

Para uma crítica lexicográfica das microestruturas da Munduruku Word List

For a lexicographic critique of the microstructures of Munduruku Word List

Renan do Socorro dos Santos BORGES (UFPA)
renanborges100@gmail.com

Jorge Domingues LOPES (UFPA)
jdlopes@ufpa.br

Recebido em: 13 de fev. de 2018.
Aceito em: 02 de jun. de 2018.

BORGES, Renan do Socorro dos Santos; LOPES, Jorge Domingues. Para uma crítica lexicográfica das microestruturas da Munduruku Word List. **Entrepalavras**, Fortaleza, v. 8, n. 2, p. 442-469, maio/ago. 2018.

Resumo: O presente trabalho, que descreve e analisa o tratamento lexicográfico dado às microestruturas da *Munduruku Word List*, de Crofts (1986a, 1986b, 1986c, 1986d), baseia-se nas ciências do léxico – lexicologia, lexicografia e metalexicografia –, e parte, sobretudo, de reflexões metalexicográficas, tomando como pressupostos teórico-metodológicos Haensch (1982), Welker (2004) e Atkins e Rundell (2008), para reflexões de teoria lexicográfica, e em Lopes (2014), para a descrição de microestruturas segundo um esquema de descrição lógica que apresenta os elementos constituintes das microestruturas. O *corpus* constitui-se do conjunto de entradas iniciadas pela letra A da *Munduruku Word List*. As análises mostram que, nessa parte do material, existem 12 microestruturas básicas diferentes, com elementos variantes no lema, no equivalente e exemplo, além da presença de notas do autor referentes ao lema e ao equivalente. Outros recursos que se encontram no material são traduções literais e remissivas.

Palavras-chave: Mundurukú. Material lexicográfico. Lista de palavras. Microestrutura.

Abstract: This work describes and analyzes a lexicographic material of a Brazilian indigenous language, focusing on the lexicographic treatment given to the microstructures of *Munduruku Word List* (CROFTS, 1986a, 1986b, 1986c, 1986d). This work was based on the science of the lexicon – lexicology, lexicography and metalexicography – since the work starts, above all, from metalexicographic reflections, taking as theoretical and methodological assumptions Haensch (1982), Welker (2004) and Atkins and Rundell (2008) for reflections of lexicographic theory; and Lopes (2014), for the description of microstructures according to a logical description scheme that presents the elements of the microstructures. The *corpus* consists of all entries beginning with the letter A of *Munduruku Word List*, and the analyzes show that, in this part of the material, there are 12 different basic microstructures, with variant elements in the headword, in the equivalent and example, and presence of notes of author referring to the headword and its equivalent. Other resources that are found in the material are literal translations and cross-references.

Keywords: Munduruku. Lexicographic material. Word List. Microstructure.

Introdução

Considerando a produção lexicográfica no que tange às línguas indígenas brasileiras, sobretudo no século XX, fazemos uma abordagem descritiva com reflexões e críticas metalexicográficas acerca da *Munduruku Word List*, um material produzido por Marjorie Crofts¹ sobre a língua Mundurukú, família Mundurukú, tronco Tupi (RODRIGUES, 1970).

Decorridos mais de 30 anos de sua produção, percebemos que esse material não somente se configura como importante repositório de dados linguísticos, mas também pode ser, ele próprio, objeto de descrição e análise do ponto de vista metalexicográfico. Nesse sentido, tal lista constitui o *corpus* aqui analisado, em que buscamos compreender como foram gerenciadas as escolhas para a compilação do material e que orientações tipológicas e teóricas guiaram o processo de sua construção lexicográfica.

Assim, o presente estudo busca fazer uma descrição da microestrutura lexicográfica da obra *Munduruku Word List* (CROFTS, 1986), para compreender sua estruturação interna e também as escolhas feitas pela autora ao produzir o referido material.

Como não temos a pretensão de fazer uma análise exaustiva, a ponto de abarcar todas as entradas de todos os volumes do material, a análise considera o conjunto das entradas iniciadas pela letra A, que corresponde à quantidade de 600 verbetes. A presente pesquisa, de cunho quanti-qualitativo, faz uso de concepções metalexicográficas para alcançar os objetivos propostos.

¹ Crofts fez parte do grupo de missionários ligado ao Summer Institute of Linguistics, realizando pesquisas de línguas indígenas brasileiras entre as décadas de 1960-1980.

Compreendendo lexicografia

A lexicografia, segundo Haensch, é “[...] todo domínio da descrição léxica que se concentre no estudo e descrição dos monemas e sinmonemas² individuais dos discursos individuais, dos discursos coletivos e dos sistemas linguísticos individuais [...] e coletivos”³ (1982, p. 93). Esse conceito acerca das finalidades com que fazemos lexicografia abrange os discursos individuais e coletivos, pois a comunicação é feita por meio deles nos sistemas linguísticos que se inserem. A lexicografia é, pois, a descrição do léxico dos discursos de comunidades linguísticas segundo Haensch, considerando que ligamos a lexicografia, principalmente, à construção de dicionários. Com base nessa concepção, Haensch (1982) chama de teoria da lexicografia os métodos científicos para reflexão e análise do léxico.

Em outra perspectiva, Welker (2004, p. 11) define o termo lexicografia em dois sentidos: “[...] o estudo de problemas ligados à elaboração de dicionários, a crítica de dicionários, a pesquisa da história da lexicografia, a pesquisa do uso de dicionários e ainda a tipologia”.

Parece-nos também adequado, para os fins deste trabalho, esse conceito apresentado por Welker (2004), tanto na perspectiva de produção de obra de descrição léxica (como lexicografia) quanto no campo teórico científico (como metalexicografia). A respeito da lexicografia teórica ou metalexicografia, Atkins e Rundell (2008) declaram que, para os metalexicógrafos, o próprio dicionário é seu objeto de estudo, constando então a metalexicografia como o estudo sobre os dicionários em geral. O uso desse termo será feito, portanto, sempre que nos referirmos à teoria de qualquer obra de cunho lexicográfico.

Por fim, é importante destacarmos que essas ciências têm como objeto o léxico em si, que pode ser entendido como “[...] o conjunto de monemas do discurso individual, do discurso coletivo, do sistema linguístico individual ou do sistema linguístico coletivo” (HAENSCH, 1982, p. 91),⁴ ou seja, o léxico é descrito como a totalidade de unidades

² Segundo Dubois *et al.* (1993, p. 471, grifos dos autores), “Na terminologia de A. Martinet, o termo *monema* é a unidade significativa elementar. Pode ser uma palavra simples, um radical, um afixo, uma desinência”. Já sinmonemas, de acordo com Werner (1982, p. 71), são “significantes que se compõem de vários monemas”.

³ “[...] todo dominio de la descripción léxica que se concentre en el estudio y la descripción de los monemas y sinmonemas individuales de los discursos individuales, de los discursos colectivos y de los sistemas linguísticos individuales y [...] colectivos.”

⁴ “[...] el conjunto de monemas del discurso individual, del discurso colectivo, del sistema linguístico individual o del sistema linguístico colectivo”.

que são produzidas pelos falantes de qualquer sistema linguístico em seu discurso, seja individual, seja coletivo, tal como o léxico do falante ou de toda a comunidade a que pertence. É nesse contexto que se inscreve o material lexicográfico aqui analisado.

Microestrutura lexicográfica: conceitos necessários

A partir do domínio da lexicografia, faz-se mister introduzirmos alguns conceitos relacionados à microestrutura dos verbetes de material lexicográfico monolíngue ou bilíngue: o primeiro deles é a própria noção de microestrutura. Por microestrutura (artigo, verbete), consideramos “[...] o conjunto das informações ordenadas de cada verbete após a entrada” (REY-DEBOVE, 1971, p. 21 *apud* WELKER, 2004, p. 107). Compreende, portanto, toda a organização “horizontal”⁵ utilizada em uma obra lexicográfica e tem um padrão lógico de estruturação. Na terminologia lexicográfica, comumente encontramos microestrutura como sinônimo de entrada, artigo, verbete; dependendo do autor e da teoria.

Cada microestrutura contém a “cabeça” e o “corpo” do verbete. Na cabeça do verbete, encontramos a palavra-entrada, ou usualmente conhecida como *lema*. Este, por sua vez, engloba o conjunto de paradigmas de uma palavra, isto é, reconhecemos uma forma dita canônica para representar todas as formas de flexão possíveis para tal palavra. Como consequência, conhecemos como lematização o uso do lema para entradas em dicionários, que, de acordo com Biderman (2001, p. 84), “[...] a lematização é o processo de agrupar todas as variantes morfológicas de uma palavra sob uma única entrada (lema) e a tarefa correlata de separar os homônimos”. Conforme o critério de lematização e a língua em que produzimos a obra lexicográfica, as escolhas são diferentes, como o caso do Português em que usamos o infinitivo para os verbos e masculino singular para os nomes, dentre outras escolhas. Para Haensch (1982, p. 462), “[...] o lema é a parte enunciativa de um artigo, cujo objeto é a descrição e explicação do lema”⁶. O que se segue ao lema é a definição, parte descritiva-explicativa; nos dicionários monolíngues, a definição descreve o que significa, explicando-o e, nos bilíngues, costuma dar o seu equivalente na língua-alvo, ou seja, na língua que está a parte definidora. O conjunto que forma o lema e a definição compõe a microestrutura que contém, além desses elementos, outros recursos relacionados também ao lema.

⁵ Cf. Santiago (2012, p. 5).

⁶ “[...] el lema es la parte enunciativa de un artículo, cuyo objeto es la descripción y explicación del lema.”

Como este trabalho trata de uma obra bilíngue, buscamos definir esse tipo de material lexicográfico e, principalmente, as teorias sobre os dicionários bilíngues. Opomos, inicialmente, dicionários monolíngues e plurilíngues (bilíngues, trilíngues), sendo o primeiro constituído de apenas uma língua em que os verbetes são compostos basicamente de uma palavra-entrada, sobre a qual é dada alguma informação que equivale à sua definição/ao seu significado; o segundo, por sua vez, constitui-se de mais de uma língua.

A partir disso, buscamos delimitar a conceituação dos materiais bilíngues nos quais este trabalho foca. Esses materiais lexicográficos são caracterizados por apresentarem informações sobre mais de uma língua, pois, como descrevem Coward e Grimes (2000, p. 70), “[...] concentram-se em fornecer equivalentes de tradução (aqui denominados ‘glosas’) com referência a outro idioma. A questão é fornecer informações suficientes para que o usuário saiba quais glosas são adequadas (e inadequadas) em contextos específicos”⁷. Logo, dicionários bilíngues são consultados, em geral, para encontrarmos a tradução de uma palavra em outra língua, diferente da consulta ao dicionário monolíngue em que procuramos o significado/a definição para determinada unidade lexical.

Outra questão necessária para a compreensão de lexicografia bilíngue é a noção de equivalência. Werner (1982, p. 286) diz que “[...] o dicionário bilíngue dá, no lugar da definição monolíngue do significado, equivalentes por tradução”⁸, ou seja, isso significa que a definição do dicionário bilíngue não é exatamente uma definição, mas um item lexical que equivale à palavra-entrada em outra língua. Além do mais, quando fazemos a “definição” em um material bilíngue, a metalinguagem dessa definição não faz uso de significantes da língua descrita, mas da outra língua, o que confere uma diferença pontual entre um dicionário monolíngue e um bilíngue (WERNER, 1982, p. 286).

Por fim, Haensch (1982) destaca que, em um artigo de um dicionário bilíngue, ou seja, na microestrutura, alguns recursos comumente aparecem, são eles: o enunciado do lema, as indicações fonéticas, ortográficas, gramaticais e léxicas, o equivalente na língua de destino e uma parte sintagmática e, às vezes, paradigmática. Segundo esse mesmo autor, “[...] o artigo é a menor unidade autônoma de um

⁷ “[...] focuses on providing translation equivalents (here called ‘glosses’) with reference to another language. The trick is to provide enough information so the user knows which glosses are appropriate (and inappropriate) in particular contexts”.

⁸ “[...] el diccionario bilingüe da – en lugar de la definición monolingüe del significado –, equivalentes por traducción”.

dicionário, e pode ter uma fisionomia muito variada, desde poucas palavras sem subdivisões [...] até várias colunas com uma série de divisões e subdivisões”⁹ (HAENSCH, 1982, p. 462). Embora possa agrupar todos esses recursos, uma microestrutura simples de um material bilíngue costuma conter, no mínimo, apenas duas informações: o lema e seu equivalente na língua-alvo. Portanto, apesar de fazer parte de uma estrutura maior, um artigo de dicionário é, de certa maneira, independente em relação à macroestrutura do material, uma vez que há toda uma ordenação que deve ser seguida nesse nível de organização.

Dados esses conceitos básicos, passamos às considerações sobre a tipologia do *corpus* em análise.

Algumas considerações sobre a ‘lista’

Ao considerarmos a nomenclatura adotada por Crofts (1986a) para seu material, percebemos que a ‘lista’ se trata de um tipo de material lexicográfico de características singulares, sobre o qual há poucas referências teóricas.

Para compreendermos melhor o sentido e o contexto de uso desse termo, recorreremos a formas congêneres relacionadas à tipologia de materiais lexicográficos. O primeiro elemento dessa tipologia é “dicionário”, termo amplamente conhecido, cuja definição, segundo Atkins e Rundell (2008, p. 2) é “[...] uma descrição do vocabulário usado por membros da comunidade de fala”¹⁰. Essa concepção é embasada a partir de que itens são usados na comunidade de fala para comunicação entre seus membros. Já segundo Dubois *et al.* (1993, p. 186), o dicionário é “[...] uma obra que registra certa descrição do léxico de uma língua ou de muitas línguas postas em paralelo”, ou seja, o dicionário é o material responsável pelo registro léxico de um sistema de fala, visto que o léxico preferível é o que está em uso pelos falantes. A lista, se comparada a um dicionário, tomando por base as entradas de sua macroestrutura, seria mais breve. Além de dicionários, outros termos muito difundidos são “glossário” e “vocabulário”.

As definições que mais se aproximam de uma lista são as de *glossary* e *finderlist* apresentadas por Coward e Grimes (2000, p. 67): “Um glossário geralmente não é mais do que uma listagem da palavras-

⁹ “[...] el artículo es la más pequeña unidad autónoma de un diccionario, y puede tener una fisionomía muy variada, desde pocas palabras sin subdivisiones [...] hasta varias columnas con una serie de divisiones y subdivisiones”.

¹⁰ “[...] a description of the vocabulary used by members of a speech community”.

entrada (lexemas) e uma ou duas simples glosas. [...] é um mínimo necessário para arquivar idiomas e culturas em declínio”¹¹.

Em relação ao dicionário, um *glossary* tem diferença na apresentação de suas entradas, principalmente. Já uma *finderlist*, que tem uma proximidade com *glossary*, apresenta estrutura como um *index* ou uma lista para encontrar formas equivalentes em outras línguas. Dentre outros objetivos, pode ser somente uma lista de lexemas ou formulário, sem outras informações. Ademais, pode ser encontrada em reversões de dicionários bilíngues e também como listas comparativas para línguas geneticamente aparentadas (COWARD; GRIMES, 2000, p. 67).

A partir dessas ideias, as listas de palavras são feitas com o propósito de registrar o léxico de uma comunidade linguística, geralmente de uma língua pouco conhecida, e de servir para comparar o léxico de línguas genética e possivelmente aparentadas, sendo breves em suas entradas e, com frequência, configurando o passo inicial para a documentação da língua de um povo. Além disso, uma variação desse tipo de material lexicográfico, tais como: as listas de Loukotka e de Swadesh, que tinha como propósito a comparação entre línguas, foram utilizadas em larga escala para a documentação de muitas línguas da América e contribuíram para o estabelecimento de um ramo do léxico chamado de teoria léxico-estatística¹².

Ainda nesse contexto de listas comparativas, em um estudo sobre listas de palavras e lexicografia comparativa, com base em línguas do Oeste da África, Naden (1993, p. 171) afirma que “[...] grande parte da amostragem inicial de línguas foi na forma da coleção de listas de palavras”¹³. Convém destacarmos que tais listas também foram usadas com bastante frequência no Brasil, desde a época da colonização, para registrar línguas dos povos indígenas.

Outras espécies de listas de palavras que encontramos são aquelas que foram produzidas por viajantes e etnógrafos europeus, a exemplo de Spix e Martius, que fizeram várias expedições por terras

¹¹ “A glossary is usually no more than a listing of the headword (lexeme) and a simple gloss or two. [...] A glossary is sometimes a necessary minimum for archiving dying languages and cultures”.

¹² A finalidade primeira das listas comparativas, como o próprio nome já diz, é de fazer comparações para estabelecer relações genéticas para o agrupamento de possíveis línguas de mesma família ou tronco linguístico, além de estabelecer se línguas não agrupadas em famílias constituem isolados linguísticos. Essas listas contêm entre 100 e 200 itens (cf. RODRIGUES, 1963; 1964).

¹³ “[...] much of the early sampling of languages was in the form of the collection of word lists”.

brasileiras, sendo grandes documentadores e difusores das línguas indígenas do Brasil (MARTIUS, 1867). Essas listas eram breves e apresentavam a palavra na língua indígena e em línguas europeias (alemão, francês, inglês, português, latim, dentre outras) ou somente na língua indígena, não abordando nenhum aspecto complexo na língua, apenas elencando os vocábulos. Por conseguinte, a compilação de listas de palavras no Brasil foi, muitas vezes, o resultado de trabalhos de campo de pesquisadores que entenderam, por essa forma, fazer a documentação das línguas brevemente.

Esse tipo de material para registro de língua, que possui microestrutura menor que as encontradas nos dicionários, foi amplamente utilizado no Brasil durante os séculos XVIII, XIX e XX. Aliás, com relação à prática de nomear/classificar materiais lexicográficos no Brasil, muitos autores usam a terminologia (dicionário, vocabulário, léxico, glossário, lista) muitas vezes de maneira pouco evidente. Por isso, não é difícil encontrar materiais com estrutura e extensão bem semelhantes, mas denominados diversamente um do outro.

A respeito da tipologia 'lista', no âmbito das línguas indígenas do Brasil, foram produzidas várias delas. Lopes (2014, p. 169-170) apresenta os vários títulos encontrados desse tipo de material, como: lista breve, lista comparativa, lista de animais, lista de espécies, lista de itens lexicais, lista de palavras, lista de peixes, lista de plantas, lista de substantivos, lista de Swadesh, lista de verbos, lista de vocábulos, lista geral, lista lexical, lista padrão, lista resumida, lista vocabular. Em suma, como as listas foram criadas para vários objetivos e conteúdos, não há consenso quanto ao uso dessa terminologia.

Essa discussão versa muito em torno de questões formais, no entanto, de acordo com Haensch (1982, p. 461), “[...] um dicionário ou um glossário é um extrato padronizado do conjunto do léxico existente ou de um subconjunto léxico”¹⁴, ou seja, reunimos em um material dados lexicais de uma comunidade linguística, e isso depende da padronização que damos, seja em um dicionário, um glossário, seja qualquer outro material, considerando sua organização. Embora com diferenças formais, as obras lexicográficas têm um mesmo propósito, que é o de registrar parte do léxico de uma comunidade linguística a partir de um exemplário logicamente estruturado.

¹⁴ “[...] un diccionario o un glosario es un extracto estandarizado del conjunto del léxico existente o de un subconjunto léxico”.

Listas de palavras são, finalmente, materiais que reúnem lemas (lexemas) ou são seguidos, em geral, de uma breve definição ou, o que é mais comum, uma tradução, cujo objetivo é registrar de maneira preliminar o léxico de uma comunidade de fala. Portanto, antes de um objetivo comparativo, as listas de palavras são, principalmente, meios de documentação da língua de uma comunidade, enquanto os dicionários seriam a melhor forma para descrever com mais detalhes todos os elementos necessários a respeito da língua, pois apresenta descrição mais completa sobre ela.

Esses são os conceitos que nos norteiam quanto à lista de palavras neste trabalho, haja vista que a tradição metalexicográfica brasileira pouco desenvolveu ideias sobre essa tipologia, ao contrário da atenção maior dispensada a dicionários, vocabulários e glossários.

A produção lexicográfica indígena brasileira

A lexicografia indígena produzida no Brasil teve início, grosso modo, com a publicação de um material considerado de cunho lexicográfico de autoria de Jean Corbier e compilação de Jehan Lamy no ano de 1540, e não passava de uma breve lista de palavras em duas línguas, mas é somente no século XVII que ela vai se firmar de maneira definitiva. A produção lexicográfica no Brasil existe, portanto, desde o início de povoamento por estas terras (ISQUERDO, 2011, p. 115). Inicialmente, várias listas de palavras foram produzidas por viajantes europeus, muitas vezes como resultados de expedições no território brasileiro. Os escritos produzidos por esses viajantes configuram-se também como os pioneiros na produção de materiais lexicográficos das línguas indígenas brasileiras. Esses materiais produzidos no âmbito das línguas nativas do Brasil, com variadas tipologias, têm servido de base para análises fonéticas e fonológicas, morfológicas, sintáticas e semânticas, bem como para reflexões de ordem sincrônica e diacrônica. Grande exemplo dessa leva de produtores de materiais lexicográficos é o alemão Martius, que foi grande documentador e difusor das línguas indígenas brasileiras (MARTIUS, 1867).

Além dos viajantes, várias expedições foram feitas por missionários jesuítas, que muito descreveram sobre os povos autóctones, sua cultura e seu modo de falar. Por consequência, essas descrições continham, quase sempre, listas de palavras das línguas faladas por essas populações. Segundo Gonçalves e Murakawa (2010, p. 213 *apud*

ISQUERDO, 2011, p. 117), “[...] as descrições da natureza exótica (fauna e flora) realizadas pelos missionários nos seus relatos de viagem, quando analisados à luz da lexicografia, revelam semelhanças com os produtos lexicográficos”. Apesar de esses materiais não terem sido produzidos como sendo propriamente lexicográficos, isto é, fazendo uso de teorias lexicográficas e com esse fim em si, eles aparentam proximidade quanto a esse ramo de descrição, o que configura que os materiais produzidos com essa característica são de caráter paralexicográfico. A ideia de materiais paralexicográficos é corroborada por características presentes nos materiais que são análogas a uma definição enciclopédica, pois, de acordo com o que escreve Isquerdo (2011, p. 117), os viajantes “[...] exercitavam, de certa forma, a prática da definição de palavras”, estando aproximado, desse modo, dos princípios lexicográficos.

Com o passar dos anos e o avanço nos estudos descritivos das línguas nativas do Brasil, a documentação dessas línguas mostrou-se necessária, principalmente em virtude do uso para estudos históricos comparativos. Daí a importância da documentação do léxico das várias línguas indígenas brasileiras. O século XX é visto como um marco para a produção lexicográfica no Brasil, com pesquisas de campo para descrição e análise linguística das línguas indígenas brasileiras, que começam a descrever as variadas línguas nativas do Brasil e, por consequência, registram esses dados em listas, vocabulários, glossários, dicionários, entre outros. De acordo com Isquerdo (2011, p. 120), foram produzidos “[...] muitos vocabulários bilíngues [...] pelos jesuítas como material de apoio para a catequização dos povos autóctones, etnias nativas da terra recém-descoberta pelo colonizador português”. Isso corrobora o fato de que os materiais compilados pelos missionários e jesuítas sobre as línguas nativas do Brasil muito contribuíram para a “civilização” das populações nativas, mais ainda pelo valor linguístico-cultural que carregam, por registrarem o léxico de tais comunidades.

Já com traços mais modernos, a lexicografia indígena brasileira produz, no século XX, dezenas de materiais, com ênfase em dicionários, muitas vezes resultantes de pesquisas no âmbito de dissertações e teses dos programas de pós-graduação em instituições de ensino superior, e produções de profissionais estrangeiros. Nesse contexto, vale destacarmos a atuação da organização Summer Institute of Linguistics (SIL) nas áreas indígenas, que reuniu significativa quantidade de documentação acerca de várias línguas indígenas brasileiras, usada, em geral, para fins de proselitismo religioso. Convém acrescentarmos ainda o convênio do SIL

com o Museu Nacional do Rio de Janeiro na construção do *Formulário dos Vocabulários Padrões para Estudos Comparativos Preliminares nas Línguas Indígenas Brasileiras*¹⁵, material-base que serviu de guia para a pesquisa linguística no Brasil.

No que concerne à língua Mundurukú, vários materiais lexicográficos foram produzidos, como dicionários, vocabulários e listas, dentre os quais destacamos a *Munduruku Word List*, de autoria de Marjorie Crofts, pesquisadora a serviço do SIL Internacional.

Análise do corpus: Munduruku Word List

O *corpus* analisado neste trabalho foi extraído da *Munduruku Word List*, que foi compilada por Crofts (1986a, 1986b, 1986c, 1986d), cuja versão consultada foi obtida no acervo da Biblioteca Curt Nimuendajú, da Fundação Nacional do Índio (FUNAI). Este material lexicográfico contém grande parte do léxico Mundurukú, totalizando 12.912 entradas na direção Mundurukú-Português e é de ordem semasiológica, ou seja, a organização é pelas letras do alfabeto, sendo dividida em quatro volumes: o primeiro, contendo as letras A, B, C, D, E e H; o segundo, com as letras I, J, K, L, e M; o terceiro, com as letras N, Ñ, O, P e R; e, no último, as letras S, T, U, W, X, Y; todos de acordo com a ortografia do alfabeto Mundurukú (CROFTS, 2004).

Dentre algumas singularidades em torno da produção dessa obra, destacamos a inexistência de demarcação ou divisão na mudança de entradas iniciadas por uma letra para outra dentro do mesmo volume; além de Crofts (1986a, 1986b, 1986c, 1986d) usar algumas notas para referir-se à mudança de item lexical quando da mudança de tom, mas sem a numeração que usa em Crofts (2004, 2006), por exemplo.

Percebemos, também, na *Munduruku Word List*, uma arrolagem de itens lexicais como *radicais* e *palavras* (CROFTS, 2006), que são tomados como entradas nesse material, além de palavras derivadas. Outro aspecto importante na macroestrutura da lista é o tratamento de homonímia/polissemia que é feito da mesma forma, sem diferenças na apresentação dos dois fenômenos, sendo o significado da tradução o elemento denotador de um ou outro fenômeno. Essa exposição de entradas homônimas e polissêmicas difere dos orientados em alguns métodos de tratamento lexicográfico, como, por exemplo, o de Haensch (1982).

¹⁵ Esse formulário caracterizava-se por alternar vocábulos – termos gerais e específicos – e frases da língua indígena em uma média de 321 itens.

Método de descrição da *Munduruku Word List*

Para descrição e análise das microestruturas dos verbetes da *Munduruku Word List*, usamos o modelo de descrição de microestruturas desenvolvido por Lopes (2014), que propõe modelos lexicográficos para línguas indígenas brasileiras, a partir de um esquema de descrição dos recursos utilizados nas microestruturas, em que levamos em conta todos os elementos presentes nesse tipo de estruturação interna.

Desse modo, fundamentamo-nos em Lopes (2014), que procura entender a organização da microestrutura de materiais lexicográficos produzidos para as línguas indígenas brasileiras ao longo dos anos; no caso deste trabalho, aplicamos a descrição para compreender o funcionamento da lista de palavras Mundurukú aqui analisada. Assim, o esquema visa obter o máximo da informação microestrutural dos materiais lexicográficos, considerando que “[...] toda microestrutura representa um conjunto organizado de informações dispostas intencionalmente” (LOPES, 2014, p. 160), ou seja, essa proposta considera que os componentes da microestrutura respeitam orientações e métodos em que o autor procura organizar os verbetes. Nesse sentido, a proposta de Lopes (2014) sintetiza os elementos esquematicamente para a melhor análise de determinada microestrutura.

Outro aspecto importante é que o referido trabalho está inserido no campo da lexicografia das línguas indígenas brasileiras, o que está inteiramente de acordo com os objetivos desta pesquisa. Portanto, seguindo a justificativa para tal modelo de descrição, de acordo com Lopes (2014, p. 163), uma microestrutura é composta de elementos básicos, tais como: (i) *lema*, parte que inicia o verbo e relaciona-se com os outros elementos que compõem a microestrutura; (ii) *descriptor*, parte que contém as informações metalinguísticas relacionadas ao lema; nos dicionários monolíngues, são as definições e, nos bilíngues, são os equivalentes; (iii) *operador*, parte que se associa ao lema ou ao(s) descriptor(es), pode ser a indicação de variação, pronúncia, remissões, notas, etimologias, fontes, entre outros; por fim, (iv) *exemplo*, haja vista que os materiais podem apresentar exemplos ou abonações sobre o lema ou o descriptor (grifos nossos). Com essa apresentação, o esquema de descrição de microestruturas que o pesquisador propõe se apresenta no que ele chama de *Expressão da Microestrutura Lexicográfica* (EML), que é parecida com uma expressão numérica matemática, com elementos entre chaves { }, colchetes [] e parênteses ().

A descrição das EMLs será feita da seguinte maneira, segundo o proposto por Lopes (2014, p. 164): iniciamos com o lema (L) separado do corpo do verbete por dois pontos; os operadores (O) relacionados diretamente com o lema aparecem em seguida e, caso haja mais de um, vêm separados por vírgulas e em ordem alfabética, não importando a ordem que aparece na microestrutura; após os operadores, destacamos, entre chaves, os descritores (D); os operadores associados ao descritor devem vir entre parênteses. Caso haja mais de um, são separados entre vírgulas e em ordem alfabética; por fim, fora do campo dos descritores, devem ser apresentados os exemplos, entre colchetes. A partir disso, a EML básica para a descrição das microestruturas dos materiais lexicográficos é a seguinte:

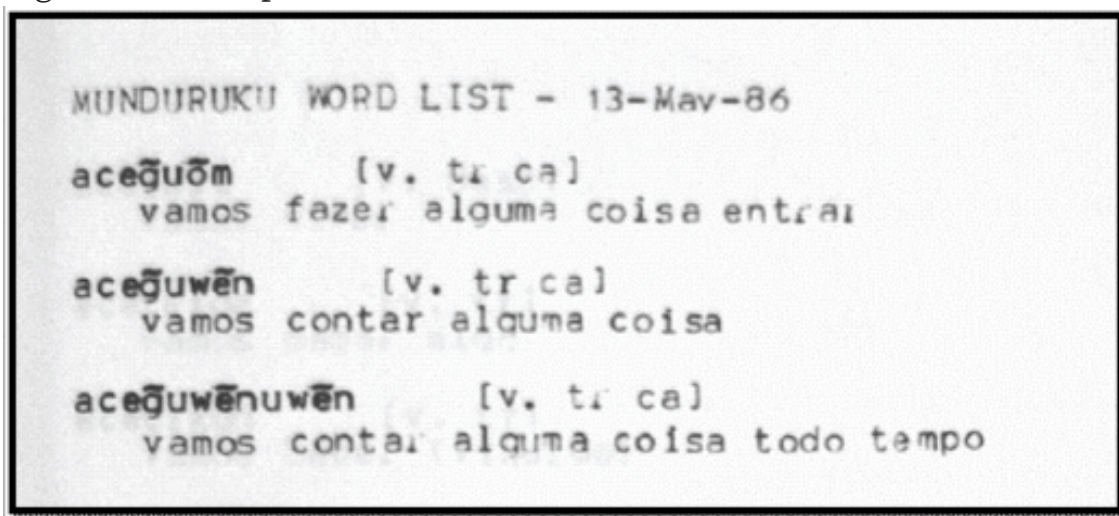
EML: L: O {D} [E]

Portanto, a descrição das microestruturas da *Munduruku Word List* é apresentada com base nesse esquema.

Descrição e análise de EMLs

454

Na análise e descrição, limitamos o *corpus* da pesquisa somente aos verbetes constantes na letra A. Essa escolha está relacionada ao fato de não podermos abarcar todas as entradas, considerando os limites formais deste trabalho e, por haver, ao todo, 12.912 entradas na *Munduruku Word List* completa (contando os quatro volumes), o que conferiria um gênero diferente para análise. Neste caso, como o conjunto da letra A contém 600 verbetes, alcançamos uma delimitação mais consciente para a obtenção de bons resultados na análise dos dados. Para análise da microestrutura do material, usamos o modelo de descrição proposto por Lopes (2014), como já exposto em seção anterior e, com base nele, enumeramos para descrição as possibilidades de microestrutura encontradas no *corpus*. A apresentação dos verbetes é coordenada da mesma maneira como constam na *Munduruku Word List* e, abaixo deles, a EML para descrição da microestrutura encontrada. Assim, antes de serem descritas, mostramos um exemplo de como constam no material.

Figura 1 – Exemplo de microestrutura da *Munduruku Word List*

Fonte: Crofts (1986a, p. 5).

A microestrutura mais básica encontrada no *corpus* é de uma palavra-entrada (lema) em Mundurukú, seguida da sua categoria gramatical e a glosa em português, que pode ser representada pelo seguinte exemplo:

[1] abikbikap [subs] cadeira¹⁶ (CROFTS, 1986a, p. 3).

Com a seguinte EML obtida: **LId: Gr {DPt}**¹⁷.

A partir desse exemplo de descrição, dá para depreendermos que a EML do verbete anterior conta com um lema na língua indígena (Mundurukú) (LId), um operador que indica categoria gramatical – neste caso, um *substantivo* (Gr) associado ao Lema –, e um Descritor em Português (DPt), que corresponde à glosa ‘cadeira’. Essa mesma microestrutura [1] encontrada na *Munduruku Word List* é a mais básica que consta no material; no entanto, há outras em que cada elemento a mais acarreta dificuldades de compreensão da ideia a ser repassada.

Em outro nível de análise proposto em Lopes (2014), esse esquema é reduzido de acordo com um agrupamento por tipologia de microestrutura, em que Lema é representado por M; Operador, por O; descritor, por D; e Exemplo/Abonação, por E, sendo logicamente alocados

¹⁶ Todos os verbetes descritos neste trabalho foram obtidos em Crofts (1986a).

¹⁷ Siglas utilizadas na descrição das microestruturas: **LId**: Lema em língua indígena, **Gr**: Indicação da categoria gramatical [Operador], **DPt**: Descritor em português, **EId**: Exemplo em Português, **EPt**: Exemplo em Língua Indígena, **Na**: Nota do autor [Operador], **Rm**: Remissiva [Operador], **TL**: Tradução literal [Operador], **Vr**: Variante/Variedade [Operador] (Cf. LOPES, 2014).

segundo sua ordem apresentada no verbete. Portanto, a microestrutura mencionada anteriormente poderia ser descrita também como MOD, mas esse modo de descrição, menos detalhado, não será considerado neste trabalho.

A seguir, apresentamos amostras de cada tipo de microestrutura encontrado no conjunto de verbetes iniciados pela letra A da Lista.

[2] abi [subs] o fim de.../o começo de... (CROFTS, 1986a, p. 3).
EML: **LId: Gr {DPt (Vr)}**.

No dado [2], a EML é composta de Lema em língua indígena (LId), seguido de apenas um Operador na forma de Categoria gramatical (Gr). Já o Descritor, que apresenta glosa em Português (DPt), possui um registro de Variação (Vr), ou seja, há dois sentidos equivalentes em Português para somente um elemento na língua Mundurukú¹⁸.

[3] aceap [v. intr] vamos todos morrer (num dia só) (CROFTS, 1986a, p. 4).
EML: **LId: Gr {DPt (Na)}**.

O verbete [3] constitui EML composta por LId com a sua Gr e o Descritor em Português. Ao Descritor está relacionada, nessa microestrutura, uma Nota do autor (Na), disposta entre parênteses após DPt. Neste caso, o verbete contém uma nota do autor destacada para a glosa em português, uma vez que o sentido da glosa está de acordo com o que a nota destaca. As notas no verbete servem para complementar um sentido que falta para que a definição/tradução (ou outro elemento) seja bem compreendida. Por isso, existem verbetes no *corpus* que têm notas para outros elementos, não somente para o descritor, como na microestrutura acima, mas, referente ao lema em muitos casos.

[4] abuat [perg] quem é que...? (Ex: abuat oe'ũ? Quem é que morreu?) (CROFTS, 1986a, p. 3).
EML: **LId: Gr {DPt}{EId/EPt}**.

¹⁸ Nesse descritor duplo, as duas equivalências constituem formas antônimas em português, correspondendo a uma glosa somente em Mundurukú. (cf. “**abibimap** [subs] de começar/o começo” em que a glosa constitui sinônimos em português) (CROFTS, 1986a).

A microestrutura [4] agrega um LIId, seguido da Categoria gramatical (Gr) que se insere e, após esse operador, da glosa referente a esse Lema em Português (DPt). Além disso, no final desse verbete, está o elemento Exemplo em Português e na língua indígena, mostrando o uso ou abonação da palavra arrolada.

Crofts (1986a) faz uso de um exemplo para mostrar o uso do Lema. Mais ainda, depreendemos que as reticências representam, nos descritores da microestrutura dessa lista, uma lacuna para algum complemento da estrutura sintática que já contém, em si, o significado ou equivalência do lema, como, por exemplo, um núcleo verbal da oração subordinada: ‘Quem é que *morreu*?’, conforme o exemplo citado no verbete.

Apesar de ter papel crucial em um verbete, o elemento Exemplo raras vezes aparece nas microestruturas analisadas da *Munduruku Word List*. Sendo assim, os casos em que há Exemplo nas microestruturas são em pequena quantidade, salvo casos como [4] e [5], visto que o contexto de uso é necessário para tirar qualquer dúvida quanto ao significado do lema.

[5] ãg [part] deixa... (p. ex: ãg ieenũy. Deixa ele estar./Não mexe com ele) (CROFTS, 1986a, p. 32).

EML: **LIId: Gr {DPt [EId/EPt (Vr)]}**.

Esta microestrutura, além dos mesmos elementos da microestrutura anterior — isto é, LIId: Gr {DPt}[EId/EPt] —, apresenta Variação (Vr) no elemento Exemplo em Português. Esse elemento Vr não é exatamente uma abonação quanto ao uso do equivalente, porém, apresenta-se como uma variação de forma/sentido do enunciado-base do Exemplo traduzido. A abonação referente ao Lema é somente o elemento EPt, embora sua variação seja uma extensão de sentido metafórico.

[6] abu [perg] quem? (singular) (CROFTS, 1986a, p. 3).

EML: **LIId: Gr, Na {DPt}**.

Nesta sexta amostra de microestrutura, a EML é composta de LIId com informação gramatical (Gr) e um Descritor em Português como equivalente (DPt), seguido de uma Nota do autor (Na) ao final do artigo. Na descrição esquemática da microestrutura, a nota do autor está disposta ao lado do operador, significando relação direta com o lema em Mundurukú.

Visualmente, nada indica a que se refere a nota entre parênteses. Apesar de essa nota vir ao final do artigo, ela está relacionada ao Lema, visto que não existe flexão do pronome ‘quem’ no Português, existindo, portanto, somente no singular. Resta-nos inferir que essa partícula interrogativa ‘abu’, quando usada no singular, em Mundurukú, corresponde à glosa “quem?”, no Português, também como partícula interrogativa. Vale ressaltarmos a importância da nota do autor nesse verbete para destacar elementos gramaticais e lexicais próprios da língua Mundurukú¹⁹.

[7] acejo’o [v. tr] vamos comer algo (Cf: [v. je] acekõn. vamos comer) (CROFTS, 1986a, p. 5).

EML: **LId: Gr, Rm {DPt}**.

Este verbete é composto de Lema em língua indígena (LId), Operador gramatical (Gr) e Descritor em Português (DPt), além de uma Remissiva (Rm), cuja ocorrência dentro do verbete remete a palavra à outra diretamente relacionada, podemos, nesse caso, dizer que o elemento remissivo tem relação com a microestrutura inteira. Esse tipo de remissão, segundo Wiegand (*apud* WELKER, 2004, p. 179), é “facultativa”, pois não conduz necessariamente o consulente ao lema a que remete.

Em outro contexto, para que façamos uso de uma remissão como na microestrutura anterior, Welker aponta que devam existir lexemas relacionados semanticamente, como sinônimos, antônimos, por exemplo (WELKER, 2004, p. 179). Desse modo, ao fazermos a remissão a outro verbete, como em [7], remetemos não só à palavra relacionada, mas a toda a microestrutura em que aparece.

[8] acetip’uk [v. tr] vamos capinar (lit: tirar o capim) (CROFTS, 1986a, p. 8).

EML: **LId: Gr [Tl]{DPt}**.

O verbete [8] apresenta organização estrutural com lema em língua indígena (LId), sua categoria gramatical (Gr) e a glosa em Português (DPt). Essa microestrutura apresenta ainda uma Tradução literal (Tl) para a glosa em Português.

¹⁹ Cf. “**abuyũ** [perg] quem? (plural)” (CROFTS, 1986a) com nota ressaltando que o lema está no plural, mas com o mesmo equivalente em Português.

Isso acontece porque, na maioria das vezes, na tradução de uma língua para outra, como uma língua nativa brasileira e o português, ocorre a adaptação da tradução, pois a tradução literal de uma palavra ou sentença não tem um valor semântico que caiba na língua-alvo, como é o caso da microestrutura anterior. Logo, para qualquer falante do Português, a sentença “tirar o capim”, que é a tradução literal para o Lema, não expressa completamente o sentido que costumamos repassar, que é o ato de capinar; esse Lema é melhor traduzido por “vamos capinar”, mais aceito e falado no Brasil.

Outro elemento que, assim como o Descritor, tem variantes no *corpus* analisado é o Lema. Neste material, aparece separado por barras para expressar variação. Vide exemplo a seguir:

[9] agõwakaiceh/agõwakaiceeh [voc] uma personagem numa história antiga (CROFTS, 1986a, p. 21).

EML: **LId: Gr, Vr {DPt}**.

A EML obtida a partir dessa microestrutura é composta de Lema em língua indígena (LId) e a variante do lema (Vr), seguida da categoria gramatical (Gr) e a tradução para o Português (DPt). Nessa EML, o elemento (Vr) está fazendo referência ao lema, por isso a disposição ao lado do operador (Gr), separados por vírgulas.

A presença do elemento Vr nesse verbete expressa, segundo Atkins e Rundell (2008, p. 206), “[...] uma grafia alternativa ou ligeira variação na forma desta palavra”²⁰, sendo que a forma variante, muitas vezes, é uma pequena variação do Lema, que pode ser a forma mais conhecida que outras. Vale ressaltarmos que a forma variante não implica em mudança na tradução, diferente de mudanças na forma da palavra com flexão e derivação, por exemplo, que geram novos sentidos à palavra; neste caso, o que ocorre é uma variação na pronúncia e, por consequência, na grafia.

[10] aceḡuymũḡ [v. tr ca] vamos fazer alguém calar/vamos fazer alguém encostar (depende do tom) (CROFTS, 1986a, p. 7).

EML: **LId: Gr, Na {DPt (Vr)}**.

²⁰ “[...] an alternative spelling or slight variation in the form of this word”.

O verbete acima apresenta uma microestrutura com o lema na língua indígena (LI_d) e um operador como Categoria gramatical (Gr), seguido do Descritor a que corresponde no Português. Nessa microestrutura, encontramos o Equivalente no Português (DP_t) e uma forma variante (Vr), além de uma Nota do autor (Na) ao final do verbete, fazendo referência ao tom, presente na língua Mundurukú. Temos, nesta EML, o uso de variante referente ao descritor e nota referente ao lema, apesar de a nota estar ao final do artigo. Nesse contexto, a variante apresentada destaca uma importância a mais para um consulente falante do Português do que a um falante Mundurukú, embora sirva para falantes de ambas as línguas. Já no caso da nota do autor, a informação sobre a tonalidade na pronúncia da palavra em Mundurukú é relevante, entretanto, em se tratando de fala, é mais útil a um consulente Mundurukú.

Mesmo que não conheçamos por completo as questões fonéticas da língua Mundurukú, uma questão complexa é até que ponto o tom interfere na variante da equivalência em Português, porque, nesse verbete, as glosas têm propriedades semânticas distintas (a saber: calar/encostar). O que inferimos é que a mudança de tom é responsável por essa variação, pois, indo mais a fundo na estrutura da língua, Picanço (2005, p. 311) expõe que o idioma Mundurukú “[...] é uma língua tonal, isto é, faz uso de variações de tom para diferenciar lexicalmente os itens”²¹. De fato, a presença distintiva do tom²² nessa língua é elemento pontual para a mudança de unidade lexical na microestrutura anterior, embora não esteja claro onde está a diferença no tom. Em contrapartida, Crofts (1986a, 1986b, 1986c, 1986d) não faz marcação de tom na própria palavra, como o faz em outros trabalhos (CROFTS, 2004; 2006), mas sim, por meio de nota ao final do verbete, apenas fazendo referência quanto ao seu traço distintivo. Esse tipo de escolha fez mais sentido para a teoria lexicográfica usada pela compiladora, que procurou ser clara e precisa na arrolagem do lema, embora tenha exposto com brevidade as informações, uma vez que uma numeração dos quatro níveis de tom presentes na língua – dois, para Picanço (2002, 2005) –, geraria uma nova entrada para cada item lexical ou duas palavras-entrada. Em consequência disso, os descritores precisariam ser especificados

²¹ “[...] is a tonal language; that is, it makes use of pitch variations to lexically differentiate items”.

²² No conjunto das línguas indígenas brasileiras, destacamos como línguas tonais – para as quais essa reflexão também pode ser útil – o Juruna (Cf. FARGETTI, 2001), o Guató (Cf. PALÁCIO, 1984) e o Tikuna (Cf. SOARES, 2001).

para cada lema a que se referem, porque, sem isso, causaria dúvidas quanto aos seus respectivos lemas, especificados pela tonalidade; ou, ainda, as palavras poderiam ser entendidas como sendo homônimas (homógrafas). Entretanto, essas discussões, que são complexas, não foram levadas em consideração na *Munduruku Word List*, apesar de não sabermos se foi com base nisso a escolha para o não uso de marcação de tom no referido material.

Outro tipo de microestrutura presente na *Munduruku Word List* é quando a variação existe tanto no Lema quanto no Descritor, ao contrário da anterior. Vejamos o exemplo a seguir:

[11] ak’adit/ak’adit’ũma [v. desc] não responde/não tem interesse em... (CROFTS, 1986a, p. 15).

EML: **LId: Gr, Vr {DPt (Vr)}**.

Overbete [11] apresenta EML composta de Lema em Mundurukú (LId) e uma Variante (Vr), operador gramatical (Gr), seguidos da glosa em Português e outra variante. Neste caso, há uma dupla variação no lema e no descritor. A variação no lema é expressa pela adição do sufixo *’ũma*, que é o elemento sufixal privativo em Mundurukú, mas isso se refere a um sistema interno de variação lexical na língua Mundurukú, não está relacionado, necessariamente, a que cada variante da língua-fonte corresponde, respectivamente, à variante da língua-alvo.

A variação no lema é uma variação apenas na forma da palavra, enquanto que no descritor ocorre no nível semântico, referindo-se a somente um item lexical na língua-fonte. Apesar dessa “dupla variação”, ou seja, no lema e na equivalência, essas variações não têm, necessariamente, relação respectiva a cada lema, ocorrendo a primeira somente na língua Mundurukú e a segunda sendo uma tradução opcional para o lema em questão. A complexidade de entendimento desse verbete esbarra na implicação dos elementos variantes, que ocorrem no lema e no descritor. Além do mais, não há sequer um texto introdutório no material sobre o arrolamento de palavras em casos como esse.

[12] a’õkukuḡ [v. tr] discutindo (lit: mexendo com a voz)/imitando a voz/imitando o cântico (dum pássaro). (CROFTS, 1986a, p. 18).

EML: **LId: Gr, [Tl] {DPt (Vr, Na)}**.

Na EML obtida a partir dessa microestrutura, constam um Lema na língua indígena (LId), a Categoria gramatical (Gr) e o Descritor em Português (DPt). Este, por sua vez, apresenta dois elementos subordinados, que é uma Variante (Vr) e uma Nota do autor (Na). Além disso, há uma tradução literal da glosa em Português (Tl).

Em outro aspecto, percebemos múltiplos significados para o lema na língua Mundurukú, por exemplo, a especificação “dum pássaro” (Na) fez-se necessária por se tratar de uma glosa que é mais distante semanticamente das outras, já que se referem à imitação de voz, porém, a última refere-se a pássaro, especificando a imitação de cântico do pássaro. Esta última variante da glosa contém a nota do autor para especificar o seu significado para não falantes de Mundurukú, já a primeira glosa é seguida também de uma tradução literal, a qual está diretamente relacionada com os outros elementos variantes da glosa. A partir dessas análises, foram encontradas as seguintes microestruturas com a letra A na lista.

Quadro 1 – Tipos de microestruturas encontradas nas entradas iniciadas pela letra A da *Munduruku Word List*

	Microestrutura (EML)
1	LId: Gr {DPt}
2	LId: Gr {DPt (Vr)}
3	LId: Gr {DPt (Na)}
4	LId: Gr {DPt}[EId/EPt]
5	LId: Gr {DPt}[EId/EPt (Vr)]
6	LId: Gr, Na {DPt}
7	LId: Gr, Rm {DPt}
8	LId: Gr, Tl {DPt}
9	LId: Gr, Vr {DPt}
10	LId: Gr, Na {DPt (Vr)}
11	LId: Gr, Vr {DPt (Vr)}
12	LId: Gr, Tl {DPt (Vr, Na)}

Fonte: Autoria própria.

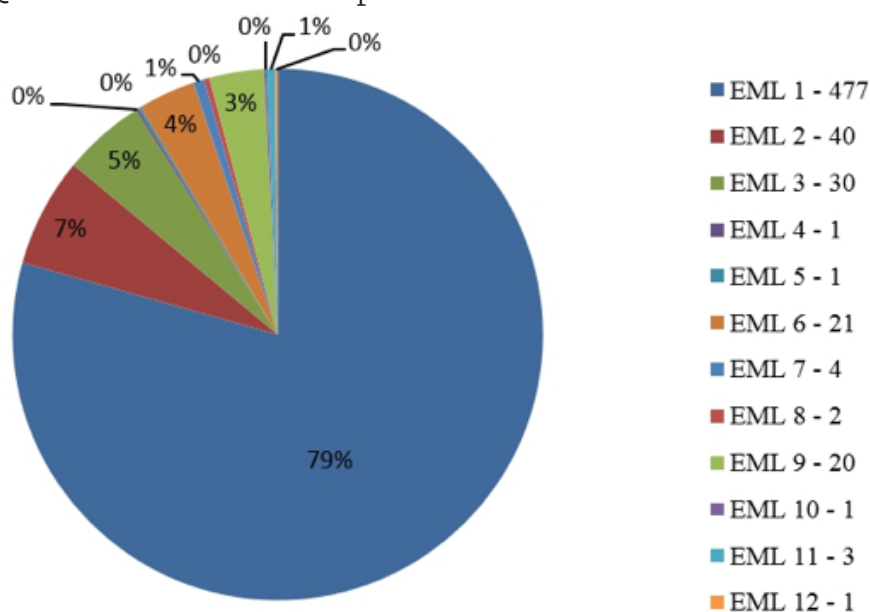
O levantamento de tipos de microestruturas leva em conta as diferenças que acarretam mudanças na descrição das EMLs. Por isso, quando ocorreu mais de uma variação no lema ou no descritor, estas

não foram consideradas, em virtude de a EML não admitir descrição de mais de uma variação em um mesmo elemento. Por exemplo, existem verbetes cujos lemas ou descritores contêm mais de duas possibilidades de variação que não cabem nesse nível de descrição. Ao todo, segundo essa ideia, foram depreendidas 12 microestruturas.

Após a amostragem dos dados, percebemos que, dentre as microestruturas descritas anteriormente, ocorreu variação em vários elementos, como o caso das entradas [2], [5] e [9], que apresentaram variação no lema, descritor e exemplo, respectivamente.

Abaixo, mostramos, em gráfico, a quantidade de entradas para cada tipo de microestrutura.

Gráfico 1 – Quantidade de entradas para cada EML na *Munduruku Word List*



Fonte: autoria própria.

Esse gráfico expressa a predominância, em muito, da EML 1 na *Munduruku Word List* com 79% dos dados arrolados nessa microestrutura. Do contrário, as EMLs [4], [5], [10] e [12] apresentam apenas uma microestrutura cada. Já outras têm quantidade considerável de microestruturas, apresentando, assim, 40, 30, 21 e 20 entradas, respectivamente, as EMLs [2], [3], [6] e [9].

Em se tratando de métodos (escolhas) utilizados por Crofts (1986a, 1986b, 1986c, 1986d), destacamos as notas, em sua maioria ao final dos verbetes, as quais são apresentadas de várias maneiras. Uma delas refere-se à mudança de unidade lexical (lema), acarretada

pela mudança de altura no tom (se alto ou baixo, principalmente). Crofts (1986a) expõe essas mudanças de tom apenas por nota, fazendo referência à implicação deste na ortoépia do lema e, portanto, sua distintividade na língua Mundurukú. Além disso, existem notas referentes aos descritores, para especificar alguma informação que interfira diretamente na compreensão ao consulente.

No campo onde consta a informação gramatical na *Munduruku Word List*, esse recurso aparece entre colchetes e apresenta informação diversificada acerca da gramática Mundurukú; exemplo disso é a distinção entre nome e substantivo dada nesse material, que são expostos com alguma diferença, estando apresentados como “nome” e “subs”, respectivamente. No entanto, os especificados como “nome” correspondem a nomes próprios e não necessariamente à categoria gramatical de que fazem parte substantivo e adjetivo. Outros tipos de informação gramatical aparentes na Lista são os verbais, tais como: os verbos transitivos (“v. tr”), verbos transitivos causativos (“v. tr ca”), intransitivos (“v. intr”), verbos descritivos (“v. desc”) e a voz média verbal “je-”²³, que aparece como (“v. je”) no campo de informação gramatical.

Todas as informações gramaticais obtidas a partir das microestruturas da *Munduruku Word List* (verbetes iniciados pela letra A) são as seguintes²⁴:

²³ Para mais informações sobre a voz média em Mundurukú, recomendamos a leitura de Gomes (2007, p. 317-323).

²⁴ Inserimos o Lema e Descritor para expor o contexto em que aparece a informação gramatical.

Quadro 2 – Os vários tipos de informação gramatical (Operador) presentes na *Munduruku Word List*²⁵

Informação Gramatical	Abreviatura	Lema	Descritor
Advérbio	adv	aceôm	vamos entrar
Aspecto	asp	adi	num dia
Canto	cant	awaycekũrũbububu	parte do cântico do Daydo
Exclamação	excl	acõõõõh	não há
Ideofone	ideo	adabot	estava dobrando
Locative	loc	akiju	terreiro da casa
Nome	nome	abraão'ũm	o finado Abraão
numeral	num	adesũ	mais ou menos muitos
partícula	part	acã	Somente
pergunta	perg	apoce	aonde?
pergunta/resposta	perg/resp	aaaah	o que?
relativo/relacional	rel	ase	em cima de
resposta	resp	asũmẽn	e assim
substantivo	subs	adeay'ũm	os que serão muitos
verbo descritivo	v. desc	adaodao	você é doido
verbo existencial	v. exist	a'a'e	há alguma coisa redonda
verbo intransitivo	v. intr	aceôm	vamos entrar
voz “je”	v. je	acegẽbu	vamo-nos lembrar
voz reflexiva	v. refl	acewawẽ	vamos aconselhar-nos
verbo transitivo	v. tr	atikõn	bebemos água
verbo transitivo causativo	v. tr ca	acegu'um	vamos destruir alguma coisa
vocativo	voc	aõokayũũh	homens!

Fonte: Crofts (1986a).

²⁵ Há três verbetes em que o Operador apresenta dupla informação gramatical, a saber, “perg/resp”. Neste caso, consideramos as informações separadas para quantificação no Quadro 2.

O quadro acima evidencia os diferentes tipos de informação gramatical (Operador) presentes na *Munduruku Word List*. Esse Operador aparece sempre abreviado, porém sem uma lista de abreviaturas no início do material para esclarecer o que significa cada palavra abreviada. Dessa forma, foram obtidos 22 tipos diferentes de informação gramatical, alguns em maior quantidade como os substantivos e verbos transitivos e intransitivos.

Conclusão

Diante das descrições da *Munduruku Word List*, bem como dos elementos que as compõem, percebemos que sua estruturação não é tão óbvia, nem tão fácil de ser apreendida, principalmente suas microestruturas. Constatamos que a macroestrutura, de base semasiológica, apresenta uniformidade no modo como dispõe os verbetes, estando estes sempre com o lema em negrito e verbetes ordenados segundo critério bem definido.

Já os vários verbetes, que foram descritos no decorrer deste artigo, mostram, a partir das discussões, que há uma variedade de microestruturas, com diferenças na arrolagem e com a presença de fenômenos diversos, visto que a apresentação microestrutural pouco consegue ser uniforme. De fato, segundo as descrições, foram encontradas na *Lista* doze EMLs diferentes, com predominância da EML 1. Além das microestruturas mais básicas, foi possível identificarmos elementos como variação no Lema e no Descritor e Notas do autor.

A presença das variações é um traço importante e implica na compreensão do verbete. Ressaltamos que as variações acontecem tanto no Lema quanto no Descritor, mas não pudemos descrever se, em duplas variações, no caso do Lema e Descritor, elas correspondem aos respectivos elementos variantes. Já no caso de variação somente no Descritor, mas com nota destacando a mudança de unidade lexical implicada pelo tom, não conseguimos detectar qual seria a sílaba que implicava o alteamento ou abaixamento de tom, já que Crofts (1986a) não faz marcação de tom neste material dentro do lema. Pode haver certa dificuldade para o consulente do material em compreender as microestruturas mais complexas, principalmente desambiguar as duplas equivalências referentes às entradas cujas notas remetem à mudança de tom, que é elemento distintivo para mudança de unidade lexical.

Quanto à tipologia do material, esta não implica no conteúdo, pois, embora tenha o título de lista, apresenta diversificada e extensa descrição do léxico da língua Mundurukú, sendo o material com maior quantidade de entradas produzido no âmbito dessa língua.

Em suma, as análises do tratamento lexicográfico na *Munduruku Word List*, apesar de não terem sido extensas, a ponto de abarcar todas as letras e todos os volumes do material, demonstram a diversidade lexical dessa lista e permitem compreender as principais estratégias e formas usadas na construção das microestruturas, evidenciando que se trata de um material construído a partir de uma perspectiva teórica contemporânea, apesar de não apresentá-la explicitamente.

Referências

ATKINS, Beryl T.; RUNDELL, Michael. **The Oxford Guide to Practical Lexicography**. New York: Oxford University Press, 2008.

BARROS, Maria Cândida Drumond Mendes. O contexto político e intelectual da entrada do Summer Institute of Linguistics na América Latina (1930-1960). **Revista Internacional de Linguística Iberoamericana (RILI)**, n. 4, p. 149-210, 2004. Disponível em: <<https://dialnet.unirioja.es/ejemplar/121122>>. Acesso em: 15 dez. 2017.

BIDERMAN, Maria Tereza Camargo. **Teoria Linguística: teoria lexical e linguística computacional**. 2. ed. São Paulo: Martins Fontes, 2001.

COWARD, David; GRIMES, Charles. **Making Dictionaries: a guide to lexicography and the Multi-Dictionary Formatter**. Waxhaw, North Carolina: SIL International, 2000.

CROFTS, Marjorie. **Munduruku Word List**. [s.l.], 1986a. Ms. V. 1.

_____. **Munduruku Word List**. [s.l.], 1986b. Ms. V. 2.

_____. **Munduruku Word List**. [s.l.], 1986c. Ms. V. 3.

_____. **Munduruku Word List**. [s.l.], 1986d. Ms. V. 4.

_____. **Aspectos da língua Munduruku**. Cuiabá: Sociedade Internacional de Linguística, 2004.

_____. **Gramática Munduruku**. Tradução de Mary I. Daniel. Cuiabá: Sociedade Internacional de Linguística, 2006.

DUBOIS *et al.* **Dicionário de Linguística**. São Paulo: Cultrix, 1993.

FARGETTI, Cristina Martins. **Estudos fonológico e morfossintático da língua Juruna**. 2001. 317 f. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, SP, 2001.

GOMES, Dionei Moreira. Voz média em Mundurukú: uma análise do morfema *je-(ɔ̃e-)*. In: RODRIGUES, Aryon D.; CABRAL, Ana S. A. C. (Orgs.). **Línguas e culturas Tupi**. Campinas, SP: Curt Nimuendajú; Brasília: LALI/UNB, 2007. v. 1. p. 317-323.

HAENSCH, Günther; WOLF, Lothar; ETTINGER, Stefan; WERNER, Reinhold. **La lexicografía**: de la lingüística teórica a la lexicografía práctica. Madrid: Gredos, 1982.

ISQUERDO, Aparecida Negri. Os estudos lexicográficos no Brasil: um percurso histórico. In: CARDOSO, Suzana; MEJRI, Salah; MOTA, Jacyra (Orgs.). **Os dicionários**: fontes, métodos, novas tecnologias. Salvador: Vento Leste, 2011. p. 113-140.

LOPES, Jorge Domingues. **Uma interface da documentação linguística e modelos lexicográficos para línguas indígenas brasileiras**: uma proposta para o Suruí-Aikewára. 2014. 599 f. Tese (Doutorado em Linguística) - Programa de Pós-Graduação em Linguística, Instituto de Letras, Universidade de Brasília, Brasília, DF, 2014.

MARTIUS, Carl Friederich Phillip Von. **Lingua Tupi**: Dialecti varie. Dialectus vulgaris. Der gemeine Dialekt, oder die Mundrucûs. 1867. p. 18-20.

PICANÇO, Gessiane Lobato. O sistema tonal de Mundurukú revisitado. In: RODRIGUES, A. D.; CABRAL, A. S. (Eds.). Línguas Indígenas Brasileiras: Fonologia, Gramática e História. ENCONTRO INTERNACIONAL DO GT DE LÍNGUAS INDÍGENAS DA ANPOLL, 1. **Anais...** Belém: Editora da UFPA, 2002. v. 1. p. 243-253.

_____. **Mundurukú**: phonetics, phonology, synchrony, diachrony. 2005. 424 f. Thesis. Doctor of Philosophy-Linguistics. The University of British Columbia. 2005.

NADEN, Tony. From Wordlist to Comparative Lexicography: The Lexinotes. **Lexikos**, v. 3, p. 167-190, 1993. Disponível em: <<http://dx.doi.org/10.5788/3-1-1105>>. Acesso em: 16 mai. 2017.

PALÁCIO, Adair Pimentel. **Guató**: a língua dos índios canoeiros do rio Paraguai. 1984. 155 f. Tese (Doutorado em Linguística) - Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas-SP, 1984.

RODRIGUES, Aryon Dall'Igna. Os estudos de Lingüística Indígena no Brasil. **Revista de Antropologia**, v. 2, n. 1-2, p. 9-21, jun./dez., 1963.

_____. A classificação do tronco lingüístico tupí. **Revista de Antropologia**, v. 12, n. 1-2, jun./dez., 1964.

_____. **Línguas Ameríndias**. In: GRANDE ENCICLOPÉDIA DELTA LAROUSSE. Rio de Janeiro: Delta, 1970. p. 4034-4036.

SANTIAGO, Márcio Sales. Análises contrastivas de microestruturas em dicionários escolares. **Pesquisas em Discurso Pedagógico**, Rio de Janeiro, n. 1, p. 1-14, jan./jun. 2012.

SOARES, Marília Facó. Subespecificação tonal e tom default: o caso Tikuna. *In*: CABRAL, Ana Suelly Arruda Cabral; RODRIGUES, Aryon Dall'Igna (Orgs.). **Estudos sobre Línguas Indígenas I**. Belém: Universidade Federal do Pará, 2001. p. 9-35.

WELKER, Herbert Andreas. **Dicionários**: uma pequena introdução à lexicografia. 2. ed. Brasília, DF: Thesaurus, 2004.

WERNER, Reinhold. Léxico y teoría general del lenguaje. *In*: HAENSCH, Günther; WOLF, Lothar; ETTINGER, Stefan; WERNER, Reinhold. **La lexicografía**: de la lingüística teórica a la lexicografía práctica. Madrid: Gredos, 1982. p. 21-94.