

O caráter do pensar autóctone: construção de uma base linguística tecnológica semasiológica e onomasiológica para termos indígenas

The character of indigenous thinking: building a semasiological onomasiological and technological basis for indigenous linguistic terms

Águida Aparecida Gava

Universidade Federal do Mato Grosso – Cuiabá – Mato Grosso – Brasil

Fabricio Gava de Almeida Jorge

Universidade de São Paulo – São Paulo – São Paulo – Brasil



Resumo: Esta pesquisa objetiva a construção de uma ferramenta eletrônica que salvaguarde o conhecimento de línguas maternas indígenas brasileiras. Para tal, desenvolveremos uma base tecnológica, caracterizada por um dicionário eletrônico multilíngue, com desempenho semasiológico e onomasiológico para a inserção e manipulação online dos termos, sem que, contudo, cerceie a quantidade de línguas inseridas, pressupondo divulgar os termos da fauna brasileira, compreendendo os valores linguísticos e socioculturais a partir de características semânticas, previamente estabelecidas, de acordo com o pensar e o sentir autóctone. Devido às dificuldades geográficas, nos ateremos às línguas indígenas do estado do Mato Grosso. Com o propósito de viabilizarmos tais diligências, este projeto foi apresentado à Universidade Federal do Mato Grosso-Cuiabá, UFMT e ao CNPq. Baseamo-nos nos autores: Gava (2012), e em seu modelo de dicionários eletrônico terminológico multilíngue, organizado para seis línguas indígenas; Pottier (1985, 1987, 1992); Barbosa(2006); Fargetti (2012); Babini (2001, 2006); Cabré (1999) entre outros.

Palavras-chave: Dicionário indígena; Dicionário eletrônico multilíngue; Língua indígena; Características semânticas; Etnoterminologia

Abstract: This project aims to develop a computational system to insert a multilingual semasiological and onomasiological electronic dictionary in order to publicize and archive indigenous terms for Brazilian fauna, documenting sociocultural and linguistic semantic values. Our online database will enable the inclusion of several languages, without any limitations on the number of languages or terms, unlike previously established computational systems. Due to geographical difficulties of communication and dissemination of our work among Brazilian indigenous languages speakers, we will focus our project on the indigenous languages of the state of Mato Grosso. This project was presented to the Federal University of Mato Grosso- Cuiaba (UFMT) as well as CNPq. Our project is based in Gava (2012), in which we develop a model of dictionaries organized in six indigenous languages, and Pottier (1985, 1987, 1992), Babini (2001, 2006), Caber (1999), Fargetti (2012), Faulstich (2007) and Silva (2009).

Keywords: Indigenous dictionary; multilingual electronic dictionary; indigenous language; semantic features; etnoterminology

Introdução

Nossa pesquisa tem como objetivo produzir um software que gere um dicionário interativo, *online*, com funções semasiológica e onomasiológica, tendo como ponto de partida termos da fauna brasileira e traços semânticos de línguas indígenas mato-grossense, o qual contemple a capacidade de abarcar a grande diversidade linguística indígena brasileira, direcionada à fauna, bem como estabelecer uma gestão eficiente dos termos inseridos.

Nesse sentido, desejamos proporcionar um intercâmbio linguístico e de conhecimento entre os falantes indígenas em suas línguas maternas e identificar o pensar e o sentir autóctone, arraigados nas questões socioculturais e etnoterminológicas.

A relevância de nosso trabalho está intrínseca à importância de todo desenvolvimento voltado a salvaguardar o patrimônio cultural indígena no estado do Mato Grosso. E permanecerá acessível para novas inserções de termos indígenas por todo território brasileiro, obedecendo o recorte da fauna brasileira. Outrossim, nosso projeto mostra-se impactante em dois eixos principais: no primeiro, a pesquisa contribuirá significativamente para o desenvolvimento científico e tecnológico da área de linguística por avançar no desenvolvimento de uma base que promova a construção de dicionários e análises de diferentes línguas e suas inter-relações; em um segundo eixo, contribuirá para a preservação e fortalecimento da cultura autóctone existente no estado do Mato Grosso, evidenciando a língua e a cultura indígena, considerando seus termos e traços semânticos como ponto de partida.

O produto esperado será um conjunto de termos caracterizados semanticamente de acordo com os valores etnoterminológicos dos povos indígenas, implementado em formato digital, com base na *Plataforma Kuhi-Pei*, desenvolvida por Gava (2012). Nesse sentido, procuramos dar um passo significativo na direção de consolidar o modelo *Kuhi-Pei*. Ao mesmo tempo, é um modo de promover o intercâmbio linguístico entre povos indígenas e de divulgar a equivalência de termos entre diferentes idiomas.

O presente estudo justifica-se, portanto, por sua significativa contribuição aos estudos etnoterminológicos indígenas no Brasil. Ademais, ao considerar as diferentes realidades socioculturais existentes no desenho universal linguístico indígena brasileiro, a pesquisa permite subsidiar fortemente o debate evidenciado nos seguintes questionamentos, os quais pretendemos responder ao longo do desenvolvimento do projeto:

1. Qual a relevância do presente dicionário para as comunidades indígenas mato-grossense no

tocante ao subsídio e conservação de culturais basicamente orais?

2. Quais os valores práticos (cotidianos) de tal implementação para o processo educacional indígena?
3. Como ampliar a abrangência do projeto para todo território brasileiro?

O Brasil conta com um variado número de famílias linguísticas e idiomas indígenas, em número superior a 256 línguas distintas. Estenderemos nossa pesquisa apenas para o estado do Mato Grosso, sede da Universidade Federal do Mato Grosso, entidade executora de nosso projeto. Sobretudo idealizando o desenvolvimento do estado.

1 Fundamentação teórica

No tocante à confecção de obras lexicográficas, principalmente de dicionários indígenas, Fargetti (2012, p. 68) defende a elaboração de descrições, dicionários e gramáticas de referência para a maioria das línguas e uma pesquisa linguística aprofundada, assim como projetos de ensino entre seus falantes. Dessa forma, o momento é de atenção à ortografia na elaboração de léxicos, os quais não devem obedecer aos formatos de listas, como são conhecidos a maioria dos dicionários indígenas no país, mas estruturados com aspectos gramaticais, discussão de pronúncias, prosódia, questões de polissemia e homonímia, entre outros, provendo lematizações coerentes com a estrutura da língua.

A autora enfatiza os dicionários em formato de listas como parte de um estágio importante ao conhecimento do léxico das Línguas Indígenas, entretanto o momento é de vencer às dificuldades iniciais e objetivar obras lexicográficas expressivas, as quais documentem a língua, de maneira adequada; e proponham um estudo aprofundado, centrado em questões de como lematizar lexemas de uma língua indígena, focalizando os aspectos prosódicos, que envolvem a marcação ou não de acentos de intensidade, tons, duração e nasalidade. Entrementes, reconhece que o linguista esbarra na problemática de identificar o que é relevante e necessário grafar em línguas onde a ortografia ainda não está consolidada (FARGETTI, 2012, p. 68).

Desse modo, nosso desafio estende-se do desenvolvimento da ferramenta à criação de uma base que realmente possa atender às exigências linguísticas indígenas contemporâneas.

Direcionando a questão para a semasiologia, Biderman (1984) considera os significantes para indagar sobre os significados, ou investigar o fenômeno da significação. Em ordem inversa, a onomasiologia parte da significação ao encontro do significante, ou seja,

da designação linguística dos conceitos ou objetos considerados (BIDERMAN, 1984, p. 43).

De acordo com Bertoldi (1935), a onomasiologia é um modo particular da pesquisa linguística que, partindo de uma determinada ideia, examina as várias maneiras com as quais essa ideia encontrou expressão na palavra e trata dos aspectos ligados ao processo da denominação, que vai do conceito ao signo (BABINI, 2006, p. 38).

Tendo como pressuposto que a Terminologia delimita seu estudo em conjuntos de termos no âmbito de determinada área do conhecimento, Faulstich (2007) conclui que essa ciência descarta as significações que não são pertinentes ao estudo temático planejado, o qual privilegia uma conduta de análise que se apoia na onomasiologia. Seu discurso baseia-se na Teoria Geral da Terminologia (TGT), da qual Wüster (1985) considerou que a terminologia procede de modo onomasiológico tendo como ponto de partida os conceitos (FAULSTICH, 2007, p. 2).

Definido um dicionário como onomasiológico quando a recolha de palavras é organizada em campos conceituais, como ocorre com os dicionários analógicos, Faulstich questiona se, na organização das definições dos dicionários terminológicos, é possível separar onomasiologia de semasiologia, uma vez que a recolha de termos de uma obra terminográfica parte das relações de significação no sistema conceitual (FAULSTICH, 2007, p. 3).

No que diz respeito à estrutura semasiológica, utilizada na maioria dos repertórios lexicográficos e terminográficos, as informações só são obtidas por meio de um vocábulo, que é a palavra ou o termo. Tal método não atende às necessidades do leitor que conhece parcialmente o significado, mas desconhece o significante do termo (BABINI, 2006, p. 1). Como recurso para resolver esse tipo de busca ele vê na ciência da informática possibilidades que permitem a contemplação de um maior número de necessidades do consultante (BABINI, 2001).

Considerando que a maior parte dos dicionários trazem suas macroestruturas semasiológicas, com o advento da informática, a onomasiologia emerge permitindo, em alguns casos, uma busca mais eficiente.

Em Silva (2009), encontramos a observação de que a onomasiologia ainda carece de produções léxico-terminográficas. Segundo ele, a tradição semasiológica continua a ser mais utilizada e a orientação onomasiológica a ser preterida (SILVA, 2009, p. 25). O mesmo defende Boulanger (2001), dizendo que a lexicografia privilegia uma conduta de análise apoiada na semasiologia (BOULANGER, 2001, p. 18).

Entretanto, a teoria dos campos semasiológico e onomasiológico, concomitantemente, ganha terreno

entre os modelos propostos para análise da estruturação do significado (BIDERMAN, 2001, p. 199).

Nessa perspectiva, consideramos oportuna a criação de uma base de dados eletrônica para dicionário *online*. E propomo-nos a desenvolver uma ferramenta com atenção ao emprego de busca onomasiológica e semasiológica seguindo o modelo de Pottier, apresentado mais à frente.

1.1 Modelo de Pottier

A elaboração deste trabalho apoia-se no modelo de dicionário terminológico onomasiológico proposto por Gava (2012), que por sua vez utilizou o modelo semântico desenvolvido por Pottier (1985, 1987, 1992) e Babini (2001, 2006).

Anuímos ao modelo de Pottier (1992), apresentado em Semântica Geral, seguindo sua definição de percurso semasiológico (ou do interpretante) e percurso onomasiológico (ou do enunciador), que consiste no embasamento estrutural do modelo de dicionário eletrônico que nos propomos a desenvolver.

Para o autor, os sujeitos mudam frequentemente de posição, ora como emissores, ora como receptores, transpondo, então, o percurso onomasiológico, quando há a intenção de dizer ao enunciado, e o percurso semasiológico, quando o sujeito parte do enunciado para chegar à sua interpretação, variando de acordo com o ponto de vista do enunciador (emissor) ou do interpretante (receptor).

Na opinião de Pottier, as virtualidades da língua correspondem ao sistema semiótico da língua e aos mecanismos de enunciação que permitem as realizações discursivas. E a passagem da conceptualização à semiotização se dá pela caracterização do fenômeno de designação, por onde se estabelecem as relações entre o mundo referencial e os sistemas das línguas naturais (GAVA, 2012, p. 117).

Já o enunciador tem como ponto de partida o mundo referencial (R). Quando tem a intenção de dizer (*querer dizer*), começa a conceptualizar sua intenção (Co). Essa conceptualização deve, então, ser expressa em signos, em um processo de semiotização, que se realiza pelos meios fornecidos por um sistema semiótico (GAVA, 2012, p. 116).

No percurso semasiológico, a interpretação chega ao conceito a partir do signo e parte do discurso realizado (texto oral ou escrito) para chegar à compreensão.

Então, é através da identificação e interpretação de diferentes elementos presentes no discurso que o interpretante procura compreendê-lo (*fenômeno de interpretação*) (GAVA, 2012, p. 117).

Como o conceito pode representar o ponto de partida para o enunciador e o ponto de chegada para o

interpretante, Pottier defende que o mundo referencial (R) não deve ser o objeto representativo do ponto de partida e chegada dos dois discursos, pois para ele o mundo conceitual deve assumir tal papel (POTTIER, 1992, p. 18).

Desse modo, Pottier caracteriza o mundo conceitual (Co) como o lugar da representação mental e sede da encenação, sendo independente das línguas naturais e dos sistemas semiológicos. A língua, para o autor, é a ponte do saber através da qual se realizam a transformação das representações mentais em signos e em esquemas (no caso do percurso do enunciador – onomasiológico); e a interpretação (no caso do percurso do interpretante – semasiológico) (GAVA, 2012, p. 116).

Dois tipos de conceitos permitem realizar a encenação: os conceitos gerais e os conceitos universais, os noemas, caracterizados como uma relação abstrata universal que sustenta as operações semânticas gerais das línguas, de caráter visual, próximos a uma representação mental comum (Pottier, 1992, p. 78).

Os noemas são representações relacionais, abstraídas das experiências, mas cujo traço linguístico assume formas muito variadas nas LN (Linguagens Naturais), subjazendo às operações semânticas gerais das línguas, e é visualizável a fim de aproximar, o máximo possível, da intuição de uma representação mental compartilhada. Os semas são elementos distintivos de um conjunto e sempre analisáveis em noemas e serão tratados no próximo item.

O autor agrupa os conceitos gerais em quatro grandes categorias, a saber: os seres, as coisas, as propriedades, as atividades.

Após a transformação dos conceitos em esquemas, o processo seguinte ocorre no nível da língua, através da transformação em signos linguísticos. É a *lexemização*, que é realizada através de conceitos gerais e de noemas.

No plano conceptual, as maneiras de compreender e conceptualizar seriam as mesmas para cada língua, cultura e indivíduo (Pottier, 1992, p. 78).

Para Babini (2006, p. 78), os elementos do modelo de Pottier são de grande importância para a compreensão dos percursos representativos do processo gerativo da enunciação e o funcionamento dos dicionários semasiológicos e onomasiológicos. Mas se contrapõe a tal afirmação, retomando, para isto, Karl Wilhelm Von Humboldt (1767-1835) em sua fala de que cada língua natural se caracteriza por uma forma interna que exprime a concepção (ou a visão) do mundo próprio da nação que a fala (HUMBOLDT, 2000 apud BABINI, 2006, p. 78), e conclui que o plano conceptual de cada indivíduo é determinado, de certo modo, por influência própria a cada cultura e que não serão, obrigatoriamente, os mesmos, em todas as línguas e culturas. Por esta razão, também o plano conceptual torna-se relativo e condicionado pelas dimensões espaço-temporais como, também, o plano da língua.

A relevância dos conceitos gerais e universais de Pottier para o nosso trabalho está na estruturação das categorias e no princípio da conceptualização que cada indivíduo assume diante das influências socioculturais.

Tais conjecturas exprimem o cuidado na construção de uma base tecnológica e dos campos necessários à representação das questões semânticas compartilhadas pelos falantes de diferentes etnias indígenas. Façamos, a seguir, sobre a *Etnoterminologia*.

1.2 Etnoterminologia

Em uma idealização da língua, os elementos empregam como propriedade fundamental a referenciação de um objeto, de acordo com a função que ele desempenha na enunciação. Entretanto, a enunciação partirá do interlocutor e a compreensão do elemento enunciado estará sujeita à visão de mundo do ouvinte. Desencadeando um processo de conceptualização no qual se poderia pensar em uma equivalência entre o que o interlocutor diz e o que o ouvinte assimila.

No entanto, Barbosa (2002) expõe que uma informação potencial, para os homens, se converte em substâncias estruturadas, quando apreendida por grupos linguísticos e socioculturais, de diferentes maneiras, embora mantenham um núcleo de percepção biológica universal.

Desse modo, a relação de equivalência é contestável e estará sujeita às qualidades conceituais internalizadas anteriormente, as quais serão oriundas de diferentes níveis do percurso gerativo da conceptualização.

Porquanto a percepção das ocorrências em torno será culturalmente filtrada e orientada por recortes culturais pré-existentes (PAIS, 1994, p. 170).

Barbosa (2002) classifica três estágios semânticos:

- **As latências:** nas quais o fato observável tem os seus traços identificadores, estado semântico potencial, enquanto substância de conteúdo estruturável, apreensível;
- **As saliências:** nas quais os traços semânticos se destacam.
- **As pregnâncias:** no qual o enunciador individual e/ou coletivo seleciona os traços que configuram o conceito do fato em questão. É neste nível que a influência/intervenção da cultura de cada um agirá em sua decisão semântica.

No nível da pregnância é que são produzidos os modelos mentais, os *conceptus*, noções ou conjuntos noêmicos, que são os traços semânticos conceptuais, correspondentes aos recortes culturais construídos.

Esses três estágios, que compreendem a percepção, desde o início ao final da conceptualização, estabelecem o percurso da cognição.

A partir dos valores semânticos, Barbosa (2006b, p. 48) propõe a consolidação da Etnoterminologia como subárea da terminologia e disciplina científica, quando a atenção se concentraria no estudo da “norma relativa ao estatuto semântico, sintático e funcional do conjunto das unidades lexicais no âmbito da cultura brasileira (BARBOSA, 2006, p. 48)”.

Para a autora, o estabelecimento preciso de sua função dependerá de sua inserção em uma norma discursiva, que determinará o estatuto de vocábulo ou termo (BARBOSA, 2006, p. 49). A autora expõe a passagem da terminologia para a terminologização, na qual um núcleo sêmico é mantido na intersecção de dois sememas, a metaterminologização.

Barbosa pontua ser necessário ir ao encontro do entendimento e da familiarização com o contexto, para que ocorra a compreensão do pensamento e do sistema de valores da cultura em questão, expressos em uma “linguagem paralela” que, apenas inserida no contexto, poderá ser interpretada corretamente (BARBOSA, 2006, p. 50).

A autora ainda enfatiza a importância do valor semântico social, em concomitância com o importante processo histórico de uma cultura (BARBOSA, 2006, p. 48).

Destarte, o profundo conhecimento do pensar indígena referência o universo sociocultural ao qual ele está inserido desde o berço, e que precisaríamos fazer parte desde tenra idade para assimilar e compreendermos em completude: seus dizeres, seus medos e seus mitos.

Observados os percursos cognitivos semânticos nos princípios descritos por Pottier (1987) e Barbosa (2006), discorreremos sobre o semema.

1.3 Semas e sememas

Um semema é o resultante de vários semas que juntos formam um significado global de um lexema que, por sua vez, é constituído de palavras de mesma classe morfológica, distribuídas de forma complementar nas quais a diferença situa-se unicamente nos sufixos e indivisível em unidades menores (GAVA, 2012, p. 124).

A diferença entre o semema e o lexema reside no fato de que o semema representa o conteúdo semântico do lexema, sendo o lexema uma unidade de mais alta categoria pelas informações lexicais agregadas (POTTIER, 1987 p. 61).

O sema representa uma unidade mínima de significação dentro de um campo semântico, que não se realiza fora do campo lexical. Sua presença consiste na ausência ou distanciamento de outros semas e sua função é estritamente relacional e não substancial, sendo que ele poderá se relacionar a elementos constituintes

de sememas distintos (POTTIER, 1987, p. 60). Como exemplo, o semema <cadeira>, possui os semas: “para sentar”, “com pés”, “com encosto”, “sem braços”.

De acordo com Pottier, os semas são de dois tipos: denotativos e conotativos. Por denotativo entendem-se os semas que determinam “de uma maneira estável e com ampla aceitação social a significação de um signo”. Eles se subdividem em semas específicos, aqueles que permitem distinguir semas vizinhos e semas genéricos, que indicam “uma categoria geral” (POTTIER, 1985, p. 29).

Os semas conotativos caracterizam-se pela “maneira instável e frequentemente individual à significação de um signo (POTTIER, 1985. p. 30)”.

No desenvolvimento de funções onomasiológicas, empregaremos os semas mais utilizados de acordo com nossa análise, os quais pretendemos agregar à nossa base de dados.

Para Gava (2012), os semas empregados a um termo podem ser oriundos de valores particulares de cada língua. Nas culturas indígenas, as questões de relevância para os sememas transmitem valores míticos, alimentares, culturais, intrínsecos às questões de sobrevivência, crenças ancestrais e valores socioculturais.

2 Materiais e métodos

De mais de 3 milhões de indígenas da época do descobrimento, temos hoje no Brasil cerca de 330 mil indígenas contatados. O estado do Mato Grosso é habitado por cerca de vinte e oito mil indígenas, distribuídos em trinta e oito etnias diferentes, além de povos não identificados ou não contatados oficialmente.

Desse modo, nossa estratégia de trabalho, visa desenvolver um software capaz de receber todas as línguas existentes no estado. Entretanto, a confecção de um corpus e sua adequação em um dicionário não se trata de mera lista de palavras (item 1), ou de uma inserção qualquer de termos em uma base eletrônica. Exige um trabalho árduo de lexicógrafos e terminógrafos, fontes seguras, buscas exaustivas a significados e recorrências, sistemas de remissivas, olhares dirigidos às questões culturais e sociológicas e a convivência com a cultura que nos faça permear a identidade cultural do outro.

Desse modo, rubricamos que o desenvolvimento do software, a que nos propomos representará apenas o ponto inicial de um trabalho árduo, de pesquisadores, estudiosos, lexicógrafos, entre outros, os quais, para a efetivação de suas pesquisas, encontrarão em nossa ferramenta lexicográfica o respaldo adequado e confiável do qual necessitam.

Portanto, focaremos nossas explicações no modo como nossa base deverá ser construída, os campos

necessários, a construção da base de dados, informações próprias do processo de desenvolvimento e disposição de dados.

Do ponto de vista da usabilidade do sistema, ao final do percurso, pretendemos expô-lo à academia, bem como às comunidades indígenas do Estado do Mato Grosso, por intermédio da Secretaria de Educação Indígena do Estado. A implantação do sistema se dará por intermédio de cursos instrutivos e demonstrativos.

O intuito primordial é a inclusão dos professores indígenas, para que possam solidificar seu conhecimento linguístico, utilizar a base terminológica para a inserção de termos, disseminando o conhecimento entre seus pares.

Exposta nossa intenção prima, ater-nos-emos, a seguir, no escopo do sistema.

2.1 A elaboração da base de dados terminológica

Para o desenvolvimento do dicionário, destacamos o necessário conhecimento em linguagens computacionais e em definições de elementos de dados e suas respectivas características. Apresentamos agora os campos da base de dados terminológica (GAVA, 2012):

- Termo: espaço para se colocar o termo.
- Língua: espaço para registrar a língua de origem do termo.
- Outras designações: espaço para registrar as outras designações, ou seja, as variantes do termo na língua pesquisada.
- Equivalência *l..n*: espaço para registrar o termo de equivalência.
- Variante *l..n*: espaço para registrar, se houver, a variante do termo equivalente pesquisado.
- Contexto *l..n*: espaço para se registrar o contexto encontrado na fonte 1.
- Classe gramatical: espaço para registrar a classe gramatical do termo encontrado.
- Gênero: espaço para registrar o gênero gramatical do termo.
- Fonética: espaço para registrar os sons do termo.
- Derivação: espaço para registrar a derivação do termo, se houver.
- Nome científico: espaço para registrar o nome científico do termo.
- Imagem: botão que permite a busca do arquivo de imagem do termo.
- Fonte do nome científico: espaço para registrar a fonte do nome científico.
- Fonte da imagem: espaço para registrar a fonte da imagem.

- Fonte *l..n* – para equivalências e variantes: local reservado para se registrar o nome da fonte da qual foi retirado o contexto que atesta a validade do equivalente pesquisado para o português.
- Observações: observações adicionais sobre o termo.
- Descrição: espaço para registrar as características onomasiológicas que compõem o termo.
- Semema: semema relativo ao termo.

Tal estrutura otimiza consideravelmente a inserção, o desempenho e a recuperação de informações.

As equivalências, no que tange ao estabelecimento de um equivalente na pesquisa terminológica multilíngue, seguem os seguintes parâmetros:

1. O termo *A* não possui sua equivalência conhecida em todas as línguas.
2. *Equivalente não encontrado*: o termo da língua *A* não encontra termo equivalente para a língua *B*, ou vice versa.
3. *Correspondente*: o termo da língua *A* recobre, apenas parcialmente, os traços conceptuais de um termo da língua *B*, ou vice-versa.

Os verbetes de remissivas são destinados tanto a termos em língua portuguesa quanto a termos nas línguas indígenas da seguinte forma (Babini, 2001, p. 95):

Campos	→	Conteúdo
• Termo	→	Entrada
• Categoria	→	Categoria gramatical
• Fonte	→	Fonte da entrada
• Ver	→	Remissiva

2.2 Estrutura da base de dados

Na utilização de uma base de dados apoiada em um modelo, para a organização de uma nova língua ou novo léxico, apenas são necessários ajustes nas etiquetas ou nomes dos campos que aparecem para o usuário final na *interface*, ao passo que, as modificações significativas ocorrerão na estrutura da base.

O *layout* será desenvolvido em abas, com base no modelo de dicionário onomasiológico terminológico de Gava (2012), as quais proporcionam ao usuário uma visualização global dos recursos oferecidos.

O componente de inserção é composto de duas páginas *web*, as quais denominaremos: passo um – inserção do termo; e passo dois – inserção de características semânticas do termo.

No passo um, seleciona-se a língua indígena do termo, insere-se o termo na língua indígena e a sua equivalência em língua portuguesa. Se a equivalência em português não for reconhecida pelo sistema de banco de

dados, o sistema solicita que ela seja cadastrada. Nesse cadastro são solicitadas as características do termo (Língua Portuguesa): classe gramatical, gênero, nome científico, fonte do nome científico, descrição, fonte da descrição.

Se o sistema reconhecer que o termo e suas características já existem na base em língua portuguesa, ele direciona a inserção apenas na língua indígena.

As características solicitadas referentes ao termo indígena são: a fonte principal do termo, classe gramatical, gênero, fonética, derivação, observação.

Dessa forma, as informações relativas ao nome científico, referência à fonte do nome científico, descrição, fonte da descrição são solicitadas apenas no momento da inserção do equivalente em língua portuguesa, evitando que o usuário precise repetir a informação para todas as línguas.

A inserção das imagens e a referência à fonte da imagem, também ocorrem no momento da inserção da equivalência. Entendemos esse como um ponto alto no dicionário, uma vez que as figuras reforçam o conteúdo do termo.

Novos traços semânticos, ou traços não previstos poderão ser associados ao termo, utilizando a caixa de diálogo disposta ao final da página. Os novos semas devem, necessariamente, estar separados por vírgula.

O passo dois corresponde ao cadastro das características onomasiológicas, no qual uma tela é visualizada pelo usuário no momento da inserção dos atributos do termo; a partir de uma relação de características semânticas, ele poderá selecionar os semas adequados. A tela de seleção de semas é visualizada pelo usuário no momento da inserção de atributos para o termo na língua de indexação (o Português).

Outro fator de relevância é a capacidade de agregar novas línguas, que passarão a contemplar a relação de idiomas e referências já existentes. Do mesmo modo, novos semas poderão ser etiquetados ou semas atuais removidos.

A inserção de novos semas contemplará as características semânticas das culturas autóctones, implicando a ocorrência de variação semântica entre as culturas. O que implica uma ferramenta direcionada à simplicidade e fácil manuseio. Desse modo, o desenvolvimento da aplicação levou em consideração o usuário com pouco ou nenhum conhecimento de informática.

2.3 Abas e funções

O dicionário multilíngue gerado possibilitará ao consulente efetuar dois tipos de buscas: semasiológica e onomasiológica. As buscas onomasiológicas podem ser feitas pelos traços semânticos que serão apresentados ao

consulente ou através da inserção de semas em um campo apropriado.

Outra forma de busca possível, ocorre através do nome do termo ou de parte dele, por exemplo, se pedirmos um termo com a letra *a*, o sistema retornará todos os termos que contenham a letra *a* em qualquer posição do vocábulo. Quanto mais completa a informação utilizada na busca, mais preciso será o resultado (busca semasiológica).

Buscas concomitantes com partes do nome e atributos onomasiológicos também são possíveis.

A função semasiológica é representada por uma consulta, na qual o consulente poderá digitar o termo em português ou em qualquer uma das línguas cadastradas, obtendo uma resposta imediata à consulta.

O retorno à consulta, para qualquer uma das línguas indígenas, deverá ser feito por meio da equivalência em português (indexador), seguido do termo equivalente nas demais línguas, de acordo com a solicitação do consulente.

O resultado à consulta semasiológica é um termo ou uma relação de termos ordenados alfabeticamente. O diferencial do dicionário é o caráter combinatório das funções de busca.

Os casos de associações indevidas, que poderiam originar erros, serão tratados computacionalmente. São eles os atributos concorrentes, os quais não podem pertencer ao mesmo tempo às duas categorias, ditas opostas.

A função onomasiológica poderá ser modificada de acordo com as necessidades do usuário-organizador (aquele que inserirá novos termos e editará novas condições de busca), pois o modelo proposto possibilita a inserção de novos grupos de atributos e o estabelecimento de novas características e novas condições para características concorrentes.

As funções onomasiológica e semasiológica acontecem concomitantemente na busca eletrônica, mas a função semasiológica está presente na aba *Dicionários*, na qual são apresentados todos os termos dos dicionários em ordem alfabética, obedecendo o nome de entrada em língua portuguesa, seguida das equivalências em línguas indígenas.

Os elementos históricos e culturais dos falantes das línguas possuirão um lugar reservado na aba *Línguas Indígenas*. Na qual uma sucinta história da língua e de seus falantes poderão ser encontradas.

Todas as Fontes empregadas na base serão listadas automaticamente pela ferramenta em uma aba específica.

A última aba do dicionário destina-se às notas do(s) autor(es), comentários importantes, boas vindas etc. O espaço será empregado de modo arbitrário, através de uma janela de texto.

Para todas as consultas, a relação de resultados surge quando o processo de varredura interna do banco de dados é finalizado, e as implicações do processamento são listadas pelo sistema ao consulente.

O primeiro resultado encontrado é uma lista das sugestões dadas pelo sistema à consulta do usuário, em ordem alfabética. Essa relação possui um *hiperlink* que trará ao usuário a ficha terminológica do termo, se acessado. Assim, pela criação de *link* na linguagem XHTML, a relação resultante agrega um valor fundamental para ao usuário final.

Por fim, assinalamos a habilidade do *software* em comportar os sons da voz humana, relacionado ao termo e de armazenar os sons produzidos pelos animais, por exemplo, o canto de um pássaro, o uivo de um lobo, o urro de uma onça, agregando ao termo o seu fonema e vivificando um componente lúdico.

3 Considerações finais

Uma vez que a base de dados baseia-se no trabalho concluído de Gava (2012), no qual criou-se um modelo de dicionário para até seis línguas indígenas, acreditamos na viabilidade da ferramenta.

As linguagens computacionais utilizadas aliam desempenho, estabilidade, registro permanente de informações, consultas em tempo real e resultados satisfatórios. E conseqüentemente, uma melhora no processo de inserção, exclusão, varredura, consultas e resultantes terminológicas. Além de permitirem a publicação *web* do dicionário. De modo que um usuário lusófono poderá pesquisar um termo ao mesmo tempo em que outro termo é inserido em outra parte do país por um professor indígena, por exemplo, sem se preocupar com o acesso concorrente.

Acreditamos na relevância de nosso trabalho no registro, não só dos termos, mas do contexto e ideias a ele associados e na capacitação de acadêmicos e docentes indígenas para a utilização da ferramenta.

A recolha de termos levantados, de diferentes etnias, comporá o conjunto de ferramentas e facilitadores aplicados ao ensino, além de proporcionar ao aluno indígena uma visão macro do termo em outras línguas, podendo levá-lo à reflexão sobre as diferenças e proximidades entre os idiomas, bem como auxiliá-lo na compreensão da representatividade de um animal extinto ou do caráter histórico do termo para seus antepassados.

Entendemos que a ferramenta abrigará um tesouro valioso das línguas indígenas mato-grossense, um veio para novas pesquisas e um movimento a favor do resgate etnocultural de vários povos. Entretanto compreendemos tratar-se de uma trajetória árdua, a qual demandará

tempo, esforço e o envolvimento de vários interessados para que se edifique com solidez um *corpus* respeitável.

Muitas variáveis e questionamentos surgirão até o momento da replicação da ferramenta para os demais falantes indígenas do país.

Para as futuras inserções, desejamos utilizar a base aqui demonstrada como ferramenta de apoio ao pós-doutorado, iniciado no ano de 2014, o qual apresentará os sons da voz humana dos indígenas pronunciando os nomes dos animais em língua Zoró.

A ferramenta ora apresentada estará disponível aos trabalhos de acadêmicos, projetos de apoio e extensão de graduação e especializações, bem como à comunidade indígena. A todos ofertaremos o treinamento e o respaldo técnico necessário durante o período de implantação e capacitação da base ora ofertada.

Referências

- BABINI, Maurizio. Do conceito à palavra: os dicionários onomasiológicos. *Ciência e Cultura*, São Paulo, v. 58, n. 2, p. 38-41, abr.-jun. 2006.
- BABINI, Maurizio. *Onomasiologie et dictionnaires onomasiologiques*. São José do Rio Preto: Beatriz, 2001.
- BARBOSA, Maria Aparecida. Para uma etnoterminologia: recortes epistemológicos. *Ciência e Cultura*, São Paulo, v. 58, n. 2, p. 48-51, abr.-jun. 2006. Disponível em: <http://cienciaecultura.bvs.br/scielo.php?script=sci_arttext&pid=S0009-67252006000200018&lng=en&nrm=iso>. Acesso em: 06 dez. 2011.
- BERTOLDI, Vittorio. *Enciclopedia italiana di scienze, lettere ed arti*. Roma: Treccani, 1935.
- BIDERMAN, Maria Tereza (Org.). A ciência da lexicografia. *Alfa – Revista de Linguística*, São Paulo, v. 28, p. 1-26, 1984. Suplemento.
- BIDERMAN, Maria Tereza. A língua e o computador. In: BIDERMAN, Maria Tereza. *Teoria linguística*. São Paulo: Martins Fontes, 2001. p. 75-93.
- BOULANGER, Jean Claude. Convergências e divergências entre a lexicografia e a terminologia. In: LIMA, M. S.; RAMOS, P. C. (Org.). *Terminologia e ensino de segunda língua: Canadá e Brasil*. Porto Alegre: UFRGS, 2001. p. 7-28.
- CABRÉ, Maria Tereza. *La terminología: representación y comunicación, elementos para una teoría de base comunicativa y otros artículos*. Barcelona: Institut Universitari de Linguística Aplicada, 1999. <http://dx.doi.org/10.1075/tlrp.1>
- FARGETTI, Cristina Martins (Org.). *Abordagens sobre o léxico em línguas indígenas: dicionário de línguas indígenas e questões de prosódia*. Campinas: Ed. Curt Nimuendajú, 2012. p. 67-81.
- FARIAS, Virginia Sita. O emprego de ilustrações como mecanismos de elucidação do significado das unidades léxicas nos dicionários semasiológicos. In: ENCONTRO DO CELSUL, 9., 2010, Palhoça. *Anais...* Palhoça: Editora UNISUL, 2010. p. 1-19.

- FAULSTICH, Enilde; OLIVEIRA, Michelle Machado de. Para que serve um dicionário analógico? Um estudo de lexicografia comparativa. In: CONGRESSO INTERNACIONAL DE HUMANIDADES BRASIL-CHILE: DIVERSIDADE CULTURAL E INTEGRAÇÃO LATINO-AMERICANA, DIMINUINDO DISTÂNCIAS, 10., 2001, Brasília (DF). *Anais...* Brasília (DF): Universidade de Brasília, 2001. Disponível em: <<http://www.onda.eti.br/revistaintercambio/conteudo/arquivos/1221.pdf>>. Acesso em: 28 abr. 2011.
- GAVA, Águida Aparecida. Jogos Computacionais como ferramenta de ensino. *Revista UNORP*, São José do Rio Preto, v. 4, n. 1, p. 93-54, 2003.
- GAVA, Águida Aparecida. *Plataforma Kuhi pei*: proposta de um modelo de dicionário terminológico onomasiológico multilíngue para crianças, Português-Arara, Kadiwéu, Karitiana, Parintintin, Xavante, Zoró. 2012. Tese (Doutorado) – Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista, São José do Rio Preto, 2012.
- INSTITUTO SOCIOAMBIENTAL (ISA). *Povos indígenas no Brasil*. 2011. Disponível em: <<http://pib.socioambiental.org/pt/povo/arara>>. Acesso em: 21 set. 2011.
- LARA, Luís Fernando. O Dicionário e suas disciplinas. In: SIMPÓSIO INTERNACIONAL DE LEXICOGRAFIA, 1., 2002, Barcelona. *Anais...* Barcelona: Instituto Universitário de Linguística Aplicada da Universidade Pompeu – Fabra, 2002.
- ISQUERDO, Aparecida Negri; ALVES, Ieda Maria. *As ciências do léxico*: lexicologia, lexicografia e terminologia. 2. ed. Campo Grande: EdUFMS, 2001.
- PAIS, Cidmar Teodoro. O percurso gerativo da enunciação: produtividade léxica e discursiva. *Confluência*. Assis, v. 3, n. especial, p. 162-181, 1994.
- PATRIZZI, Vanessa de Paula Rodrigues. *Dicionário terminológico dos termos fundamentais da linguagem das produções telejornalísticas*. 2007. Dissertação (Mestrado) – Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista, São José do Rio Preto, 2007.
- POTTIER, Bernard. *Linguistique générale: théorie et description*. Paris: Klincksieck, 1985.
- POTTIER, Bernard. *Sémantique générale*. Paris: PUF, 1992.
- POTTIER, Bernard. *Théorie et analyse en linguistique*. 2. ed. Paris: Hachette, 1987.
- SILVA, Eduardo Batista da. *Proposta de um dicionário terminológico onomasiológico bilingue inglês-português no domínio das redes neurais artificiais*. Dissertação (Mestrado) – Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista, São Jose do Rio Preto, 2009.
- WÜSTER, Eugen. *Introducción a la teoría general de la terminología y a la lexicografía terminológica*. Barcelona: Institut Universitari de Lingüística Aplicada, 1998.

Recebido: 20 de julho de 2014
Aprovado: 27 de outubro de 2014
Contatos: guidag@gmail.com
fabriogava@gmail.com