

REFLEXÕES SOBRE A ATIVIDADE DE ELABORAÇÃO DE UM DICIONÁRIO BILÍNGUE

Marília de Nazaré Ferreira¹
(Universidade Federal do Pará)
marilia@ufpa.br

RESUMO: A investigação sobre o Léxico tem sido o foco de interesse de áreas de pesquisa como a Linguística Teórica, a Psicolinguística e o Processamento Automático das Línguas Naturais, que buscam conceituá-lo, por meio de observações acerca de sua natureza. Por meio da pesquisa que embasou este trabalho, foi possível abordar brevemente temas relacionados ao estudo do léxico, uma vez que a atividade de elaboração de um dicionário centra-se em uma perspectiva interdisciplinar, por meio da qual se pode, de modo eficiente, descrever línguas ainda não estudadas; uma vez que se parte da observação de aspectos linguísticos, extralinguísticos e culturais.

PALAVRAS-CHAVE: língua indígena, descrição linguística, dicionário bilíngue, léxico.

ABSTRACT: Investigations about the lexicon have been the object of several research fields, such as Theoretical Linguistics, Psycholinguistics and Automated Natural Language Processing, which aim to define it through observations of its inherent characteristics. This work briefly approaches a subject related to the analysis of the lexicon: the development of a bilingual dictionary on one of Brazil's indigenous languages. It is known that the development of a dictionary is an efficient way to describe a language, and that such task allows us to observe, for instance, how the lexical entries are specified regarding its content.

KEYWORDS: indigenous language, linguistic description, bilingual dictionary, lexicon.

¹ Professora Adjunta do Instituto de Letras e Comunicação vinculada à Faculdade de Letras da Universidade Federal do Pará.

0. Introdução

Os estudos sobre o léxico tem sido alvo de interesse de áreas de pesquisa como a Linguística Teórica, a Psicolinguística e o Processamento Automático das Línguas Naturais, que têm buscado conceituá-lo, por meio de observações acerca de sua natureza, dentre outras coisas. De uma outra perspectiva, há também a possibilidade de se estudar o léxico partindo-se de arcabouços teóricos da lexicologia e da lexicografia.

A tarefa de elaborar um dicionário bilíngue de uma língua é uma árdua e complexa tarefa, que engloba diferentes áreas cujo interesse é o léxico. Quando se pensa, então, em elaborar um dicionário bilíngue de uma língua ainda em descrição, mais delicada é a questão, uma vez que se tem a tarefa de, por meio da dicionarização, descrever os padrões linguísticos que constituem uma dada língua.

Inúmeros linguistas que têm trabalhado com a elaboração de dicionários concordam que a elaboração de um banco de dados possibilita uma descrição mais completa e abrangente de uma língua ainda não descrita, ou parcialmente descrita, como é o caso de várias línguas indígenas brasileiras tais como o tapajúna, as quais ainda são as línguas mais desconhecidas do mundo.

Ao elaborar um banco de dados, um linguista terá contato com aspectos diversos da língua, desde a fonética e a fonologia até aspectos semânticos-pragmáticos. Por esta razão, trabalhos dessa ordem podem contribuir para a descrição e a documentação de línguas, mesmo aquelas consideradas em perigo de extinção, que é a situação de várias línguas indígenas brasileiras atualmente.

Em termos científicos, a elaboração de um dicionário para uma língua representa também a possibilidade de contribuir para o conhecimento universal linguístico desde que se encontre um dado padrão ou ainda não documentado, ou pouco estudado. Há também as contribuições práticas decorrentes da realização desses trabalhos, uma vez que as comunidades falantes das línguas pesquisadas poderão fazer uso de um banco de dados e de um dicionário propriamente dito como instrumentos didático-pedagógicos para trabalhos de manutenção e revitalização linguísticas.

Por meio da pesquisa que embasou este trabalho, foi-nos possível abordar brevemente temas relacionados ao estudo do léxico, uma vez que a atividade de elaboração de um dicionário bilíngue traz em si a necessidade da teoria e da prática "casadas". Sabe-se que a elaboração de um dicionário é uma maneira eficaz de descrever línguas; por meio dela é possível se observar como, por exemplo, são especificadas as

entradas lexicais da língua no que diz respeito ao seu conteúdo, bem como também é possível verificar os usos de um determinado item lexical, sua significação, sua formação e constituição fonético-fonológica.

O conhecimento desses inúmeros aspectos de uma determinada língua é de suma importância também para a área de ensino e aprendizagem, que pode, juntamente com a linguística, fundamentar o trabalho de fortalecimento ou revitalização dos usos de uma língua, principalmente nos casos de línguas minoritárias.

O presente trabalho apresenta algumas reflexões sobre a atividade de elaboração de um dicionário bilíngue, cujo tema é fauna e flora de uma língua indígena brasileira: o tapajúna. O trabalho foi iniciado em 2004 por Nayara Camargo e foi tema de sua monografia de final de curso de graduação em letras, sob minha orientação. O dicionário piloto contou com a inserção de trezentas e cinquenta palavras em língua indígena e em português, com a inserção de imagens da fauna e da flora amazônicas, bem como de nomes científicos das espécies para fins de clareza quanto ao repertório verbal utilizado. O dicionário está em elaboração por Camargo, que continua a estudar a língua tapajúna em seu curso de pós-graduação na Universidade Estadual de Campinas.

Como o trabalho foi iniciado em 2004, em uma de minhas primeiras etapas de campo entre os tapajúna, quando ainda estávamos elaborando um sistema ortográfico para a língua, os itens lexicais foram escritos parcialmente nessa ortografia, uma vez que em alguns casos era necessário – até que se pudesse contar com o auxílio de um falante nativo – manter a transcrição fonética.

Neste artigo, serão apresentadas, na primeira seção, informações sobre o povo, a língua tapajúna e sua situação sociolinguística. Na segunda seção, são apresentados pressupostos teóricos que abordam algumas concepções de léxico, de definições de dicionários. Finalmente, são apresentadas, na terceira seção, algumas reflexões feitas no decorrer dessa experiência relacionadas às etapas de elaboração de um dicionário bilíngue tapajúna e português. São reflexões a que cheguei durante o complexo trabalho de dicionarizar palavras de uma língua em descrição.

De forma alguma, pretendo, por meio deste trabalhar, apresentar passos sobre como realizar a elaboração de uma obra lexicográfica, ou ainda ensinar a produzir um dicionário. Para isto, cada autor interessado em trabalhos dessa ordem necessita ter clareza quanto aos seus objetivos para chegar ao fim desejado. Neste artigo, apresento minha experiência, enquanto participante efetiva de um trabalho de descrição de uma língua indígena brasileira, por meio da elaboração de um dicionário, indicando as inúmeras perspectivas de estudo que se abrem diante de um estudioso que debruça nessa tarefa que envolve diferentes parâmetros.

1. A língua e o povo tapajúna e sua situação sociolinguística

O Brasil é um país multilíngue, que abriga cerca de cento e oitenta línguas indígenas em seu território, dentre as quais a língua tapajúna, pertencente ao tronco linguístico Macro-Jê.

O povo tapajúna, que, atualmente, soma menos de cem indivíduos, vive no estado do Mato Grosso, em uma aldeia Mebengôkrê, de índios kayapó, os quais falam uma língua geneticamente aparentada ao tapajúna.

A história desse povo remonta a um passado permeado de momentos trágicos, de mudanças de território, de luta pela sobrevivência. A última mudança que os tapajúna foram obrigados a empreender os retirou de seu território tradicional para uma aldeia kayapó, onde estão até o presente.

A herança dessa trajetória pode ser observada na situação sociolinguística atual, caracterizada pela iminente perda linguística devido ao longo contato com uma língua parente – o kayapó – e também pelo contato com a língua portuguesa. Não é possível precisar em que termos e em que exatamente a língua tapajúna sofreu influência da língua kayapó.

Pode-se falar em termos de atrito linguístico, já que a língua majoritária falada na aldeia em que vivem os tapajúna é o mebengôkrê kayapó. De acordo com Ferreira (2005), a língua parkatêjê vive uma situação semelhante, embora nesse caso a língua dominante é o português: “[...] a substituição da língua minoritária (no caso o parkatêjê) pela majoritária (no caso o português), pode levar ao completo abandono daquela, e até mesmo à sua morte.”

O mebengôkrê kayapó e o tapajúna são línguas estruturalmente parecidas, por isso há uma sobreposição entre o uso dessas línguas. Como a maior parte da comunidade tapajúna fala fluentemente o mebengôkrê, é provável que essa seja a língua mais utilizada, sendo, inclusive, a língua materna aprendida pelas crianças, uma vez que houve inúmeros casamentos interétnicos entre indivíduos pertencentes aos dois povos. Um grande número desses casamentos ocorreu entre homens tapajúna e mulheres mebengôkre e, em sendo aquelas sociedades matrilocais, isto é, o casal fixa residência junto à família da esposa, os homens tapajúna passaram a falar mebengôkre também em casa.

A língua portuguesa também é conhecida e utilizada pelos tapajúna, pois os professores indígenas anualmente participam de um curso de formação, sob a responsabilidade da Fundação Nacional do

Índio (FUNAI) e MEC (Ministério da Educação e Cultura) ministrado por profissionais não-índios. Além disso, há indígenas que são funcionários da FUNAI e por isso têm contato direto com um ambiente em que a língua portuguesa é dominante.

Assim, não é difícil reconhecer que o tapajúna encontra-se em perigo de extinção, uma vez que, entre os mebengôkre, somente vive cerca de quarenta indivíduos falantes. Trata-se de uma língua minoritária, ágrafa, cuja sobrevivência dependerá da ação da própria comunidade em querer preservá-la.

A língua tapajúna é uma língua que ainda não conta com uma descrição sistemática. Minha pesquisa com a comunidade ocorreu nos anos de 2004, 2005 e 2006, quando estive em área indígena como docente do curso de formação de professores. Com base nesse trabalho de pesquisa orientei uma dissertação de mestrado em Linguística Histórica (Rodrigues, 2009), dois trabalhos de conclusão de curso de graduação na UFPA (Camargo, 2005 e Rodrigues, 2006) e um plano de Iniciação Científica. Todos esses trabalhos incluíram dados da língua tapajúna por mim coletados nos anos em que estive realizando pesquisa de campo na aldeia Metyktíre, localizada no Mato Grosso.

A metodologia utilizada neste trabalho foi a usual da Linguística Descritiva que inclui pesquisa bibliográfica e pesquisa de campo em que se faz uma seleção de informantes de pesquisa para a coleta de dados da língua. O trabalho foi realizado em fases distintas, a saber: as viagens ao campo para a coleta de dados linguísticos, a transcrição fonética dos dados coletados, a análise morfossintática e semântica dos itens lexicais, a confirmação ou refutação de hipóteses e a inserção dos dados em um programa computacional utilizado para elaborar dicionários.

Na próxima seção, abordarei brevemente alguns aspectos dos pressupostos teóricos sobre o léxico, apresentando observações sobre a língua em estudo face às considerações teóricas utilizadas para a elaboração deste trabalho.

3. Pressupostos Teóricos

De uma perspectiva bem abrangente, o léxico de uma língua natural é concebido como um sistema abstrato de idéias e de conceitos relativos ao conhecimento prévio da realidade de um indivíduo em uma determinada sociedade. Essa apreensão de idéias, de conceitos e dos acontecimentos da sua realidade é articulada através de definições e nomeações estipuladas pelo indivíduo tendo como princípios, os acontecimentos da sociedade da qual ele faz parte.

Deste modo, a produção do acervo lexical é realizada pelo falante por meio de mecanismos psicolinguísticos e neurolinguísticos capazes de formular e articular as experiências adquiridas na sociedade. Essas experiências são armazenadas na memória e codificadas por signos linguísticos, que são parte do acervo de lexemas do indivíduo, possibilitando-o a comunicar-se com outros indivíduos de meio sociocultural equivalente.

Assim, uma língua está em constante processo de mudança e seus falantes são livres para criar itens lexicais, ou mesmo modificar itens já existentes – os quais podem cair em desuso ou terem seu significado expandido – dependendo de sua necessidade de comunicação.

O léxico mental é um dos componentes centrais dos sistemas de processamento das línguas naturais. Essa centralidade se deve ao fato de ele armazenar, além das unidades lexicais, um complexo conjunto de informações específicas de um sistema e de poder englobar propriedades fonológicas, morfológicas, sintáticas, semânticas, pragmáticas e discursivas de itens lexicais da língua. Essas informações todas são manipuladas durante o processo de interpretação de uma dada língua natural. Deste modo, a concepção de léxico como uma “lista de entradas lexicais”, um “componente marginalizado” da gramática tem sido abandonada. Ao contrário, verifica-se uma concepção de léxico como uma estrutura altamente organizada.

Tanto a Linguística Teórica, a Psicolinguística quanto o Processamento Automático das Línguas Naturais, apesar de apresentarem suas distinções peculiares no que se refere ao tratamento dado ao Léxico, convergem em um ponto: todas essas teorias concordam com a centralidade e a estruturação desse componente da gramática.

Muito embora não haja consenso em Linguística Teórica quanto à natureza do Léxico, há várias hipóteses consensuais entre os pesquisadores, dentre as quais destacamos (1) o fato de o léxico ser tido como uma rede de correspondências em vários níveis, e o item lexical apresentar em si propriedades fonológicas, morfológicas, sintáticas e semânticas e (2) a organização funcional do léxico que possibilita a recuperação do significado, características gramaticais e usos pelo falante. (Mel’cuk, 1988)

Os psicolinguistas acreditam na existência e no funcionamento de mecanismos envolvidos no processamento da linguagem humana e para compreender como ocorre o armazenamento e o acesso aos itens lexicais de uma dada língua, eles postulam a existência de um léxico mental.

Dentre inúmeros estudiosos dessa área, Bierwisch e Schreuder (1992) apontam o fato de o processamento cognitivo da linguagem envolver três tipos de operações: a conceitualização (especificação de

conceitos), a formulação (seleção de itens e construção de representações sintáticas e fonéticas) e a articulação (produção da fala), as quais abordaremos sucintamente.

Quando da produção de enunciados, a nomeação de um objeto envolve: a identificação do objeto que ativa uma vasta base de conhecimento que contém informações extralinguísticas oriundas de diversas fontes (visual, auditiva, motora, emotiva, conceitual, dentre outras), além de princípios gerais de organização conceitual. A formulação é responsável pela transformação da estrutura conceitual em enunciado linguístico. A primeira etapa da formulação relaciona-se à seleção de uma representação sintático-semântica do objeto. Para Bierwisch e Schreuder (*idem*), o resultado da primeira etapa do processo é uma forma semântica, que é transformada em forma fonológica, no segundo momento.

Por fim a articulação atua sobre a forma fonológica a fim de prover a realização fonética da forma selecionada.

Grosso modo, pode-se explicar o processamento da seguinte maneira: para nomear um objeto, o falante deve identificá-lo (conceitualização); após tê-lo identificado, a seleção de uma representação sintático-semântica é requerida; em seguida, o objeto é codificado em termos fonológicos e finalmente realizado foneticamente, o que tem como resultado o nome desse objeto.

Em tapajúna, por exemplo, de uma perspectiva icônica, o falante conceitua um animal como 'coruja', tendo como base informações extralinguísticas, que permitem a compreensão de que o som produzido pela coruja assemelha-se à forma do lexema. O nome do animal nessa língua é *hôhô*. O conceito é formulado, juntando-se o reconhecimento do animal ao item lexical proposto, o qual é articulado, conforme as regras de formas fonológicas da língua.

Com base no exposto, observa-se que, de fato, a produção de um lexema considera as experiências de um indivíduo em sua realidade. Por isto, a observação de Biderman (2001, p. 78) de que "Qualquer sistema léxico é a somatória de toda a experiência acumulada de uma sociedade e do acervo da sua cultura através das idades." é pertinente a esse ponto de nosso trabalho.

Por ser mentalmente estruturado, o léxico tem atividades de grande complexidade, como a categorização lexical. Observando-se o léxico da língua tapajúna, considerando-se informações sobre a fauna, pode-se postular que o falante procedeu a categorização lexical dos animais tendo como principal aspecto de observação sua aparência física. Assim palavras como *tracajá* e *tartaruga* têm uma forma semelhante – *kahrã* e *kahrãtxi*, respectivamente. Pode-se hipotetizar que tais nomes procedem também de um nome genérico e de uma especialização desse nome genérico, uma vez que *tracajá*, segundo

Houaiss (2001), é "tartaruga de água doce da família dos pelomedusídeos (*Podocnemis unifilis*) encontrada nos rios amazônicos, com cerca de 50 cm de comprimento, carapaça abaulada, pardo-escura, e cabeça com manchas alaranjadas." enquanto tartaruga é definida pelo mesmo dicionário como "designação comum aos répteis da ordem dos quelônios, ovíparos encontrados em quase todo o mundo, tanto na água doce ou salgada como em ambiente terrestres [...]".

A esse respeito, Biderman (2001) afirma que a taxonomia "é a elaboração própria de cada cultura e dela resultante. A não ser nos casos em que a ciência já classificou os dados da realidade segundo modelo lógico universal, cada língua possui sua classificação própria".

Em se tratando das disciplinas específicas dentro da ciência linguística que estudam o léxico, há discussões sobre os âmbitos e alcances da lexicologia e lexicografia, muito embora não esteja no escopo deste artigo discutir tais relações.

Conforme Dapena (2002), alguns estudiosos afirmam que a lexicologia e a lexicografia são como "faces de uma mesma moeda", nas quais suas diferenças corresponderiam às suas extensões e a uma diversidade de pontos de vista. Outros estudiosos afirmam que a lexicologia e lexicografia têm objetos de estudos completamente distintos. Deste modo, consideramos que essas disciplinas têm um objeto comum – o léxico – com enfoques e perspectivas diferenciadas. Dentro do âmbito da lexicografia, está a feitura de obras como um dicionário, por isso denominadas de lexicográficas.

A seguir, observar-se-á questões práticas atinentes à elaboração do dicionário bilíngue tapajúna-português.

4. A elaboração de um dicionário bilíngue

A elaboração do dicionário tapajúna constituiu-se no tema do trabalho de conclusão de curso de Camargo (2005) e teve a fauna e a flora como os assuntos selecionados para se proceder a investigação da língua, tendo em vista a possibilidade de se investigar os aspectos linguísticos e culturais desses temas. Em 2008, ela publicou um artigo sobre o assunto.

Para o trabalho de elaboração do dicionário, foram realizadas três viagens de campo, nos anos de 2004, 2005 e 2006, durante as quais foi coletado grande número de dados linguísticos. Uma parte expressiva desses dados foi coletada *in loco*, tendo-se, inclusive, como parte do trabalho, adentrado a floresta em companhia de professores indígenas e de Camargo, aluna de graduação à época, por ocasião da viagem de campo em 2005.

Após a etapa de coleta de dados em 2004 e 2005, procedi a transcrição fonética do material coletado por meio de gravações. Para a transcrição foi utilizado o Alfabeto Fonético Internacional.

Já em 2005, após a viagem de campo, Camargo iniciou a etapa de inserção dos itens lexicais em um programa computacional – o *toolbox*² – o qual é ancilar para linguistas e antropólogos na tarefa de armazenamento de dados. Esse programa permite a integração de vários tipos de dados de seus arquivos, sendo possível selecioná-los e organizá-los de diferentes maneiras. O programa integra dados lexicais, culturais, gramaticais etc. Oferece ainda a possibilidade de integrar dados de textos, favorecendo a interlinearização das informações.

Cada ficha do *toolbox* é planejada em conformidade com a necessidade do pesquisador ou do material em análise. Uma ficha contém diferentes marcadores, que contém informações acerca dos itens lexicais que forem inseridos no programa.

Os marcadores utilizados no dicionário foram: lexema (lx) utilizado para nomear a entrada, o verbete do dicionário em língua vernacular. Homônimo (hm) utilizado para quantificar o número de palavras que apresentam formas equivalentes com significados diferentes. Forma fonética (ph) para informar a transcrição fonética das palavras que constam como entrada do dicionário. Glosa (gn) termo utilizado para indicar a tradução em português. Referência (rf) para informar acerca da origem do lexema inserido no dicionário, ou ainda o nome do pesquisador que coletou o dado e o nome do assistente de pesquisa que forneceu o item lexical. Morfologia (mr) para indicar a morfologia do item lexical inserido no banco de dados. Picture (pc) refere-se às imagens de animais e plantas cujos nomes foram inseridos no banco de dados. Nota (nt) é utilizado para que o dicionarista possa incluir alguma informação a respeito do item por ele inserido. Finalmente se tem o marcador data (dt) para que se informe o dia em que aquele item foi inserido no banco de dados.

Uma questão importante foi a organização dos verbetes, considerando-se a forma ortográfica da entrada, o registro de sua transcrição fonológica, sua categoria gramatical, seguida da tradução correspondente nos seus diversos campos semânticos (vinculados especificamente à referência) e pragmáticos (relacionados especificamente ao seu uso corrente, ou ao exemplo de onde tal item foi retirado). Ficou decidido que a entrada de cada verbete seria efetuada por seu registro ortográfico em conformidade com a ordem alfabética. A transcrição fonética foi contemplada e aparece após cada entrada. Quanto à categoria gramatical, foram adotadas basicamente as normas padrão dos dicionários monolíngues. Nessa primeira etapa, não se fez a

²Para maiores informações sobre o programa pode-se consultar a homepage <http://www.sil.org/computing/toolbox/information.htm>

indexação de palavras derivadas, palavras compostas e expressões idiomáticas no corpo do verbete, uma vez que ainda não é possível se ter segurança quanto à ocorrência desses processos morfológicos na língua. Todas as palavras que pensamos ser compostas aparecem como se constituíssem um único item lexical.

Além de o linguista, ante a uma obra lexicográfica, trabalhar com uma vasta gama de aspectos teóricos dessa área de estudos, ele também precisa elaborar uma descrição de aspectos específicos de uma língua, como o fonético, fonológico, morfológico, sintático e semântico. Diante da possibilidade de se realizar uma descrição abrangente, pode-se também lançar mão dos aspectos práticos de uma tarefa como esta e pormenorizar cada questão suscitada por uma única entrada lexical. Na verdade, elaborar um dicionário de uma língua em descrição é um desafio, pois é uma tarefa que não se pode dizer concluída, tendo em vista a necessidade de se comprovar as hipóteses levantadas com base nos dados coletados em área indígena. A comprovação das hipóteses é sempre necessária junto aos falantes da língua em estudo.

Partindo-se do princípio de que o indivíduo constrói seu acervo lexical, categoriza-o e estrutura-o da maneira como ele vê o mundo, o ambiente onde vive espera-se que a maioria dos animais e dos vegetais presentes na nomenclatura referente à fauna e à flora da língua tapajúna seja classificada, em conformidade com seus aspectos físicos, que levam em conta aspectos como tamanho, forma, cor, periculosidade, dentre outros aspectos.

A taxonomia utilizada pelos tapajúna é diferente daquela utilizada em sociedades ocidentais. Desta forma, a classificação tradicional em mamíferos, répteis, insetos etc. não se aplica a essa língua. Todavia há termos genéricos como o nome *rop*, que dependendo do contexto pode ser: 'onça', 'cachorro', 'raposa' e mesmo 'gato maracajá', o qual não deixa de ser um tipo de onça.

Todas essas nuances devem ser consideradas no momento da elaboração de um dicionário. Assim essas acepções constam do verbete *rop* bem como há verbetes específicos para dados como os que são apresentados abaixo:

<i>rop</i> + <i>kamrêk</i>	=	'onça vermelha'
<i>rop</i> + <i>katêt</i> + <i>txira</i>	=	'onça pintada'
<i>rop</i> + <i>taka</i>	=	'cachorro do mato' ou 'onça maracajá'
<i>rop</i> + <i>tyk</i>	=	'onça preta'

Deste modo, entendemos que dicionários constituem repertórios léxicos organizados sistematicamente, geralmente, em ordem alfabética – o que facilita a consulta pelos falantes da língua, usuários de tal livro.

De acordo com Biderman (2001), um dicionário é um produto cultural. Para Vilela (1995, p. 78), um dicionário é definido como:

[...] o conhecimento genérico culturalmente partilhado por uma comunidade linguística e codificado no léxico, ou é a codificação desse saber, concebido de forma estática, em suporte papel ou eletrônico, arquivando esse saber [...]

Durante a execução do projeto de elaboração de um dicionário bilingue da língua tapajúna e da elaboração da obra lexicográfica em si, foi bastante recorrente a observação de polissemia e de homonímia. Deste modo, buscou-se definir esses temas a fim de se organizar os itens lexicais mais adequadamente.

De acordo com Crystal (2002, p.202-203), a polissemia é “termo usado na análise semântica para caracterizar um item lexical com uma variedade de significações diferentes, como por exemplo, manga = “parte da camisa”, “parte de um abajur”, [...]”. Sabe-se que grande parte do vocabulário de uma língua é polissêmica, por questões de economia linguística.

Um dos problemas teóricos para um linguista é como distinguir polissemia (uma forma com diversas significações) de homonímia (dois itens lexicais que têm a mesma forma fonológica). De acordo com Crystal, inúmeros critérios já foram sugeridos, dentre os quais a etimologia (em que os antecedentes de itens lexicais homônimos seriam formalmente distintos) e a proximidade da relação entre os itens em questão (as significações dos itens lexicais homônimos seriam mais distintas ou não-relacionadas). Todavia esses critérios envolvem problemas analíticos não tão simples de serem resolvidos, dos quais não trataremos neste artigo.

Todavia, Borba (2003) afirma que, pelo fato de ser o signo linguístico arbitrário, a polissemia explica a pluralidade de traços semânticos presentes no item lexical. Uma outra questão é que, muito embora a polissemia seja uma propriedade do item lexical, ela também pode apresentar um caráter discursivo, uma vez que o contexto e a situação em que um item lexical ocorre têm papéis importantes em sua conceituação.

Uma palavra como *ngô* em tapajúna, cujo sentido é originalmente água, pode fazer referência a rio e a líquidos, em geral. Do mesmo modo, a palavra *kà*, cujo primeiro sentido parece ser ‘pele’, pode referir-se a ‘seio’, ‘casca de árvore’, ‘roupa’, e ‘invólucro’ de modo geral, dentre outras significações, num claro processo de extensão de significado. Tal fato é verdadeiro para outras línguas geneticamente aparentadas ao tapajúna, como, por exemplo, o mebengôkre e o parkatêjê.

Quando um dicionarista considera o conceito de homonímia em seu trabalho, cada novo significado constituirá uma entrada independente no dicionário. Esse volume grande de acepções pode prejudicar a organização da obra. Uma possível solução é observar entre os itens lexicais polissêmicos, as acepções que tenham um traço semântico comum. Então, essas seriam apresentadas na obra lexicográfica em sequências numeradas dentro de uma mesma entrada no dicionário.

Para a elaboração de um dicionário, afirma Dapena (2002, p.78)

“[...] um dicionário deve incluir em sua parte técnica as bases teórico-linguísticas que irão fundamentá-lo, os critérios de eleição de suas entradas, o procedimento de escolha do material lexicográfico a ser analisado e estudado, assim como o estabelecimento das fontes de onde foi coletado o material que estará sendo utilizado, os métodos e características da redação, enfim, a apresentação tipográfica da obra”.

Um dicionário pode ser utilizado pela comunidade como material pedagógico, além de ser um material linguístico, cuja finalidade é ampliar o conhecimento que se tem da língua. Para o estudo da língua tapajúna, foi estabelecida uma proposta de ortografia preliminar, junto à comunidade. Materiais de leitura e de outras disciplinas como matemática também foram produzidos durante os três anos em que estive trabalhando com esse povo para que os alunos indígenas pudessem se apropriar e sedimentar a proposta ortográfica sistematizada. Além disso, outros materiais pedagógicos escritos na língua, como a apresentação preliminar desse dicionário, já estão servindo como suporte para o professor de área indígena.

Uma das formas de se ajudar na manutenção de uma língua é contribuir para que ela deixe de ser ágrafa e fazê-la objeto de estudo na escola da comunidade ao lado da língua portuguesa. Por esta razão, a atividade de elaboração de um dicionário bilíngue foi uma das formas pensadas como mais abrangentes e favoráveis à descrição, documentação e manutenção dessa língua indígena brasileira, que também proporcionou, aos envolvidos na tarefa ampla de elaborar um dicionário bilíngue, uma reflexão acerca dos vários parâmetros envolvidos no estudo do léxico de uma língua humana.

REFERÊNCIAS BIBLIOGRÁFICAS

- BIDERMAN, M. T. C. *Teoria Lingüística*. São Paulo, Martins Fontes, 2001.
- BIERWISCH, M. & SCHREUDER, R. From concepts to lexical items. *Cognition*, v. 42: 23-60, 1992.
- BORBA, Francisco da Silva. *Organização de dicionários: uma introdução à lexicografia*. São Paulo: Editora UNESP, 2003.
- CAMARGO, Nayara da S. Elaboração de um dicionário bilíngue tapajúna-português. *Estudos linguísticos*, 37 (1): 73-82. São Paulo, jan.-abr. 2008.
- CAMARGO, Nayara da S. Montagem de um banco de dados sobre a fauna e a flora da língua Tapajúna. Trabalho de Conclusão de Curso (Graduação em Licenciatura em Letras). Universidade Federal do Pará, 2005.
- CRYSTAL, David. *Dicionário de lingüística e fonética*. Tradução e adaptação [da 2ª ed. inglesa ver. e ampl., publicada em 1985], tradução de Maria Carmelita Pádua Dias. Rio de Janeiro: José Zahar Editora, 2000.
- DAPENA, José-Alvaro Porto. *Manual de Técnica Lexicográfica*. Ed. Arco/ Libros, S. L., 2002.
- FERREIRA, Marília. 2005. Descrição de aspectos da variante étnica usada pelos parkatêjê. *Revista D.E.L.T.A*, 21:1. São Paulo: EDUC, 2005, p.1-165.
- MEL'CUK, A. *Dependency syntax: Theory and Practice*. SUNY, 1988.
- RODRIGUES, Cíntia Karla Coelho. Descrição dos prefixos relacionais em tapajúna-goronã. Trabalho de Conclusão de Curso (Graduação em Licenciatura em Letras). Universidade Federal do Pará, 2006.
- RODRIGUES, Cíntia Karla Coelho. Análise histórico-comparativa das línguas Tapajúna, Suyá e Mebengôkre. Dissertação de Mestrado em Letras. Universidade Federal do Pará, 2009.
- VILELA, M. *Léxico e gramática*. Coimbra: Almedina, 1995.

Recebido em agosto de 2009
Aprovado em novembro de 2009