também contribuiria para a gestão dos elementos que apresentavam algum grau de polissemia em sua constituição.

Surge, então, a necessidade em se conceituar elemento polissêmico. O léxico polissêmico é aquele que apresenta mais de um significado para uma

mesma palavra. E como exemplo, apresenta-se os vários significados do léxico manga, conforme o Michaelis (2015c):

[1]fruta comestível da mangueira;

[2] parte do vestuário que cobre o braço, cingindo-o, parcial ou totalmente;

[3]objeto tubular munido de dois bocais, geralmente de vidro, que protege a luz nos castiçais ou nos braços de lustres;

[4]tubo plástico, de dimensões variáveis, usado para envolver condutores elétricos;

[5]Pequeno cilindro de tecido não inflamável que envolve a chama de lamparina ou de bico de gás, a fim de torná-la mais brilhante;

A polissemia é um desafio, pois exige métodos para diferenciar os significados com base no contexto em que as palavras são empregadas. No tratamento do léxico indígena polissêmico, concluiu-se que um único elemento não seria suficiente para garantir a diferenciação necessária. Para lidar com a polissemia, adotou-se uma técnica, originária da tecnologia da informação, que consiste na organização de múltiplas colunas de uma tabela, existente no banco de dados relacional, formando assim, um único campo com conteúdo equivalente à combinação das colunas selecionadas. Nesse caso, as colunas escolhidas para compor a tupla foram: palavra, classe sintática e categoria semântica. O escopo da pesquisa demonstrou que esses três atributos são suficientes para garantir a diferenciação necessária.

No Guarani-Mbyá, as palavras podem variar semanticamente, em diferentes classes gramaticais, as quais fornecem um nível de especificidade suficiente para auxiliar no tratamento da polissemia. No entanto, para compreensão da estrutura, com base no instrumento de extrações de informações, ressalta-se que o léxico foi selecionado conforme as seguintes

características: [a] palavras com sentido básico ou singular; [b] palavras polissêmicas; [c] palavras com indícios de relações de equivalência, gênero/espécie, entre outras relações significativas; [d] palavras que demonstram os aspectos da cultura e, que contemplem os elementos da cosmovisão indígena; [e] palavras que indiquem aspectos de variação e/ou regionalismos. No Quadro 2, abaixo, a aplicação do instrumento de extração (apresentado na Figura 2 acima) no termo porã :

Quadro 2 - Aplicação do instrumento de extração de informações para a entrada porã

FUNÇÃO
1. **Palavra em Guarani-Mbyá:** *Porã* **2.** **Palavra Preferida ou Palavra Candidata**: Palavra Preferida; **3.** **Sentido descritivo:** Bom e bonito; **4.** **Sentido na Cosmovisão indígena:** Sagrado; **5.** **Contexto de aplicação da palavra**: Exemplo1: xeko'ẽ porã 'amanheci bem' Exemplo2: tape porã 'caminho sagrado'
6. **Classes** **7.** **Classe Sintática**: Nome Qualitativo **8.** **Categoria Semântica:** Qualidade
9. **Variação diacrônica**: Porang Montoya (2011 [1639], p. 248])
10. **Relações entre os léxicos**: Equivalência: Porã USE Porang
11. **Variações entre parcialidades:** Não há variações.
12. **Informações Adicionais**: **13.** **Fonte:** Montoya (2011 [1639], p. 248]); Ivo (2018), Ivo et al. (2024, p.147). **14.** **Data:** 2024

Fonte: Ivo *et al.*(2024, p.147), adaptado pelas autoras (2024).

Para elucidar como as informações se relacionam, criou-se uma estruturação de classe e seus atributos, bem como, a relação entre as classes, evidenciando como as informações dialogam em uma modelagem. Entre os exemplos de classe, pode-se citar: [a] Rel – Classe Relações: o objetivo é identificar as relações entre as palavras. Para tanto, há que se registrar as

expressões que se relacionam e o tipo da relação (gênero/espécie, sinonímias, entre outras); [b] Var – Classe Variações: para um único sentido é possível identificar duas palavras. Exemplos de variações que ocorrem rotineiramente, podem-se citar os casos em que se observam a diacronia[10] e a sincronia[11]: alterações que ocorrem nas expressões indígenas, o que, sob uma perspectiva sociocultural, demanda o registro de ambas. [c] Classe Etnias – as parcialidades do grupo Guarani, por vezes, dividem o mesmo território o que confere contato lexical frequente e, até mesmo compartilham das mesmas necessidades informacionais; [d] Classe Sintática – agrupamento que reúne as informações sobre os elementos formais da língua; [e] Categoria Semântica – agrupamento que contempla as informações sobre a forma como o povo compreende o mundo a sua volta.

Como complemento, a integração com os elementos funcionais, identificados e possíveis de serem implementadas em um Sistema de Recuperação da Informação (SRI), os métodos a seguir, foram divididos conforme o seu aproveitamento nos vários ambientes:

[1] métodos reutilizáveis: cadastrar, listar, atualizar e apagar.

[2] métodos únicos: algumas funcionalidades acontecem em apenas uma estrutura. É o caso do controle de polissemia que surge apenas nas inserções de novas palavras ou expressões, por exemplo, e o módulo busca, necessário em um ambiente navegacional. Após estas construções, a classe sintática tem origem na

[10] "Termo adotado por Ferdinand de Saussure (1857–1913) para a transmissão de uma língua através do tempo e das gerações, sofrendo, nesse transcurso, mudanças fonéticas, mórficas, sintáticas, semânticas e léxicas. A diacronia define o caráter dos fatos linguísticos considerados em sua evolução ao longo do tempo" (Diacronia, 2015a).

[11] "Estado de uma língua num determinado momento, sem levar em conta sua evolução histórica" (Sincronia, 2015b).

pesquisa documental do corpus desta pesquisa, o Dicionário Guarani-Mbyá/Português.

O corpus, atualmente, apresenta uma classificação em bases lexicais, composta por elementos como verbos, nomes, pronomes, marcadores morfológicos, entre outros. Contudo, optou-se por uma classificação adicional, baseada em categorias básicas (classes semânticas), o que resultou no uso das categorias como flora, fauna, alimentação, partes do corpo, relações de parentesco etc. Percebe-se que esta ação universaliza o conhecimento para outras áreas, assim como confere funcionalidades adicionais, a exemplo da recuperação da informação e da navegação facetada.

Destaca-se que, nesta fase da pesquisa, estão sendo realizadas estratégias para o SRI, por exemplo, o cadastramento de expressões polissêmicas, visando contribuir com a consistência na gestão desse léxico, bem como a compreensão e a representação adequada da diversidade de significados na língua. Outro fator considerado é a variação e a mudança linguísticas, reconhecendo-se a dinamicidade da língua Guarani-Mbyá ao longo do tempo, incluindo-se o cadastro da diacronia e sincronia, considerando-se tanto a perspectiva atual quanto histórica da língua, conforme apresentado no script de gestão dos dados (Figura 3). As soluções propostas para aprimoramento do Prombyá, que além das citadas, também considerou o cadastramento de palavras correlacionadas, como relações de equivalência e classes, contribuindo para a organização e compreensão do léxico da língua Guarani-Mbyá e fornecendo diretrizes para a representação eficaz e abrangente da língua indígena.

Figura 3 - Proposta de Organização das Informações

+Novo Termo

Letra	Termo	Semântica	Aplicação do Termo	Classe Sintática	Categoria Semântica	Tipo do Termo	Relações	Termo Correlato	Etnia	Variação do Termo conforme Etnia	Guarani Antigo	Nota	Data
A	-aguyje	1. 'maturidade', plenitude'; 2. abundância; 3. conclusão; 4.	[JK]: xãjau ijaguyje ma 'a mel	Nome possuível intransferível	parentesco	preferencial	genero/especie	Aguyjevete				Fonte: Ivo et al (2024).	2024-06-13
A	Aguyjevete	Gratidão	[JK] aguyjevete 'muito obrigad	Nomes não possuíveis	comprimento_diario	preferencial	equivalência	javy ju 'nos levant	Guarani-Mbyá			Fonte: Ivo et al (2024).	2024-06-13
A	-a	cair	ha'a 'eu caí'	Verbos	não_aplicável	preferencial	----------		Guarani-Mbyá			Fonte: Ivo et al. (2024)	2024-06-13
A	-a	fruta	narã'a 'fruta da laranjeira'	Nome possuível intransferível	natureza	preferencial	----------		Guarani-Mbyá			Fonte: Ivo et al. (2024)	2024-06-13

Fonte: Elaborado pelas autoras.

Na próxima seção, apresentam-se as considerações finais sobre a pesquisa, em desenvolvimento.

6 CONSIDERAÇÕES FINAIS

A modelagem navegacional para a língua Guarani-Mbyá, aqui apresentada, contribui para fornecer uma estrutura que permite a realização da gestão do léxico, contemplando os requisitos de consistência, coerência e preservação da essência Guarani-Mbyá.

Por meio da proposta, destaca-se a relevância da organização do conhecimento para a valorização, revitalização e difusão da língua Guarani-Mbyá, evidenciando e alcançando, assim, por meio da linguagem científica, outras perspectivas e horizontes para uma compreensão mais abrangente e sensível das necessidades informacionais e culturais das comunidades Guarani-Mbyá.

A abordagem interdisciplinar adotada, em particular, a OC como campo teórico-metodológico, em diálogo com a Linguística, teve um papel central no

desenvolvimento da estrutura classificatória, pois, ajudou a compreender os elementos essenciais da Língua Guarani-Mbyá para que se pudesse identificar, definir e organizar corretamente as expressões presentes da língua indígena, destacando-se os variados aspectos e perspectivas de usos e significados entre a comunidade Guarani-Mbyá e contribuindo para uma categorização mais representativa e detalhada. A integração da abordagem antropológica enriqueceu ainda mais a estrutura classificatória, uma vez que considera a cultura, valores e perspectivas dos membros da comunidade indígena, com destaque para o aspecto etnográfico, observando-se a garantia literária em termos de representação da língua e cultura Guarani Mbyá.

A estrutura classificatória aqui apresentada, como parte dos resultados da pesquisa qualificada, categoriza o léxico indígena semântica e sintaticamente, identifica relações entre as palavras e trata ambiguidades e sinonímias, contribuindo, assim, para a manutenção e valorização do patrimônio linguístico e cultural indígena. Os resultados obtidos com a modelagem da língua Guarani-Mbyá mostram a viabilidade e a relevância da pesquisa, fornecendo um modelo conceitual para o desenvolvimento de um SRI no futuro.

Destaca-se a importância da colaboração com a comunidade Guarani-Mbyá, especialmente na construção do Dicionário Bilíngue Guarani-Mbyá/Português, corpus de análise no presente estudo que, ao considerar a cultura e práticas linguísticas no convívio e escuta dos falantes nativos, possibilitou uma representação mais próxima possível da língua indígena. A pesquisa, enquanto trabalho científico em nível de pós-graduação, aponta para a continuidade dos estudos sobre a organização do conhecimento indígena, visando promover a inclusão e valorização das línguas e culturas indígenas no contexto acadêmico e social. No âmbito da inovação social, por meio da

aplicação do Prombyá, espera-se promover a vitalização, aprendizado e análise da língua e cultura Guarani-Mbyá, atendendo às necessidades de diversos públicos, a saber, membros da comunidade indígena, falantes bilíngues, pesquisadores e sociedade em geral.

REFERÊNCIAS

ABA/Associação Brasileira de Antropologia. *In*: REUNIÃO BRASILEIRA DE ANTROPOLOGIA, 2., 1955, Salvador. **Anais** [...] Salvador: S. A. Artes Gráficas, 1957. Disponível em: http://www.aba.abant.org.br/conteudo/ANAIS/conteudo-967268. Acesso em: 09 jan. 2025.

ALMEIDA, Rita Heloísa. **O diretório dos índios**: um projeto de civilização no Brasil do século XVIII. Brasília: Editora Universidade de Brasília – UnB, 1997.

BARITÉ, Mário; FERNÁNDEZ-MOLINA, Juan Carlos; GUIMARÃES, José Augusto Chaves; MORAES, João Batista Ernesto de. Garantia literária: elementos para uma revisão crítica após um século. **Transinformação**, Campinas, v. 22, n. 2, p. 123–138, 2010. Disponível em: https://www.scielo.br/j/tinf/a/prtXbCcMkLD48hdnHR4tcfS/. Acesso em: 08 jan. 2025.

CABRÉ, Maria Teresa C. El princípio de poliedricidad: la articulación de lo discursivo, lo cognitivo y lo lingüístico en Terminología. **Revista de la Asociación Europea de Lenguas para Fines Específicos**, p. 9-36, 2008. Disponível em: https://www.redalyc.org/articulo.oa?id=287024065003. Acesso em: 08 jan. 2025.

CABRÉ, Maria Teresa C. Terminología Y documentación. *In*: CABRÉ, Maria Teresa C. (org.). **La terminologia:** representación y comunicación: elementos para una teoría de base comunicativa y otros artículos. 1. ed. Barcelona: ULA/Universitat Pompeu Fabra, 1999.

CABRÉ, Maria Teresa C. **Terminology:** theory, methods, and applications. Países Baixos/Amsterdã: John Benjamins Publishing Company, 1998.

DAHLBERG, Ingetraut. Fundamentos teórico-conceituais da classificação. **Revista de Biblioteconomia de Brasília**, Brasília, v. 6, n. 1, p. 9-21, 1978. Disponível em: https://periodicos.unb.br/index.php/rbbsb/article/view/29057. Acesso em: 07 jan. 2025.

DIACRONIA. *In:* MICHAELIS **—** Dicionário Brasileiro da Língua Portuguesa, [*s. l.*], 2015a. Disponível em: https://michaelis.uol.com.br/moderno-portugues/busca/portugues-brasileiro/diacronia/. Acesso em: 06 jan. 2025.

IVO, Ivana Pereira. **Características fonéticas e fonologia do Guarani no Brasil**. 2018. Tese (Doutorado em Linguística) — Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 2018. Disponível em: https://hdl.handle.net/20.500.12733/1636701. DOI: https://doi.org/10.47749/T/UNICAMP.2018.1090892. Acesso em: 30 jun. 2022.

IVO, Ivana Pereira. **Dicionário Bilíngue Guarani-Mbyá – Português/Português – Guarani-Mbyá**, 2024. Colaboração: COSTA, Julia Lima, RIOS, Jonedson Costa, REZENDE, Karen Soledade. Consultores Indígenas: Joel Kuaray, Simone Takuá, Sara Katu, Valério Karaí, Iraci Nunes. Organização e elaboração da versão eletrônica: Micheline Maria Costa de Azevedo (em fase de elaboração).

LIMA, Gercina Ângela de. Organização e representação do conhecimento e da informação na web: teorias e técnicas. **Perspectivas em Ciência da Informação**, Campinas, v. 25 (Especial), p. 57-97, 2020. Disponível em: https://periodicos.ufmg.br/index.php/pci/article/view/22283. Acesso em: 06 jan. 2025.

LITAIFF, Aldo. **As divinas palavras**: Identidade étnica dos Guarani-Mbyá. Florianópolis: Editora da UFSC, 1996.

LITTLETREE, Sandra; METOYER, Cheryl A. Knowledge Organization from an Indigenous Perspective: The Mashantucket Pequot Thesaurus of American Indian Terminology Project. **Cataloging and Classification,** Quarterly, p. 640-657, 2015. Disponível em: DOI: 10.1080/ 01639374.2015.1010113. Acesso em: 16 jan. 2025.

MACHADO, Jorge; FERREIRA, Carlos H. (Orgs.). **Resistência Guarani** — uma Vivência na Aldeia Rio Silveiras. São Paulo: Tendenz, 2016.

MELIÁ, B. La lengua Guaraní del Paraguay. *In*: POTTIER, Bernard (coord.). **America Latina en sus Lenguas Indigenas**. Madrid: Editorial Mapfre, 1983. p.43-59.

MINAYO, Maria Cecília de Souza. O desafio da Pesquisa Social. *In*: MINAYO, Maria Cecília de Souza. (Orgs.) **Pesquisa Social** - Teoria, método e criatividade. 2. ed. Petrópolis, RJ: Vozes, 2009. p. 9-29.

OLSON, Hope. Cultural Discourses of Classification: Indigenous Alternatives to the Tradition of Aristotle, Durkheim and Foucault. **Classification Research Workshop**, 1999. Disponível em: https://journals.lib.washington.edu/index.php/acro/article/view/12484/11022. DOI: 10.7152/acro.v10i1.12484. Acesso em: 13 jan. 2025.

POMBO, Olga. Interdisciplinaridade e integração dos saberes. **Liinc em Revista**, v. 1, n. 1, p. 1–13, 2006. Disponível em: https://revista.ibict.br/liinc/article/view/3082. DOI: 10.18617/liinc. v1i1.186. Acesso em: 13 jan. 2025.

RANGANATHAN, Shiyali Ramamrita. **Prolegomena to library classification**. Mumbai: Asia Publishing House, 1967. 305 p.

RODRIGUES, Aryon. D. **Fonética Histórica Tupi Guarani**. Diferenças fonéticas entre o Tupi e o Guarani. Curitiba: Arquivos do Museu Paranaense. Coleção Aryon Rodrigues, 1945. v. 4. p. 333-354.

SAUSSURE, Ferdinand. **Curso de Linguística Geral**. Tradução de Antônio Chelini, José Paulo Paes e Izidoro Blikstein. São Paulo: Cultrix, 1995.

SINCRONIA. *In:* MICHAELIS — Dicionário Brasileiro da Língua Portuguesa, [*s. l.*], 2015b. Disponível em: https://michaelis.uol.com.br/moderno-portugues/busca/portugues-brasileiro/sincronia/. Acesso em: 06 jan. 2025.

TRIVIÑOS, Augusto Nibaldo Silva. **Introdução à pesquisa em ciências sociais**: a pesquisa qualitativa em educação. São Paulo: Editora Atlas S.A., 1987.

Copyright: Esta obra está licenciada com uma Licença Creative Commons Atribuição 4.0 Internacional.

tpbci@ancib.org @anciboficial