

Estudo das Línguas Indígenas Brasileiras

Trabalhando com dicionários em comunidades indígenas: a importância da contrapartida

Nathalie Pires Vlcek

RESUMO: O trabalho de pesquisa com línguas indígenas envolve não somente as atividades de descrição e documentação, mas também um comprometimento da parte do pesquisador com as necessidades da comunidade com a qual se realiza a pesquisa. Não se trata apenas de conciliar três trabalhos paralelos, mas de fazer com que esses três pontos de partida formem um tripé de sustentação para um trabalho cooperativo entre pesquisador e comunidade. Tomando como exemplo o início da produção de dicionários temáticos bilíngues Tuyuka-Português, demonstro como descrição, documentação e contrapartida são três atividades não apenas complementares como também interdependentes.

PALAVRAS-CHAVE: dicionários temáticos bilíngues, Tuyuka, pesquisa participativa, documentação linguística, lexicologia

ABSTRACT: Working with indigenous languages is not about just describing and documenting them, but also about assuming responsibilities in keeping with the language community's needs. These three goals should not be viewed as parallel efforts, but as a tripod foundation for cooperative work between the researcher and the community. Using examples based on my own experience with the initial stages of creating bilingual Tuyuka-Portuguese theme dictionaries, I will demonstrate how description, documentation and community-based activities are not only complementary but interdependent activities.

KEYWORDS: bilingual theme-based dictionaries, Tuyuka, collaborative research, language documentation, lexicology

Introdução

Ao trabalhar com línguas indígenas o pesquisador de campo se envolve com duas atividades complementares: a descrição e a documentação (FARGETTI, 2013). Com esses objetivos cheguei eu à língua Tuyuka - uma língua Tukano Oriental falada por parte do seu povo homônimo. Meu trabalho, porém, restringe-se às

comunidades brasileiras abrangidas pela área da escola Tuyuka, no Alto rio Tiquié, noroeste amazônico, região de fronteira com a Colômbia.

No início da minha experiência junto aos Tuyuka, em alguns momentos me sentia como uma grande administradora de tempo e recursos dividindo-me em uma tríade de atividades. Por um lado, cabia a mim gerar um trabalho acadêmico de análise de diversos aspectos da língua Tuyuka; por outro lado, também era esperado de mim a produção de um acervo de dados primários duradouro, acessível e organizado, capaz de servir de corpus de análise para futuras pesquisas, bem como de material de registro da língua (GIPPERT, HIMMELMAN, MOSEL, 2006; DIXON, 2010; AUSTIN, SALLABANK, 2011; STENZEL, 2008). Mais do que isso, também sabia eu que uma função importante do meu trabalho seria promover benefícios reais à comunidade a depender das suas necessidades - o que a literatura chama de 'contrapartida' (STENZEL, *inédito*).

A contrapartida, na minha mente ainda inexperiente com trabalho de campo, era compreendida como uma espécie de troca de habilidades e recursos, em que o pesquisador levaria à comunidade seu saber técnico para auxiliar o alcance de objetivos determinados pela comunidade. O que descobri ao longo de um pouco mais de um ano trabalhando ao lado do povo Tuyuka na produção de pequenos dicionários temáticos, é que a contrapartida na verdade é um processo muito mais complexo e rico do que uma simples retribuição. A contrapartida na verdade significa uma integração de esforços, uma superação dos objetivos iniciais e isolados de um pesquisador em prol de um trabalho integrado com a comunidade. Durante um processo de documentação, as atividades se complementam e se impulsionam de tal forma que não é mais possível dizer para quem serão os benefícios de uma determinada atividade, já que todas elas – se o processo estiver dando certo – são boas para todos os envolvidos.

2. O papel dos dicionários temáticos no processo de documentação

O final da década de 80 foi um período de virada na luta pela preservação e documentação da diversidade linguística do mundo (STENZEL, 2008, p. 66). Somando-se à divulgação e conscientização da causa, surgiram políticas públicas de defesa da cultura e língua dos povos indígenas. O reconhecimento dos direitos de preservação de suas línguas e culturas é assegurado no Brasil a partir da

Constituição de 1988 e em atos legislativos como a Lei de Diretrizes e Bases da Educação de 1996 (STENZEL, 2008, p. 68). A partir destas novas políticas, grupos indígenas reconheceram novas oportunidades e meios para fazer valer seus direitos e valorizar suas línguas e culturas. Algumas histórias bem sucedidas (FRANCHETTO, 2011) fizeram com que outros grupos se inspirassem e fizessem o mesmo. Já entre os linguistas, diante desta nova demanda de revitalização linguística por parte dos grupos indígenas, o modelo 'clássico' ('Boasiano') de documentação - pautado em compor uma gramática descritiva, um dicionário e uma pequena coletânea de texto - começa a ser expandido. O novo modelo de pesquisa de documentação incorpora novos elementos, como a essencial participação das populações nas decisões e no processo de documentação de suas línguas e culturas. A nova documentação linguística concentra grande parte dos seus esforços nos dados primários: coleta, acessibilidade, reprodução, armazenamento e anotação (STENZEL, 2008, p.71,72).

Não é o caso de se abandonar a descrição, mas sim de torná-la um produto científico que constitua também um acervo de reflexões e análises dos falantes a respeito de sua própria língua. Trata-se de aliar ao trabalho descritivo uma preocupação em manter ao seu lado um acervo que permita não só o registro da língua, mas sua possível análise por outros pesquisadores e o seu acesso a futuros membros da comunidade. Para isso, é imprescindível o pesquisador estar em campo – e estar equipado com gravadores e cartões de memória, além de ferramentas de anotação. A partir de 2000 surgem centros de recursos (como *Hans Rausing Endangered Languages Project - HRELP*) para pesquisa de documentação linguística cujo foco passa a ser a gravação e o armazenamento digital de dados primários. Esses agentes financiadores de pesquisa não foram concebidos para durarem por tempo indefinido e, mesmo tendo conseguido prover fundos por mais tempo do que era previsto inicialmente, hoje infelizmente estão se fechando. Com isso, o foco dos pesquisadores tem mudado mais uma vez em direção à capacitação de equipes indígenas para a atividade autônoma da documentação linguística.

Ao dar início a um trabalho de documentação linguística no interior de uma comunidade, é bastante produtivo, portanto, que se desenvolvam oficinas de trabalho conjunto entre o linguista e os membros da comunidade. Um dos primeiros objetivos das oficinas iniciais é justamente estabelecer quais seriam as prioridades

para aquela população. Há comunidades para as quais o desenvolvimento de uma ortografia é uma questão central, o que não aconteceu com minha experiência entre os Tuyuka, que já possuíam uma história de vinte anos de prática ortográfica devido à criação da escola diferenciada em 2000, com colaboração da Federação das Organizações Indígenas do Rio Negro (FOIRN) e do Instituto Socioambiental (ISA) (CABALZAR, 2012). A prioridade de nossos esforços correspondeu às necessidades da escola Tuyuka naquele momento: a produção de materiais didáticos. Atendendo então a demandas da comunidade, no nosso caso, uma importante contrapartida da minha pesquisa seria o desenvolvimento dos dicionários temáticos multimídia. Com dicionários, não somente materiais seriam criados ao fim do projeto, mas a atividade de desenvolvimento em si já é trabalhada na escola e na comunidade com momentos de pesquisa, integração e valorização da língua e da cultura.

Como vimos, no modelo de documentação linguística atual, não se deve tratar o dicionário de uma língua como um simples apêndice a uma gramática descritiva, como muitas vezes se fazia no modelo clássico. Os dicionários são importantes fontes de informação sobre a língua e trazem em si muitas reflexões de outras naturezas, como história, filosofia, biologia e etnografia.

Lembremo-nos que a língua não é um objeto abstrato, pelo contrário, está altamente inserida no contexto histórico e social da sua cultura (SEKI, 2012, p. 14). Desta forma, ao pesquisador de campo é dada a abrangente função de articular objetivos acadêmicos com objetivos sociais. Em termos da realidade Tuyuka, isso se traduz em duas grandes demandas. Se por um lado há um anseio por parte da comunidade em trabalhar políticas de auto-afirmação de sua língua através da produção de materiais e de pesquisas metalinguísticas, por outro lado há também uma necessidade crescente do domínio do português por diversos indivíduos em busca de maiores oportunidades profissionais.

3. Tomadas de decisão que se desvelam na produção dos dicionários temáticos

A lexicografia, prática de construir dicionários, exige um bom entendimento de lexicologia, que consiste em compreender os itens lexicais em sua constituição morfológica e fonológica (FARGETTI, 2013).

Acerca da organização de dicionários há de se pensar na macroestrutura e na microestrutura. A principal questão de macroestrutura está na decisão de organizar os verbetes em ordem alfabética ou em campos semânticos. Apesar de a escolha da organização em campos semânticos ser bastante propícia para ampliação de corpus e contraste de sentidos, a organização por campo semântico traz questões relevantes (SEKI, 2012, p. 18). Imaginem meu choque em uma primeiríssima viagem a campo ao perceber que ao invés de categorias familiares de nossos dicionários brasileiros como “mamíferos, plantas, lazer”, deparei-me com outras, tais quais “doenças e curas” e “trabalhos de homens”, sugeridos pelos membros da comunidade Tuyuka. Partindo da observação de que culturas diferentes categorizam seu respectivo universo referencial de forma particular, percebe-se que a organização dos verbetes por campos semânticos seria uma para o falante de português e outra completamente diferente para o falante de Tuyuka. Para resolver este problema, o ideal é imprimir coleções distintas de dicionário para cada língua de origem. Uma outra opção é a produção de dicionários em formato multimídia, como CDs. Há programas de computador, como o *Lexique Pro*, que permite esta dupla categorização de verbetes dependendo a língua fonte selecionada, mas por detrás deste facilitador está a realidade das difíceis condições de se manter uma base acessível de equipamentos em locais remotos e altamente úmidos. No caso da escola Tuyuka, estamos desenvolvendo uma pequena DVD-teca que irá contar com acervos de CDs e DVDs executáveis em leitores para tais itens, que exigem menor consumo de energia para recarregar baterias do que computadores, além de serem imunes a vírus e mais resistentes a poeira e umidade.

Já no que tange à microestrutura, a maior decisão está na escolha entre entradas (itens lexicais) versus subentradas (derivações previsíveis destes itens pela gramática da língua). A inclusão de palavras derivadas pode gerar redundância de entradas, mas a sua exclusão omitiria restrições possíveis na derivação. Em Tuyuka optamos por múltiplas entradas. Por exemplo, temos a existência de classificadores, como <-wi> para ‘edificação’, que aparece em verbetes como a palavra ‘maloca’, que em Tuyuka pode ser expressa por <basawi¹> ou <yawi²> dependendo do que irá se realizar nela, celebração ou uma refeição coletiva,

¹ <basa-> ‘celebrar’ + <-wi> ‘classificador de edifício’

² <ya-> ‘comer’ + <-wi> ‘classificador de edifício’

respectivamente. Apesar de ser econômico inserir somente o classificador e não as duas palavras distintas para maloca, não estaria contemplando esta importante construção para a cultura Tuyuka, nem tampouco garantindo aos leitores falantes de português saber os termos que podem ser compostos por este classificador, que não pode se combinar com toda e qualquer raiz verbal.

Ao idealizar um dicionário, também pensamos imediatamente em qual a natureza da informação que devemos incluir. No caso de dicionários bilíngues em comunidades indígenas a noção de tradução é problemática, pois traz à tona “a problemática que essas línguas, tão distintas estruturalmente de línguas indo-européias e vinculadas a culturas igualmente distintas, colocam aos pesquisadores” (SEKI, 2012, p. 13).

O pesquisador e a comunidade estão diante, nos casos de dicionários bilíngues como Tuyuka-Português, de uma grande falta de isomorfismo entre os léxicos dessas duas línguas, que são altamente distintas em sua tipologia. Embora para algumas palavras exista uma equivalência, pode acontecer de um item lexical em Tuyuka significar vários em português. Pode também ser o caso de diversos itens lexicais em Tuyuka corresponderem a um só item em português (como <opare>, ‘carregar’ com o ombro e <omare>, ‘carregar’ com as costas). Há casos mais complexos, em que simplesmente o campo semântico do Tuyuka denota itens diferentes do campo semântico do português, como acontece com os termos de parentesco, em que, por exemplo, o termo <pakomaku> teria como tradução primeira ‘o filho da irmã da mãe’ (CABALZAR, 2009).

Outro caso recorrente de não correspondência biunívoca vocabular está nos elementos que descrevem o mundo natural, como plantas e animais, em que muitas vezes é preciso recorrer à ajuda de um biólogo para encontrar o nome científico do referente. Claro que mesmo na ausência de um profissional da área vale à pena manter a pesquisa em andamento, e esta é uma das funções das fotos e das definições neste tipo de verbete, garantindo a possibilidade de o trabalho ser continuado no futuro. Além disso, a catalogação de elementos naturais pelos membros da comunidade muitas vezes gera um movimento de revitalização e valorização de conhecimentos da cultura oral tradicional, dada a necessidade de os mais jovens (normalmente envolvidos nos projetos de documentação pelo atrativo da tecnologia dos instrumentos) consultarem os conhecedores mais velhos para

resgatar saberes que estão sendo esquecidos. No caso do dicionário Tuyuka, o tema “remédios e doenças” promoveu uma série de entrevistas e ainda é tema de debate nas oficinas, visto que o saber das plantas medicinais estava praticamente perdido pelos habitantes das comunidades Tuyuka do Brasil, e o registro deste conhecimento, hoje em posse dos Tuyuka habitantes da Colômbia, fez com que este conhecimento pudesse ser mais preservado para as gerações futuras.

Não é só no caso dos elementos do mundo biológico que o dicionário deve ser visto pelo pesquisador e pela comunidade como um eterno processo em andamento, e não como uma etapa a cumprir para se chegar a um produto final. O dicionário guarda também entradas de unidades menos lexicais na língua e mais funcionais, tais como partículas modais, aspectuais, evidenciais etc. Até o momento, ainda não decidimos como essas entradas aparecerão nos nossos dicionários, mas as muitas descobertas a este respeito estão sendo anotadas. Se até hoje ainda descobrimos construções ou repensamos estruturas em línguas amplamente estudadas como o inglês e o francês, imaginem quantas coisas não continuaremos descobrindo em línguas cuja documentação ainda está em fase inicial.

4. A entrada de unidades lexicais é fonte e é produto de análise das categorias funcionais

Os temas para o dicionário Tuyuka foram selecionados pela comunidade, que é o principal público alvo dos dicionários. Os diferentes temas foram divididos entre os bolsistas do projeto, e com ajuda da antropóloga Flora Cabalzar, estes jovens começaram a mobilização da pesquisa e seleção vocabular. Com a minha participação, foram feitas discussões acerca da natureza da palavra e de seus elementos componentes, tais como a distinção entre raízes e sufixos, muito cara ao Tuyuka, uma língua polissintética-aglutinante de padrão OV. Os pesquisadores indígenas registravam as palavras levantadas por pesquisa, junto com algumas informações adicionais para o verbete, no programa *WeSay*. Este programa – de interface simples, produzido para membros da comunidade sem treinamento prévio – permite exportação para o programa *Flex* – desenvolvido para linguistas comporem uma vasta e complexa base de dados que dialoga com programas de anotação de dados primários e com programas de dicionário. Junto com os pesquisadores indígenas, revisávamos e compilávamos no programa *Flex* os

exemplos obtidos no *WeSay*, os áudios e as imagens, mapeando informações acerca do verbete. Ao fim de tudo, a base *Flex* pode ser exportada para programas de gerar dicionários, como o *Lexique Pro*, ou simplesmente pode ser impressa para uso imediato na comunidade.

Muito embora eu acredite que nossa maior conquista foi o alto nível de treinamento de uma equipe de documentação linguística Tuyuka, temos já finalizado também um primeiro tema de dicionário pronto para publicação, seja em mídia de CD ou livro. Trata-se do tema <umũ padere>, ‘trabalho dos homens’, e suas trinta entradas lexicais, conta com cada verbete em ortografia Tuyuka, sua transcrição fonética, tradução, imagem, anotações pertinentes e áudio, além de pelo menos um exemplo, por entrada do dicionário, da palavra em uso.

Reparem que a iniciativa dos dicionários temáticos foi a principal contrapartida minha como linguista da equipe do projeto *Documentação Cultural e Linguística do Tuyuka*, do IPHAN, proposto e coordenado pela antropóloga Flora Cabalzar, pelo Instituto Socioambiental, não estando diretamente relacionado à minha tese junto à UFRJ. Porém, além de engrandecer minhas relações com os membros da comunidade através de oficinas e semanas de árduos trabalhos, o desenvolvimento dos dicionários temáticos revelou-se uma fonte de pesquisa para descrição da língua capaz de evidenciar muitas características de difícil elicitación.

Tomemos a expressão da categoria gramatical de ‘evidencialidade’ como exemplo. Tendo revisado a literatura sobre a língua, sabia eu antes de ir a campo que a expressão da evidencialidade em Tuyuka constitui um sistema gramatical com morfemas sufixais dispostos num paradigma e cuja presença é obrigatória em verbos de sentenças principais ou independentes (BARNES, 1984). O sistema de expressão de evidencialidade, que trata da relação entre o falante e a fonte da informação, é partido em cinco categorias na língua Tuyuka. A categoria VISUAL é utilizada quando o falante é testemunha visual ou participante da ação que enuncia, tendo presenciado a mesma visualmente. A categoria NÃO-VISUAL é usada quando o falante percebe aquela informação através de um sentido que não o visual, seja ele o auditivo, o olfato, ou outro. A categoria INFERENCIAL é usada quando o falante infere aquela informação através de resultados consequentes da mesma. Já marcadores da categoria REPORTADO aparecem quando uma outra pessoa deu a informação ao

falante, e o morfema da categoria PRESUMIDO é usado quando o falante supõe aquele evento baseando-se em seu conhecimento do mundo.

Vejamos um exemplo de EVIDENCIAL VISUAL (01), usado para descrever estados ou eventos que o falante viu ou está vendo, incluindo aqueles em que é o ator.

- (01) *wasorĩ opag#tia*
ũãsõ-rĩ opa-g# ti-a
caibro-PL³ carregar.no.ombro-SGM fazer-VIS.1/2.PRES
'Eu estou carregando caibros'.

Em (01), o verbo *opa* 'carregar no ombro' principal é nominalizado pelo morfema *-g#*, indicando um sujeito singular masculino, e o sufixo de evidencialidade *-a* se liga ao verbo auxiliar *ti* 'fazer'. Repare que na glosa do morfema evidencial consta o valor temporal PRES compreendido dentro de uma construção de sentido progressivo atestada por Barnes (1984) e Malone (1988). Esta construção é utilizada em eventos em que o falante está presenciando a ação no momento em que esta é reportada, e a noção de 'presente' do evento é inferida a partir do fato de que o falante obteve a informação no mesmo tempo da fala.

Em outras palavras, alguns valores gramaticais dos evidenciais – como pessoa, número e gênero – concordam com o sujeito gramatical da oração, estando ligadas ao evento. Mas outras categorias – como tempo e evidencialidade – estão ligadas à evidência em si. Logo, as sentenças com evidenciais devem ser interpretadas sob a luz de Timberlake (2007) e Payne (1997), vistas não como expressão de tempo absoluto, mas sim como índice da relação temporal entre o evento e o momento da fala. Portanto, o valor "passado" da glosa do sufixo evidencial *-wi* em (02) indica que o falante obteve a informação num momento anterior à elocução.

- (02) *y#bai botari seg#wawĩ*
d#u⁴-bai bota-ri se-g# ~wa-wi
1SG-irmão menor esteio-PL coletar-SGM ir.para.longo-VIS.3SGM.PST
'Meu irmão foi pegar esteios'.

³ Glosas: PL: PLURAL; SG: SINGULAR; M: MASCULINO; 1,2,3: PESSOAS DO DISCURSO; PST: PASSADO; PRES: PRESENTE; VIS: CATEGORIA EVIDENCIAL VISUAL

⁴ A confirmação de análise do morfema *d#u* '1SG' com possível *status* de clítico ainda está dependendo de futuras investigações.

Em (02), temos uma construção de finalidade, formada pela raiz verbal *se* ‘coletar’, nominalizado com marca do sujeito *-g#* (singular masculino) complemento do verbo *~wa* ‘ir para longe’ sufixado pelo evidencial visual *-wi*, que também concorda com o sujeito.

Observe que tanto a frase em (01) como a frase em (02) foram criados pelos pesquisadores indígenas como exemplos de unidades lexicais em uso para o dicionário. Imaginem a dificuldade de se elicitarem padrões de elementos tão veiculados a situação real de fala como os evidenciais. Além disso, outras informações gramaticais se depreendem acerca não somente de partículas funcionais, mas das construções de evidencialidade como um todo na língua.

Em (01) fica clara uma noção de finalidade produzida pela construção de raiz verbal nominalizada complemento do verbo ‘ir’ marcado de evidencialidade, mas o trabalho com este exemplo na construção do dicionário trouxe ainda outras surpresas ricas para análise. O falante em (02) viu que o irmão saiu para fazer esta ação, mas não o viu *fazendo* a mesma. Esta interpretação se verificou quando o grupo de consultores desenhou apenas um homem andando para ilustrar a palavra *<botari>*, visto que não achavam correto desenhar os esteios, já que estes não estavam presentes no evento da elocução em (02).

Uma breve conclusão:

Os trabalhos de documentação, descrição e de realização de contrapartidas para a comunidade não devem ser vistos como atividades paralelas, mas sim como momentos complementares de um mesmo processo. Mais do que isso, são atividades inter-dependentes, visto que é impossível fazer um bom trabalho de documentação que não envolva a reflexão e o engrandecimento da comunidade, bem como é igualmente inviável auxiliar a comunidade em suas demandas sem debruçar-se em diferentes esforços de análise.

Referências:

1. AUSTION, P., & SALLABANK, J. (2011). *The Cambridge Handbook of Endangered Languages*. New York: Cambridge University Press.
2. BARNES, J. (1984). Evidentials in the Tuyuca Verb. *International Journal of American Linguistics*, pp. 225-271.

3. CABALZAR, A. (2009). Filhos da Cobra de Pedra - Organização social e trajetória tuyuka no rio Tiquié (noroeste amazônico). São Paulo, Rio de Janeiro: Unesp/ISA, Nuti.
4. CABALZAR, F. D. (org.) (2012). Educação Escolar Indígenas do Rio Negro: relatos de experiências, lições aprendidas. São Paulo, São Gabriel da Cachoeira: Instituto Sócio Ambiental, Federação das Organizações Indígenas do Rio Negro.
5. DIXON, R. (2010). Basic Linguistic Theory. Volume 1: Methodology. Oxford: Oxford University Press.
6. FARGETTI, C. M. (2013). Pesquisa de línguas indígenas - questões de método. In: F. A. Del Ré, F. Komesu, L. Tenati, & A. Vieira (orgs.), Estudos linguísticos contemporâneos. Diferentes olhares. (pp. 115-). Araraquara: Cultura Acadêmica.
7. FRANCHETTO, B. (org.). (2011). Alto Xingu: uma sociedade multilíngue. Rio de Janeiro: Museu do Índio - FUNAI.
8. GIPPERT, J., Himmelmann, N., & Mosel, U. (org.) (2006). Essentials of language documentation. Hague: Mouton de Gruyter.
9. MALONE, T. (1988). The Origin and development of Tuyuca evidentials. *International Journal of American Linguistics* 54 , pp. 119-140.
10. PAYNE, T. (1997). Describing morphosyntax: a guide for field linguists. Cambridge: Cambridge University Press.
11. SEKI, L. (2012). Discutindo dicionários bilíngues: o caso Kamaiurá. In: C. M. Fargetti (org), Abordagens sobre o léxico em línguas indígenas - I encontro do grupo LINBRA (pp. 13-36). Campiinas: Curt Nimendajú.
12. STENZEL, K. (2008). Novos horizontes da documentação linguística no Brasil. *Revista de Estudos e Pesquisas - FUNAI* - vol. 5 n.1/2 , 49-99.
13. STENZEL, K. (*inédito*). The pleasures and pitfalls of a 'participatory' documentation project: an experience in northwestern Amazonia. *Language Documentation & Conservation*.
14. TIMBERLAKE, A. (2007). Tense, aspect and mood. In T. Shopen (org.), *Language typology and syntactic description*, Vol.3 (pp. 280-333). Cambridge: Cambridge University Press.